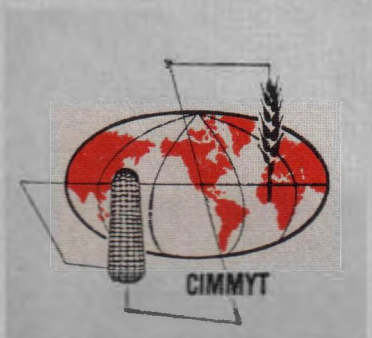
Biotechnology Applications to Plant Breeding

a workshop offered by the
Applied Biotechnology Center
to CIMMYT Staff



CENTRO INTERNACIONAL DE MEJORAMIENTO DE MAIZ Y TRIGO

INTERNATIONAL MAIZE AND WHEAT IMPROVEMENT CENTER

México

Biotechnology Applications to Plant Breeding

a workshop offered by the
Applied Biotechnology Center
to CIMMYT Staff

9 - 18 December 1997
CIMMYT Headquarters, El Batán, Mexico

The Rockefeller Foundation
International Program on Rice Biotechnology

Introduction to RFLP Mapping and Plant Breeding Applications

Prepared by
Gary Kochert
Department of Botany
University of Georgia
Athens, GA 30602

Introduction

In recent years gene cloning technology has led to fundamental advances in many areas of biology. One of the most exciting of these is the use of cloned fragments of chromosomal DNA as genetic markers, usually termed "RFLP mapping", with RFLP denoting Restriction Fragment Length Polymorphism. In this technique, which depends on natural variation in DNA base sequence, DNA is digested with a restriction enzyme. Homologous restriction fragments of DNA which differ in size, or "length", can be used as genetic markers to follow chromosome segments through genetic crosses. This new technology promises to revolutionize some areas of plant genetics and plant breeding; the present publication is intended to briefly explain RFLP analysis and some of its applications in plant breeding. A listing of current research literature is available for further reading.

What is an RFLP?

The genetic information which makes up the genes of higher plants is stored in the DNA sequences of the nuclear chromosomes and the organelle genomes. Plants are able to replicate their DNA with high accuracy and rapidity, but many mechanisms are operative which cause changes in the DNA. Simple base pair changes may occur, or larger scale changes as a result of inversions, translocations, deletions or transpositions. There is such an enormous amount of

DNA Terminology

DNA is a giant molecule, far bigger than any other molecule in the cell. DNA amounts are usually expressed as the absolute amount per cell (in picograms) or the number of base pairs (often expressed as thousands of base pairs, kb, or millions of base pairs mb). Since DNA molecules are all quite similar (they all consist of alternating A-T or G-C nucleotide pairs), it is easy to convert from one form of measurement to another using the following relationship:

One picogram = 0.965×10^9 bp = 6.1×10^{11} dalton = 29 cm

It can be seen from this relationship that human cells, which contain about 3×10^9 base pairs per haploid genome, each contain 1.8 meters of DNA! Plants show a great deal of variation in DNA content; the table below lists DNA content for several organisms and organelles.

DNA Content of Some Representative Organisms and Organelles

	DNA content (per genome)	
Organism	picograms	kilobase pairs
E. coli	.0047	4.2×10^3
Chloroplast (*Zea mays*)	.0002	1.6×10^2
Mitochondrion (*Zea mays*)	.0007	5.7×10^2
Arabidopsis thaliana	0.07	7.0×10^4
Oryza sativa	0.6	5.8×10^5
Lycopersicon esculentum	0.7	7.1×10^5
Zea mays	7.5	7.2×10^6
Homo sapiens	3.2	3.9×10^6

Restriction Enzymes

Restriction enzymes are a class of DNA endonucleases which are found in various microorganisms. Type II restriction enzymes, which are able to cleave DNA at sites where a specific base sequence (restriction site) occurs, are the type used most often in molecular biology. Restriction enzymes are thought to have evolved as a defense mechanism against foreign DNA, such as that introduced into a bacterial cell by a bacteriophage. Such exogenous DNA is rapidly cleaved into smaller fragments by restriction enzymes and the bacterium is thus protected from viral infection. Bacteria protect their own DNA from being degraded by attaching methyl groups to nucleotides in potentially sensitive restriction sites.

The first three letters in the name of a restriction enzyme generally denote the genus and species of the organism in which the enzyme was first characterized. Subsequent letters and numbers denote the specific strain of the organism or the serial order in which restriction enzymes were characterized. For example:

- *Pst*I - the first restriction enzyme characterized from *Providencia stuartii*
- *Eco*RII - the second restriction enzyme from *Escherichia coli* strain R245

Restriction sites for various restriction enzymes vary from 4 to 8 base pairs length, and are usually palindromic. The restriction site for *Pst*I is thus:

```
      ↓
CTGCAG
GACGTC
      ↑
```

(The arrows denote the site where the phosphodiester backbone of the DNA is cleaved.)

DNA in a higher plant cell, that no two organisms are likely to be identical in DNA base sequence. Thus there is a tremendous amount of DNA variation present in natural populations of plants, but until recently there has been no direct way to utilize this variation in plant genetics.

Natural variation in DNA sequence can be detected in several ways. One way, of course, is to directly sequence the DNA and make detailed comparisons. Unfortunately, this method is still very cumbersome and time consuming. Another way to access this variation is by using a special class of enzymes called restriction enzymes. These are nucleases produced by a variety of microorganisms, and they have the ability to recognize target sites (called restriction sites) made up of specific base sequences in the DNA. If the required base sequence is present in the target DNA, the restriction enzyme will cleave the DNA at the target site. A large piece of DNA will thus be reduced to a series of smaller fragments of defined size by digestion with a restriction enzyme. The number of fragments produced and the sizes of each fragment will

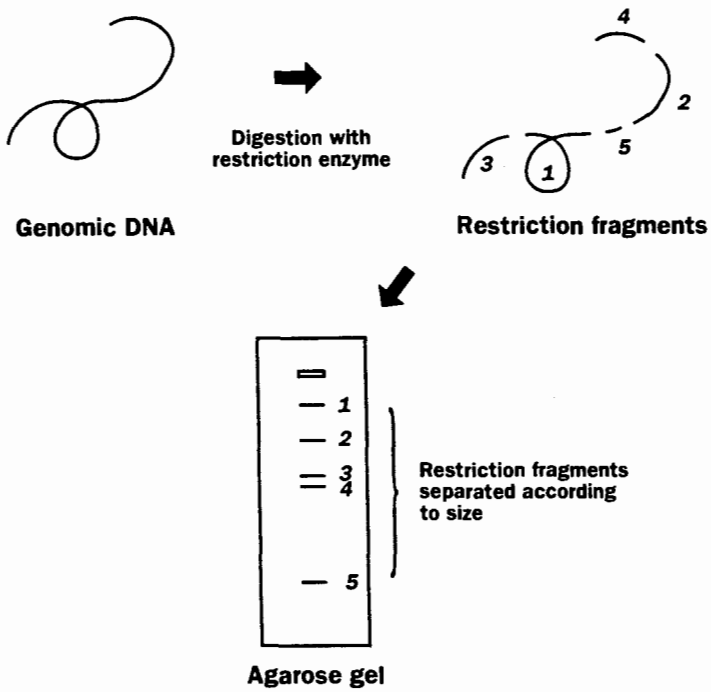
Plant Genome Organization

Plant chromosomes contain a large amount of DNA, and its organization is very complicated. One such complication involves the copy number of individual sequences. Early in the study of eucaryotic genomes it was discovered that DNA sequences may be present in virtually any number from one copy to millions of copies per haploid genome. Such a repeated DNA sequence may be present in tandem copies at one chromosomal locus, or may be dispersed thoughout the genome. Repeated DNA sequences are often observed to be concentrated in centromeres, telomeres, or other heterochromatic parts of chromosomes, and may thus have some structural role.

Most RFLP research projects use single copy sequences as hybridization probes, because markers specific to single chromosome segments are desired. It is therefore necessary to screen probe libraries to remove repeated DNA sequences (and organelle sequences).

Agarose Electrophoresis

Since DNA has a large number of phosphate groups which are negatively charged at neutral pH, it will migrate toward the anode in an electric field. The basic phosphodiester backbone of all DNAs is the same so they have a uniform charge density. Therefore, if the DNA is placed in a porous medium such as agarose or polyacrylamide, it will migrate toward the anode at a rate which is proportional to its molecular weight. This is the technique which is used to separate DNA fragments of different sizes produced by digestion of DNA with restriction enzymes.



reflect the distribution of restriction sites in the DNA. The fragments produced will thus be specific for each target DNA /restriction enzyme combination and can be used as a "fingerprint" specific for a given target DNA (or for the organism containing that DNA).

Relatively small DNAs, such as chloroplast DNA, will usually produce about 40 discrete restriction fragments when digested with a typical restriction enzyme, such as EcoRI. The restriction fragments produced by digestion of purified chloroplast DNA can be separated according to size by subjecting the DNA to agarose gel electrophoresis after digestion with a restriction enzyme. After the gel is stained with ethidium bromide, the pattern of restriction fragments can be directly observed or photographed in ultraviolet light. Chloroplast DNAs which differ from one another in base sequence, or have been rearranged by insertions, deletions, or inversions will produce restriction fragments of different sizes. Such differences in fragment size, arising from restriction enzyme digestion of nuclear, organellar, or total DNA, are called restriction fragment length polymorphisms (RFLP) and can be used as a direct measure of genetic variability.

Many studies have utilized RFLPs of chloroplast DNA to study phylogeny and systematics in various plant groups. These studies have shown that the analysis of RFLP variation in chloroplast DNA is very useful in unraveling systematic relationships in plants. Unfortunately, however, the utility of chloroplast DNA in plant breeding is extremely limited. Most genes of agronomic importance are located on nuclear chromosomes; few are found in chloroplast DNA.

RFLP analysis can be applied to chromosomal DNA also, but is more complex because of the much greater complexity of nuclear DNA. Digestion of the nuclear DNA from a higher plant with a typical restriction enzyme produces millions of discrete DNA fragments in a continuous range of sizes. If digested DNA is subjected to gel electrophoresis and stained with ethidium bromide no distinct fragments can be visualized. So many fragments are produced that the DNA appears to run as a continuous smear. However, individual restriction fragments are still well-resolved in the gel, and RFLPs are still present between DNAs from different organisms. More complex techniques have to be used to resolve them.

These techniques involve the use of cloned DNA probes and DNA hybridization.

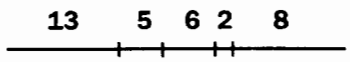
Libraries of Cloned Probes

Since RFLPs of nuclear DNA cannot be directly visualized, the usual procedure is to use small pieces of chromosomal DNA as probes to detect individual restriction fragments. Using the high specificity of DNA-DNA hybridization, such probes can detect individual restriction fragments in the complex mixture of fragments of nuclear DNA present in a restriction digest. To use this technique a set of chromosomal DNA fragments is prepared to use as probes. Such a set of probes is called a library. DNA isolated from the species of interest is digested with a restriction enzyme, and relatively small fragments (usually 2-5 kb) are used as DNA hybridization probes. Individual restriction fragments can be used as a probe, but a supply of the individual fragments in pure form is needed. It would be very difficult to directly isolate individual restriction fragments, but fortunately it is not necessary to try to do this. Instead, one utilizes gene cloning technology and the ability of a bacterium such as *E. coli* to very accurately replicate DNA. In this technique, individual restriction fragments are ligated into a bacterial plasmid, and the plasmid is then transformed into a bacterial cell. The bacterium replicates the plasmid as it grows and divides. By growing a culture of these transformed bacteria and then isolating the plasmid, one obtains a large supply of a single plant DNA restriction fragment which is suitable for use as a hybridization probe. The bacterial strain carrying a fragment of interest can be maintained for a long period of time, and can thus be repeatedly used or distributed to other researchers for use in their experiments.

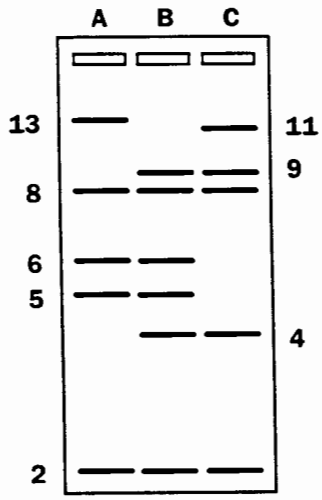
Phylogenetic Analysis using RFLPs

In the simplified example illustrated here, RFLP analysis is applied to a phylogenetic problem. The restriction map for a small region of the chloroplast genome of Plant A is shown, along with RFLP patterns for Plant A and two other plant accessions (B and C). If plant A is assumed to be least changed from the progenitor species, two mutations can be deduced:

- Mutation 1, a new restriction site appeared in the 13 kb fragment producing 9 and 4kb fragments
- Mutation 2, a restriction site was lost in the line leading to plant C, causing the appearance of an 11kb fragment and the loss of the 6 and 5 kb fragments.

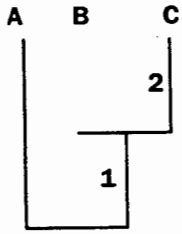


Restriction map of DNA from plant A



Agarose gel

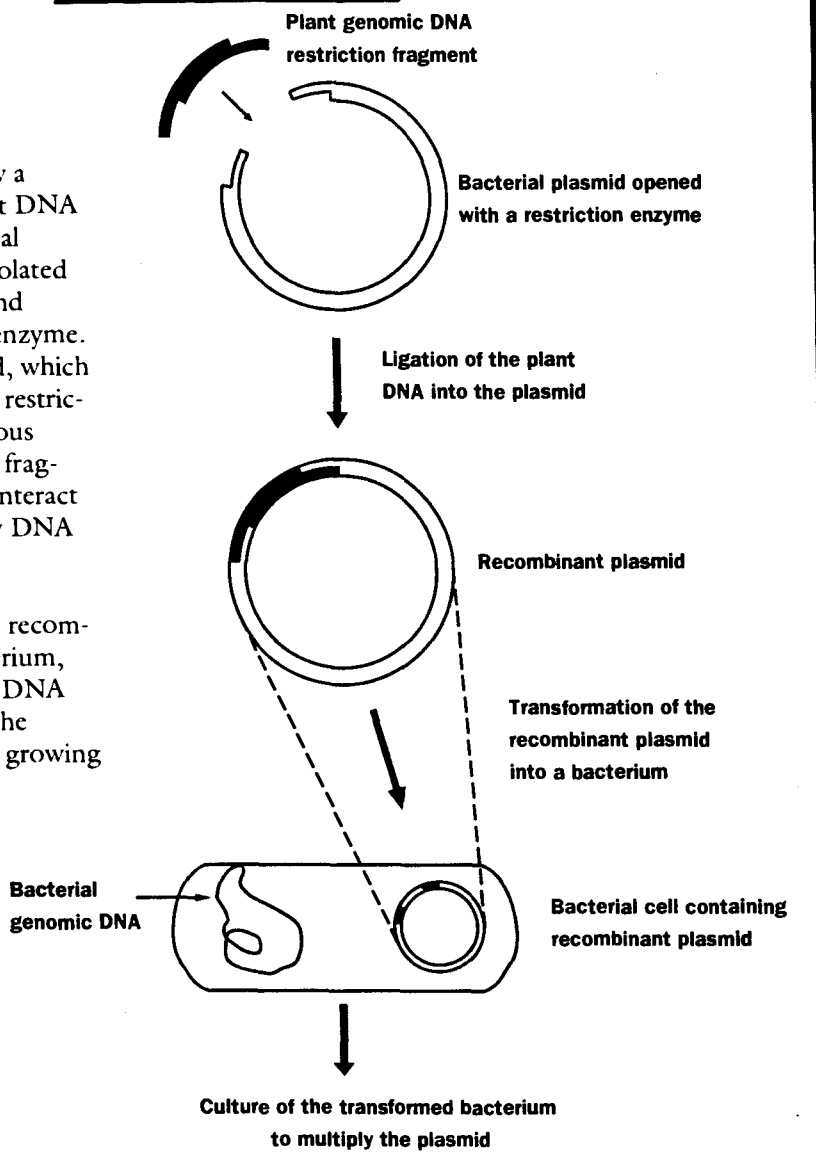
A phylogeny based on this series of events would be:



Cloning of Plant DNA

This diagram illustrates how a restriction fragment of plant DNA can be cloned into a bacterial plasmid. DNA would be isolated from the plant of interest and digested with a restriction enzyme. When mixed with a plasmid, which has been cut with the same restriction enzyme, the homologous single-stranded ends of the fragment and the plasmid can interact and be covalently joined by DNA ligase.

After transformation of the recombinant plasmid into a bacterium, large amounts of the plant DNA fragment are produced as the plasmid is replicated in the growing culture of bacteria.



Using cloned probes to detect RFLPs in nuclear DNA preparations is done in the following manner. DNA from the plants to be compared for RFLP differences is isolated, digested with a restriction enzyme, and then fractionated on an agarose gel. As explained above, the DNA will now be in the form of millions of restriction fragments fractionated in the gel by molecular weight. In order to use DNA-DNA hybridization to detect specific fragments, the probe DNA and the DNA in the gel must be single-stranded (denatured). The gel is soaked in a base such as NaOH to denature the DNA, and to facilitate hybridization the DNA is transferred out of the gel onto a membrane filter by a procedure called a "Southern" transfer (or blot).. A filter cut to the same size as the gel is placed directly against the gel, and the DNA is eluted out of the gel onto the filter. Since the filter is in direct contact with the gel, the pattern of restriction fragments in the gel is maintained on the filter. The denatured DNA binds very tightly to such filters, and the filter can be repeatedly used for hybridization experiments.

After a filter is prepared by Southern transfer, it can be probed with one of the cloned probes from the library. Basically, this procedure consists of denaturing the cloned probe and allowing it to hybridize with the filter containing the nuclear DNA. Under the proper conditions of temperature and salt concentration, the

denatured probe will specifically hybridize to any nuclear restriction fragments on the filter which are homologous to it. However, in order to see where the probe hybridized, it is necessary to label it in some way. The most common way to do this has been to radioactively label the probe with phosphorous-32 (^{32}P). The labeled probe is allowed to hybridize to the DNA filter, excess unhybridized probe is washed off the filter, and the filter is subjected to autoradiography to detect the restriction fragments to which the probe has hybridized. In this way, individual restriction fragments can be detected in complex preparations of higher plant DNA. Currently, probe labeling techniques which do not require the use of radioisotopes are being developed.

Construction of RFLP Genetic Maps

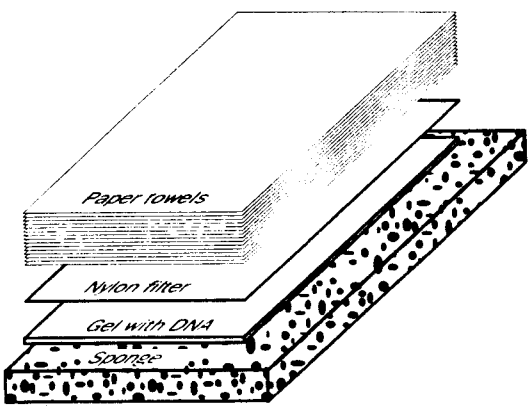
In order to discuss the utility of RFLP markers in genetic mapping, it is useful to think of genetic mapping in a slightly different fashion. Consider a genetic cross in a diploid organism in which two parents are crossed to produce an F_1, and the F_1 is selfed to produce an F_2 population. To simplify the discussion we will assume the the parent plants have been selfed through several generations and are homozygous at all loci. It is clear that the F_1 hybrid from two such parents will have one set of chromosomes from each of the parents and that the two members of a homologous chromosome pair will be different from one another to the degree which the parents differ in DNA base sequence.

When the F_1 plant undergoes meiosis to produce gametes, its chromosomes will undergo recombination by crossing over. This

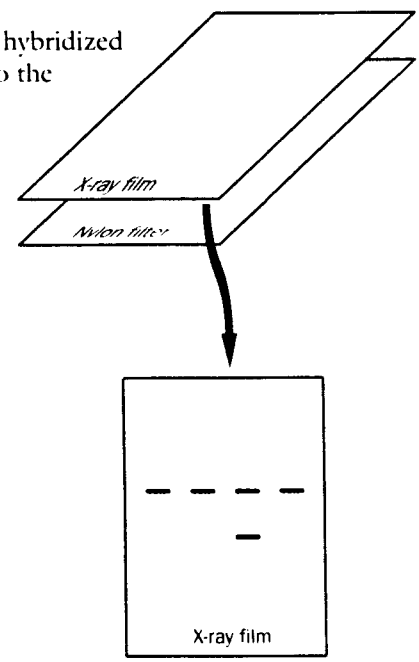
Southern Transfer and Autoradiography

After DNA has been fractionated by agarose electrophoresis, it can be transferred out of the gel and on to a membrane filter by the process called Southern transfer. An exploded view of this process is diagrammed below. The sponge would be partially immersed in a suitable buffer. As the buffer passes through the gel on its way to being absorbed in the paper towels, the DNA is carried out of the gel and binds to the filter.

After the filters are prepared, radioactive RFLP probes are hybridized to them. Location of restriction fragments homologous to the probes is determined by autoradiography, and RFLP data can be scored from the developed film.



Southern Transfer



Autoradiography

recombination process will form gamete chromosomes which are mosaics with segments from each of the two parental chromosomes, and no two chromosomes will have an identical array of segments. In F_2 plants produced by selfing the F_1, homologous pairs will contain different mosaics of parental chromosome segments.

This recombination process is the basis of conventional genetic mapping, which depends on two observations. 1. Chromosome segments which are not on the same chromosome will undergo random recombination. 2. Chromosome segments on the same (homologous) chromosome will undergo recombination according to a function which depends on their physical distance apart. Chromosome segments which are close together (closely linked) will undergo recombination less commonly than will segments which are further apart.

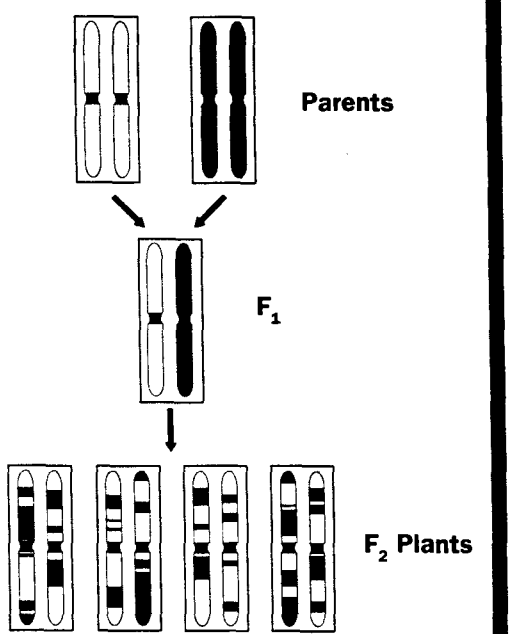
Thus genetic distance or map distance is defined as a function of the recombination which occurs during gamete formation.

Previously, the only way to follow chromosome segments and observe their recombination during gamete formation was to observe the phenotype caused by the action of genes which happened to be on the segments. By observing a phenotype such as flower color, plant height, insect resistance, or characteristics of the endosperm, one could infer the behavior of the respective chromosome segments during meiosis and recombination. Following chromosome segments in this indirect manner has been utilized to construct fairly detailed genetic maps of several plant species. However, the method is cumbersome and extremely time-consuming.

Fortunately, RFLP markers provide a way to directly follow chromosome segments during

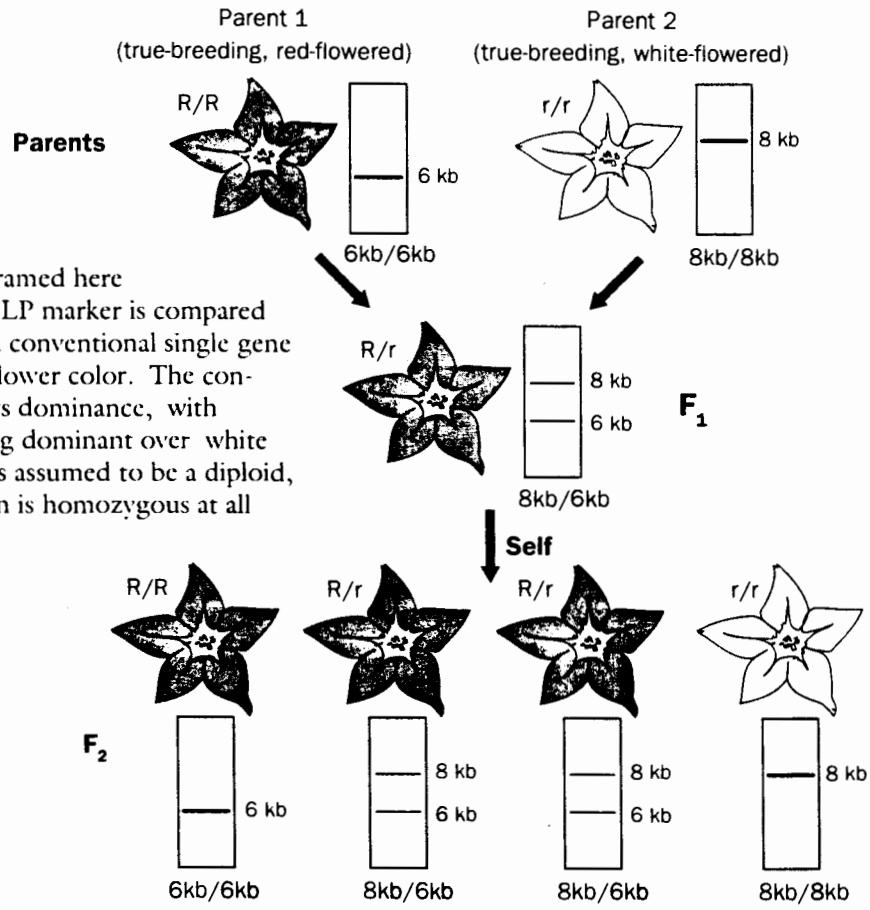
Inheritance of Chromosome Segments

This diagram illustrates the way chromosome segments are inherited in a genetic cross. Two inbred, completely homozygous parents are crossed to produce an F_1, which is heterozygous at all loci (only a single chromosome pair is shown). During selfing of the F_1, recombination produces gamete chromosomes which contain varying parental segments. Four F_2 plants illustrate the distribution of chromosome segments possible in such a segregating population.



Since large numbers of RFLP markers can be analyzed in a single mapping population, it is possible to construct chromosome segment maps of a segregating population. These maps show the distribution of chromosome segments from each of the parents for the entire genome of each individual in the population. This allows direct selection of those individuals whose propagation will lead to the fastest advance toward the goals of the breeding program. For example, if the goal was to introduce a single gene in a backcross program, individuals should be selected to have the maximum number of chromosome segments from the recurrent parent while still retaining the segment containing the gene of interest.

RFLP Inheritance Patterns



In the example diagramed here inheritance of an RFLP marker is compared with inheritance of a conventional single gene marker controlling flower color. The conventional gene shows dominance, with colored flowers being dominant over white flowers. The plant is assumed to be a diploid, selfing species, which is homozygous at all loci.

recombination and greatly simplify the construction of this type of genetic map. Instead of looking at the phenotype caused by the presence of a gene on a chromosome segment, one looks directly at an RFLP marker on the segment itself. Thus one is looking directly at the genotype of the plant, rather than indirectly, through the phenotype produced by gene action. This has a number of advantages, which are discussed below.

When one considers the inheritance pattern of RLFP markers, it will be seen that they segregate exactly in the same way as do conventional gene markers and follow strict Mendelian rules. This means that segregation patterns of RFLP markers can be analyzed by conventional Mendelian methods, and that maps of RFLP markers can be constructed in the same fashion as can maps of conventional markers.

Comparison of Conventional Genetic Markers and RFLP Markers

When conventional genes and RFLP markers are compared, it is apparent that RFLP markers have many advantages for the construction of genetic maps. Some of the most obvious of these are:

Natural variation - To construct a genetic map by following chromosome segments through segregation, one must be able to distinguish segments from each of the two parental chromosomes. If conventional genes are used to follow the chromosome segments, then each parent must have different alleles of the gene in question (sometimes referred to as wild type and mutant). Often, one cannot depend on natural variation of conventional markers, and variants will have to be produced by mutagenesis. RFLP mapping, on the other hand, utilizes

abundant natural variation in DNA sequence, and no mutagenesis is necessary. Unlimited numbers of RFLP markers are available in virtually any plant species, whether or not any formal genetics has been done with it.

Developmental stage or plant organ - To infer the segregation of a chromosomal segment by analyzing the segregation of conventional genetic markers, one has to depend on gene expression. This can lead to many complications. For example, genes may be expressed only in certain tissues (flower color genes can normally be scored only when one has flowers present), or at certain developmental stages, or under certain environmental conditions. In some cases, scoring for a conventional gene may be expensive or time-consuming. For example, to score for an insect resistance gene, it is often necessary to build cages around the plants and introduce insects (which must be the right biotype in a suitable developmental stage). Since RFLP markers directly detect RFLP variation in the chromosomal DNA, any developmental stage and virtually any plant part from which a small amount of DNA can be isolated is suitable for analysis.

"...any developmental stage and virtually any plant part from which a small amount of DNA can be isolated is suitable for analysis."

"Since scoring of RFLP markers does not depend on gene expression, it is not affected by the environment"

Phenotypic effect - RFLP markers are reflections of the natural variation present in the DNA sequence of plants. In the vast majority of cases, they will have no phenotypic effect at all. An effectively unlimited number of RFLP markers are naturally present in any plant. Contrast this with phenotypic markers. In most cases only a few phenotypic markers can be maintained in any one plant, either because the mutant phenotypes are harmful to the plant or because the mutations have been induced in different plants, and it would be very time-consuming to try to assemble these by crosses into any one plant line. These facts have important consequences for the construction of genetic maps.

Since individual plants contain a great deal of RFLP variation, one can often construct a detailed RFLP map utilizing only one cross, one mapping population, and one modest-sized research group to gather and analyze the data. Conventional genetic maps, however, always require the cooperation of many research groups and represent years of effort and hundreds of crosses.

Environmental effect - Conventional genes depend on gene expression for scoring. The DNA sequence comprising the gene must be transcribed into mRNA, the mRNA must be translated to protein, and the protein must assume the proper configuration and localization to exert its effect. The effect exerted by the protein is often due to its enzymatic activity or ability to bind to some cell component. All these stages can, and often are, affected by the environment. Since scoring of RFLP markers does not depend on gene expression, it is not affected by the environment. RFLP markers simply represent the presence or absence of characteristic base sequences.

Interaction with other genes - Since conventional gene markers depend on gene expression to be scored, their phenotypic effect is often modulated by other genes. One of the simplest cases is dominance; many mutations of common genes are recessive. These can only be scored in plants homozygous for the mutant allele, and heterozygotic plants cannot be distinguished from homozygotic wild-type plants. RFLP markers are codominant; the three possible genotypes for one marker can always be distinguished. In addition, conventional markers will often display different phenotypes in different genetic backgrounds due to gene interactions. Since RFLP markers are a direct reflection of the genotype , they are independent of one another, and an RFLP marker will maintain its identity in any genetic background.

Construction of Genetic Maps using RFLP Markers

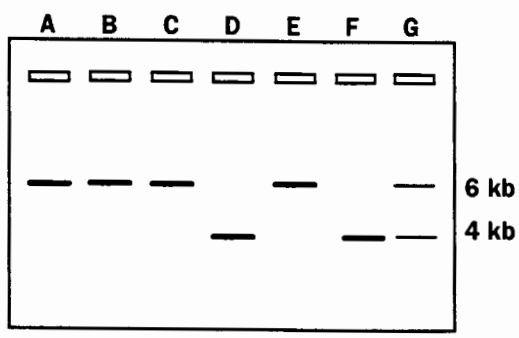
To construct an RFLP map a fairly well-defined series of steps is followed. These are outlined below. For several crop species, such as rice, tomato, and maize, RFLP maps are already available, and one can simply use these without having to construct a new one.

1. Select the parent plants - As parent plants in a cross for RFLP map construction, one wants to select plants which are genetically divergent enough to exhibit sufficient RFLP, but not so far distant as to cause sterility of progeny. It would be advantageous also to select a cross in which some desirable agronomic traits are segregating, but this has not always been possible. In most plants thus far mapped, it has been possible to find sufficient variation between varieties of a single species. For instance, in the Rockefeller rice RFLP mapping project, a cross between indica and javanica types of rice was used. In tomato, however, the genetic base of the cultivated types is too narrow, and to obtain a usable level of polymorphism, a cross had to be made between a cultivar and a wild species.

Since one does not know *a priori* how much RFLP variation will be present in a crop species, it is necessary to make a survey using a random selection of cloned probes. Such a survey should include cultivars or plant introductions thought to represent the spectrum of variability present in the taxon. Related wild species which can be crossed to the cultivar should also be included. One should utilize any available information to select the plants to be surveyed. Existing systematic or evolutionary studies should be consulted; isozyme surveys may be available.

After selecting a range of plants to be surveyed, DNA is isolated from individual plants of each of the accessions, digested with restriction enzymes, and screened for polymorphisms by the usual method. Each accession can then be scored for RFLP alleles present, and two accessions showing a usable amount of polymorphism can be selected for crossing.

Polymorphism Survey



This figure shows results which might be obtained by hybridizing a cloned probe to DNA from several accessions of a plant species. Accessions A, B, C, and E are homozygous for a 6kb allele, accessions D and F are homozygous for a 4kb allele, accession G is heterozygous for the two alleles.

2. Produce a mapping population - The selected parent plants are crossed to produce an F_1 plant or plants. In the case of a selfing plant such as rice, all F_1 plants should be alike and heterozygous for all RFLPs exhibited between the parent plants. A mapping population can be derived by selfing the F_1 to produce an F_2, which is then scored for segregation of the RFLPs, or by backcrossing the F_1 to one of the parents and observing segregation in the first backcross generation. It is better to use an F_2 population if this is possible, since more information can be gained than from a backcross population of comparable size. A mapping population of about 50 F_2 or backcross plants is sufficient for a fairly detailed map. With plants in which plant breeding studies are actively being carried out, suitable mapping populations might already be available as remnant seed from previous crosses. If this is the case it may be possible to make a map without having to make any crosses. This can greatly speed up the process.

3. Score RFLPs in the mapping population - Once the mapping population is obtained, DNA is isolated from each individual plant in the population. It is important to realize that the chromosomes of each plant in the mapping

population contain a unique array of parental chromosome segments. To derive an RFLP map it is necessary to determine which parent's chromosome segments are present in each plant in the mapping population. Since the unique makeup of chromosomal segments would be destroyed by recombination if the F_2 or backcross plants were to be allowed to reproduce sexually, every effort should be made to keep the individual plants in the mapping population alive so that repeated DNA extractions can be made, and a large amount of DNA can be accumulated. Mapping proceeds by sequentially scoring RFLPs in the individual plants of the mapping population by the following series of procedures.

A. *Screening for polymorphism* - Probes are selected sequentially from the library and tested against the parents in an effort to determine which restriction enzymes will detect a polymorphism between the parents. The number of enzymes required will depend on the amount of variation present in the parent plants. Very polymorphic plants, such as inbred lines of maize, will usually require only one or two enzymes to detect polymorphism with most probes. In rice, 11 enzymes were used and about 75% of the probes were polymorphic with these.

B. *Scoring* - When a polymorphic probe/enzyme combination is detected, it can then be scored in the mapping population. To accomplish this a series of agarose gels must be prepared and filters prepared from them for DNA hybridization. These filters are generally called F_2 survey filters or mapping filters. There must be one filter set for each restriction enzyme to be used in the mapping project, and each filter set must contain digested DNA from each individual plant in the mapping population. Thus one set of mapping filters would contain DNA from each of the F_2 plants digested with *Eco*RI, and another set would contain DNA from each of the F_2 plants digested with *Pst*I. If a polymorphism between the parent plants has been detected with a given probe/enzyme combination, then that probe will be scored on the corresponding mapping filter set. For example, if probe #36 is found to be polymorphic between the parents using *Pst*I, then it could be scored on the *Pst*I mapping filter set. In an F2 mapping population scored for a single RFLP, only three types of plants can be present. There are two homozygotes (one similar to each of the parents) and the heterozygotes. Thus each plant in the mapping population can be scored as a heterozygote or as one of the two possible types of homozygotes when tested with each library probe.

C. *Linkage analysis* - Data accumulated from scoring the mapping population sequentially with probes from the library is used to construct the linkage map. Linkage analysis is based on the degree to which probes tend to co-segregate. The first probe scored in a mapping study will, of course, provide no information on linkage, but beginning with the second probe one can determine whether linkage in indicated. If the second probe is linked to the first one, they will tend to co-segregate (F_2 plants which are heterozygous for the first probe will tend to also be heterozygous for the second probe; plants homozygous for the allele found in one of the parents for the first probe will also tend to be homozygous for the second probe). If no linkage is indicated, distribution of homozygotes and heterozygotes for the first two probes will tend to be random. Simple statistical tests, such as a chi square analysis, will test for randomness of segregation and hence linkage. The first few probes, since they are randomly selected, are unlikely to be linked and hence to show co-segregation. As one sequentially adds probes, however, linkages will be detected. At first there will be many more linkage groups than there are chromosomes, but the two numbers will tend to converge as more markers are added. As each probe is screened, it is tested for possible linkage to all the other markers which have been mapped before. Thus when one is testing probe #100, one can test it against the 99 other probes that have been scored before to see if it is linked to any of them. This is essentially like making a 100 point cross. It is evident that a great amount of data is rapidly accumulated in an RFLP mapping project, and a computer is required for efficient data storage and analysis.

When a new probe is scored, if linkage with one or more markers is indicated by the data, the map distance is determined by computer analysis using an algorithm such as maximum likelihood. Recombination data can be converted to map distances using any of the commonly used mapping functions.

RFLP maps and conventional maps

One of the goals of an RFLP mapping project is the production of a "saturated" map, which is a map with RFLP markers spaced every 10-20 centimorgans (cm) over all the chromosomes. On such a map, any conventional gene will be within a few cm of an RFLP marker. Since conventional plant genetic maps are about 1500 cm, about 150 well-spaced markers would be sufficient for a saturated map. Since probes are randomly selected (and will represent random chromosome locations), however, it will be necessary to map several hundred to achieve this degree of saturation.

Rationalization of RFLP maps with conventional maps - After construction of a saturated or near saturated RFLP map, the normal next step is to associate the linkage groups of RFLP markers with the conventional genetic map, if one is available. This can be done in several ways depending on the particular situation. If aneuploid lines of the plant are available, these can be very useful for rationalizing RFLP and conventional genetic maps. This technique depends on the fact that one can usually determine the number of RFLP loci present from the intensity of the hybridization signal. Thus if an RFLP marker is on chromosome 1, one would expect to see about half the signal intensity in a plant monosomic for chromosome 1 when compared with a normal plant having two copies of chromosome 1. If a complete set of monosomics is available, each of the RFLP linkage groups can be assigned to a chromosome. Trisomic plants can be used in a parallel fashion.

Placing conventional genes on the RFLP map - Once an RFLP map is constructed, conventional genes can be placed on the map. In order to do this it is necessary to have a mapping population in which both the conventional gene and the RFLP markers are segregating. By determining which RFLP markers show co-segregation with the conventional marker, one can place the conventional marker on the map.

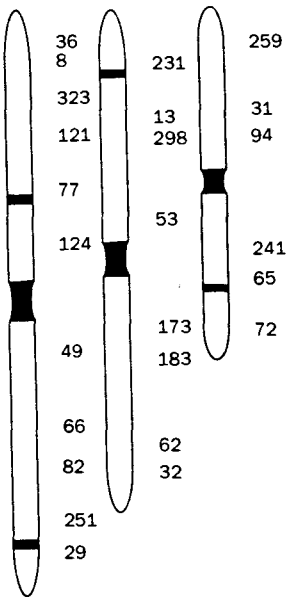
Uses of RFLPs in Plant Breeding

Construction of an RFLP map can be very interesting, and during map construction one can gather data which is useful in systematic or evolutionary studies. However, an RFLP map is not useful by itself in plant breeding. It is only useful when it is used in conjunction with analysis of conventional markers. One way in which RFLP maps can be used to supplement regular plant breeding protocols is by utilizing indirect selection. This procedure is useful when one wants to select for a conventional gene, but directly selecting for that gene would be expensive, difficult, or time-consuming. With indirect selection, one does not directly select for the gene of interest, but rather for one or more closely linked RFLP markers. If the RFLP markers are indeed closely linked, they will remain associated with the gene of interest during segregation. This allows one to select for the RFLP marker with confidence that the conventional gene will also be present, since only relatively rare recombination events would separate the two. For example, if a conventional gene is flanked by two conventional markers each 10 cm away, 99% of the time one or another of the markers would remain associated with the conventional gene of interest.

"...one does not directly select for the gene of interest, but rather for one or more closely linked RFLP markers."

Indirect selection for an RFLP marker, rather than selecting directly for the gene, can have advantages in several plant breeding scenarios.

Quantitative Traits



This figure illustrates how RFLP maps can be used to analyze quantitative traits in plants. Three chromosomes of a hypothetical plant are shown at the left. Numbers to the right of the chromosomes indicate the position of mapped RFLP markers. The position of some genes affected a quantitative trait are shown as shaded areas on the chromosomes.

Since all areas of the chromosomes are "tagged" with RFLP markers, it is possible to follow them through a genetic cross and to analyze their affect on the trait of interest. In this example, RFLP markers numbers 77, 29, 231, and 65 will show strong association with the trait.

One such situation might involve selection for a recessive gene. Suppose one were trying to introgress a recessive gene into a cultivar by using a backcross breeding program. Backcross protocols involve alternate backcross and selection phases. It is necessary to select progeny bearing the desired gene at several points in the backcross cycle. If a gene is dominant it can be directly selected, but recessive genes will not be expressed in any of the backcross plants, and it would be necessary to carry out progeny testing (probably by selfing the backcross plants to test for the presence of the recessive gene). A recessive allele could be indirectly selected, however, by selecting for a linked RFLP marker. No progeny testing would be necessary, and the process could be greatly speeded up. In some cases indirect selection for dominant genes could also be profitable in a backcross program. Plants can be scored for a linked RFLP marker while in the juvenile stage, since only a little DNA needs to be isolated, and the plant does not have to be destroyed. Plants not bearing the desired allele could be weeded out early, saving space and expense.

Another situation in which indirect selection could be of great use would be in selecting plants in a breeding protocol which have two or more independent genes which give a similar phenotype (sometimes called pyramiding). For example, it might be advantageous in breeding insect resistant crops to include more than one resistance gene in a newly developed cultivar. However, selection for more than one gene conferring resistance to a given insect is difficult since the phenotype is the same whether one, two or more resistance genes is present. Again, some sort of progeny testing would be necessary, but where more than one gene is involved progeny testing quickly becomes impractical. However, once linkage of the individual insect resistance genes to RFLP markers is accomplished, one can easily follow any number of them through a cross. Individual plants can be scored for the presence of one, two, three, or more resistance genes with no progeny testing needed.

"..the analysis of quantitative traits can be carried out by conventional Mendelian analysis, and much of the uncertainty is removed."

Quantitative traits - Unfortunately, most of the traits of agronomic significance in plants are not independent single genes, but are controlled by a number of genes scattered over the chromo-

some complement. Each of the individual genes of such a polygenic system contributes a small positive or negative effect to the trait of interest. Clear dominance is not exhibited and the phenotype has a large environmental component. All these characteristics conspire to make quantitative traits very difficult to analyze. Conventional Mendelian methods of analysis which are suitable for single gene traits cannot be applied to analysis of quantitative traits, and one is forced to use biometrical methods and extensive testing over different years and in different environments in an effort to advance toward the desired state. Fundamental advance in this area of plant breeding seemed unlikely before the advent of RFLP mapping techniques. Since one can use RFLP markers to simultaneously follow the segregation of all chromosome segments during a cross, the basic idea is to look for correlations between the quantitative trait of interest and specific chromosome segments marked by RFLPs. If correlations exist, then the chromosome segment must be involved in the quantitative trait (one or more genes determining the trait must be on that chromosome segment). The difficult part in the procedure is establishing the correlations between the trait and specific chromosome segments. The RFLP markers are easily scored, but the quantitative trait must be scored in the conventional fashion. Once this time-consuming process is completed, however, and specific chromosome segments are implicated in the trait, specific chromosome segments with a positive effect on a quantitative trait can be selected from a population of plants and incorporated into single plants with high efficiency. This is possible because of the ability to score for several RFLP markers simultaneously in a single plant in a manner which is free from environmental influence or gene interactions. Thus the analysis of quantitative traits can be carried out by conventional Mendelian analysis, and much of the uncertainty can be removed.

Summary

It is evident that RFLP markers can be directly used as probes for the presence or absence of certain chromosome segments. The ability to directly follow chromosome segments has many fundamental applications in plant breeding and molecular biology. During the next few years RFLP analysis will increasingly take its place as another tool which plant breeders will routinely use in their efforts to produce superior plants.

Glossary

cDNA - complementary DNA, a fragment of DNA which has been produced from an RNA sequence by reverse transcription. Messenger RNA is commonly used to synthesize cDNA.

kb - kilobase, 1000 bases or base pairs. DNA sizes are often expressed in kilobases.

library - a collection of cloned DNA fragments. The library may consist of cDNAs or genomic clones (fragments cloned directly from cellular DNA).

ligation - joining of DNA fragments to produce a single DNA molecule. *Ligases* are enzymes which perform this reaction. Fragments of plant DNA are *ligated* into bacterial plasmids during the cloning of probes to use for RFLP analysis.

picogram - 10^{-12} grams, a unit commonly used to express the DNA content per cell or per nuclear genome in a plant.

plasmid - DNA which is not part of the regular genome of an organism,and self-replicates during growth. Plasmids commonly occur as linear or circular elements, and may be present in one to many copies per cell.

probe - a DNA or RNA fragment used in nucleic acid hybridization experiments to detect RFLPs.

repeated DNA sequence - a sequence of nucleotides which occurs more than once in a genome. Repeated sequences may be present in a few to many millions of copies. The individual repeated sequence may be only a few nucleotides in length up to several kb.

restriction enzyme - an endonuclease with the ability to cleave DNA at the point where a certain base sequence occurs.

restriction site - a DNA base sequence recognized by a restriction enzyme.

RFLP - restriction fragment length polymorphism, a difference between samples of DNA detected as differing fragement sizes produced after treatment with a restriction enzyme.

single copy sequence - a sequence of nucleotides which occurs only once in a genome.

Southern transfer (blot) - transfer of DNA from an electrophoresis gel onto a nitrocellulose or nylon membrane. The membrane is then used in hybridization experiments.

transformation - the introduction of DNA into the cells of an organism by a method other than conventional sexual crossing. Plasmids are commonly transformed into bacterial cells by direct DNA uptake through the cell wall and membrane.

REVIEW/

RFLP MAPPING IN PLANT BREEDING: NEW TOOLS FOR AN OLD SCIENCE

S. D. Tanksley, N. D. Young, A. H. Paterson, M. W. Bonierbale
Department of Plant Breeding, Cornell University, Ithaca, NY 14853.

Breeders have traditionally improved plant varieties by selecting on the basis of phenotype. Now restriction fragment length polymorphism (RFLP) linkage maps are being constructed for most major crop plants and these maps provide a more direct method for selecting desirable genes via their linkage to easily detectable RFLP markers. The integration of RFLP techniques into plant breeding promises to: (1) Expedite the movement of desirable genes among varieties, (2) Allow the transfer of novel genes from related wild species, (3) Make possible the analysis of complex polygenic characters as ensembles of single Mendelian factors, and (4) Establish genetic relationships between sexually incompatible crop plants. In the future, high density RFLP maps may also make it possible to clone genes whose products are unknown, such as genes for disease resistance or stress tolerance.

For thousands of years man has been breeding plants by selecting individuals that are higher yielding, better tasting, or otherwise more attractive. Crop varieties first selected by early man and later improved by modern plant breeders form the basis of today's sophisticated and highly-productive agricultural economy.

As with many aspects of biology, recombinant DNA technology has had an enormous impact on plant science. Many plant genes have already been cloned, but few of these have been directly associated with characters of major agricultural importance, such as yield or product quality. This is hardly surprising since these important characters are often under complex genetic control, involving the joint action of a number of genes.

In the absence of much understanding of genes controlling characters of economic interest, plant breeders develop new varieties by selecting plants with desirable phenotypes. However, a plant's phenotype is determined, not only by its genetic composition, but also by the environment in which it is grown. In many cases effects of the environment mask those of the genotype, so the phenotype provides an imperfect measure of a plant's genetic potential. To deal with this problem, elaborate plant breeding techniques, based on statistical inference, have been developed[1]. These techniques have been very useful in producing new crop varieties, but at the cost of much time and effort. For many crops the time for development of an improved variety often exceeds 10 years.

GENETIC MAPS AND PLANT BREEDING

The science of plant genetics traces back to Mendel's classic studies with garden peas[2]. Since that time, researchers have been identifying, cataloging and mapping single gene markers in many species of higher plants. Genetic maps of crop species were among the first to be constructed[3,4] and predate the demonstration of DNA as the hereditary material. Until recently, however, the genetic markers used to develop maps in plants have been those affecting morphological characters, including genes for dwarfism, albinism, and altered leaf morphology.

Scientists have long theorized about the use of genetic maps and markers to speed up the process of plant and animal breeding. In 1923 Sax[5] proposed identifying and selecting for "minor genes" of interest by linkage with "major genes", which could be scored more easily. This idea has resurfaced many times and the concept has been extended by many workers[6]. Unfortunately, most morphological markers cause such large effects on phenotype that they are undesirable in breeding programs. Moreover, they mask the effects of linked minor gene(s), making it nearly impossible to identify desirable linkages for selection.

In recent years, isozymes have been used very successfully in certain aspects of plant breeding and genetics as nearly-neutral genetic markers[7]. Unfortunately, the number of genetic markers provided by isozyme assays is insufficient for many applications in plant breeding. As a result, even with the use of isozymes as genetic markers, the full potential of genetic mapping in plant breeding has yet to be realized[8].

GENETIC MAPPING BASED ON RESTRICTION FRAGMENT LENGTH POLYMORPHISMS

The potential impact of restriction fragment length polymorphism (RFLP) mapping on eukaryotic genetics was first described in 1980 by researchers in human genetics[9,10]. RFLPs now occupy a major role in human genetic research and the first complete genetic linkage map of the human genome, based on RFLP markers, has recently been published[11].

Mapping with RFLPs involves the application of molecular biological techniques to the basic concepts of transmission genetics. Single copy DNA is cloned from a species of interest and used as a probe to follow the segregation of homologous regions of the genome in

individuals from segregating populations such as an F2 or backcross (Fig. 1). Since many different single copy clones can be tested, genetic linkage maps can be constructed which contain a very large number of markers at close intervals.

The primary requirements for developing an RFLP-based linkage map of an organism are: (1) sexual reproduction and (2) a source of single copy DNA clones. Most crop plants are sexually-reproducing and, for most species, it is easy to obtain large numbers of offspring that can easily be maintained in a vegetative state for extended periods. Single-copy DNA can be found in virtually any organism, and a number of methods have been devised for selectively cloning this fraction of plant genomes[12].

As markers for use in plant breeding programs, RFLPs differ from morphological markers in several important ways. (1) Morphological markers normally have alleles that interact in a dominant/recessive manner. Since RFLP markers are screened directly at the DNA level, they normally behave in a codominant manner, allowing the genotype of a locus to be determined in plants derived from any mating scheme. (2) The level of allelic variation for RFLP markers in natural plant populations is much greater than that for morphological markers[13]. As a result, one can take advantage of natural variation in existing populations eliminating the need to construct special genetic stocks. (3) The majority of RFLP markers are believed to be phenotype-neutral, whereas morphological markers cause major alterations in the plant's phenotype which are usually undesirable in breeding programs. (4) Morphological markers interact epistatically, limiting the number of them that can be unequivocally scored in the same segregating population. RFLP markers, on the other hand, are apparently free of epistatic effects and a virtually limitless number can be monitored in a single population.

Over the past few years, plant geneticists and breeders have begun to develop RFLP maps for important crop species. Such maps are now available for at least seven different plant species[12,14–19], including not only crops with long histories of genetic studies and gene mapping such as tomato[12] and maize[17], but also those for which no genetic maps previously existed such as lettuce[18] and potato[19]. In the near future, RFLP maps will likely be available for most major crop plants. Figure 2 displays the current RFLP map of potato, one of the crops being studied in our laboratory.

DETECTING MAJOR GENES BY LINKAGE WITH RFLP MARKERS

The utility of RFLP markers in plant breeding is based on finding tight linkages between these markers and genes of interest. Such linkage permits one to infer the presence of a desirable gene by assaying for the RFLP marker (Fig. 1). There are a number of single gene traits that are frequently transferred from one genetic background to another by breeders. Genes conferring resistance to pathogens are a classic example[20]. Incorporation of disease resistance genes into sensitive varieties requires crosses with stocks that carry the resistance gene(s), followed by selection among the progeny for individuals possessing the desired gene combination. Traditionally, progeny are screened for the presence of disease resistance genes by inoculation with the pathogen. However, simultaneous or even sequential screening of plants with several different pathogens can be difficult or impractical. In other instances, breeders are unable to screen for resistance to new pathogens because of quarantine restrictions on their shipment and use. In contrast, detecting disease resistance genes by their linkage to RFLP probes

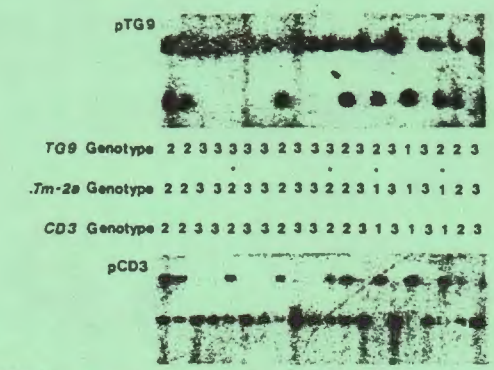


FIGURE 1 Segregation analysis to identify RFLP markers tightly linked with the *Tm-2a* gene[21]. A total of 140 F2 plants derived from a cross between a tomato variety carrying the *Tm-2a* gene from the tomato relative, *L. peruvianum*, and another variety lacking this gene, were scored for their reaction to tobacco mosaic virus and for their genotype with respect to several previously identified RFLP markers. In this system it is possible to score all three disease interaction phenotypes[51]. The results of 20 individuals from the F2 population are shown for two nearby RFLP markers, *TG9* and *CD3*, along with the inferred *Tm-2a* genotype. 1=homozygous for esculentum alleles (mosaic phenotype), 2=heterozygous (necrotic phenotype), 3=homozygous for peruvianum alleles (healthy phenotype), *=recombinant individual.

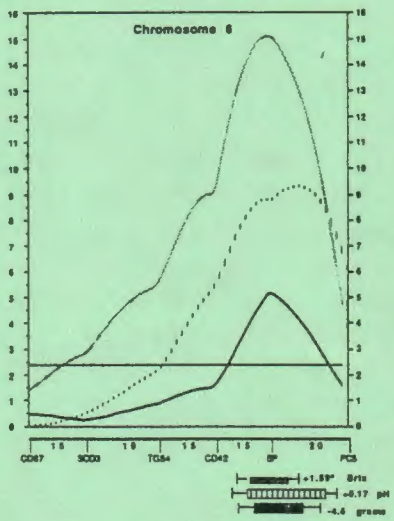


FIGURE 3 QTL likelihood map of chromosome 6 showing location of genes controlling fruit weight (solid lines and bars), soluble solids concentration (dotted lines and bars) and fruit pH (hatched lines and bars) in tomato. The map is based on segregation data from an interspecific backcross between the cultivated tomato and a wild species (*L. chmielewskii*)[22]. X-axis represents the map units along each chromosome. RFLP markers assayed are shown below axis and Y-axis gives LOD (log 10 of the odds ratio) value for probability that a QTL for each character is located at chromosomal point in question[22]. The maximum likelihood effect of a putative QTL, as well as the LOD score in favor of the existence of such a QTL, were determined at points spaced every 1 cM throughout the genome according to Lander and Botstein[25] and a smooth curve was plotted through the points. The horizontal line at the height of LOD=2.4 indicates the threshold level at which the presence of a QTL can be safely inferred. The maximum likelihood position of a QTL in a particular chromosomal region is the highest point on the curve. Bars below each graph indicate a 10:1 likelihood support interval for the position of the QTL while the line extending out from the bars indicates a 100:1 support interval. Phenotypic effects indicated beside the bars are the inferred effect of the QTL at the position on the character of interest. Large effects on all three traits map closely to *sp*, a morphological marker on chromosome 6 which profoundly alters plant development[22]. Similar maps were also constructed for the other 11 chromosomes in the tomato genome[22].

makes it practical to screen for many different disease resistance genes simultaneously without the need to inoculate the population.

In tomato, RFLP markers have been identified that are tightly linked to genes for resistance to tobacco mosaic virus[21], *Fusarium* wilt, bacterial speck and root knot nematodes, as well as major genes controlling plant habit *(sp)*, and fruit ripening properties *(u)*[22]. In maize and lettuce respectively, researchers have established linkages between RFLP markers and genes for resistance to maize dwarf mosaic virus[23] and downy mildew[18]. Similar tight linkages between RFLP markers and other major genes are currently being sought by researchers working with a number of other crop species.

RESOLVING COMPLEX GENETIC TRAITS INTO THEIR SINGLE GENE COMPONENTS

Many important heritable characters are a consequence of the joint action of several genes. Such characters are often referred to as polygenic or quantitative. Yield, maturity date and drought tolerance are examples of characters in plants that are usually quantitatively-inherited. For most quantitative traits, little is known about the number, chromosomal position or individual and interactive effects of genes controlling their expression.

While the theory and techniques of quantitative genetics[24] have proven useful in the study of quantitative traits, these characters continue to be more difficult to manipulate in breeding programs than single gene traits. If complex traits could be resolved into their individual genetic components, it might then be possible to deal with these characters with the efficacy of single gene traits. High density RFLP maps provide this opportunity by making it feasible to identify, map and measure the effects of genes underlying quantitative traits.

RFLP markers can be employed to detect genetic loci (often referred to as quantitative trait loci or QTLs) underlying quantitative traits in a conceptually simple manner. A cross is made between two plants that are genetically different for one or more characters of interest, and segregating progeny are obtained from the hybrid (commonly F2, backcross, or recombinant inbred lines). A number of progeny (usually >100) are evaluated for the character of interest and for their genotypes at RFLP marker loci at regular intervals (10–20 cM) throughout the genome. A search is then made for associations between the segregating RFLP markers and the character of interest. If such associations are found, they should be due to linkage of the RFLP marker to a gene(s) affecting the character.

The ability to detect a QTL with an RFLP marker is a function of the magnitude of the QTL's effect on the character, the size of the population being studied, and the recombination frequency between the marker and the QTL. When an RFLP marker and a QTL are far apart on a chromosome, recombination reduces the likelihood that they will be transmitted to the same progeny individuals. As a consequence, analyses based on single RFLP markers may not detect the QTL with statistical significance, or its effect may be underestimated. The approach of interval mapping[25] ameliorates this problem by basing analyses on two linked markers bounding an interval which may contain a QTL. This reduces the likelihood that genotype at the marker(s) will differ from genotype at the QTL to, at most, the square of the recombination fraction between the flanking markers[25].

Mapping of QTLs throughout an entire genome was recently carried out for the first time using interval analysis and a complete RFLP map to detect and measure the effects of genes underlying several polygenic traits in tomato[22]. The test population was a backcross of a wild tomato species, *Lycopersicon chmielewskii*, from the Andean region of Peru, to the cultivated tomato, *L. esculentum*. The traits that were studied—fruit weight, soluble solids concentration, and pH—were known from prior work to be under polygenic control and to be markedly influenced by environment. 'QTL likelihood maps'[22] were used to show the most likely chromosomal positions, effects, and strength of evidence for individual QTLs underlying each of the three traits studied (Fig. 3). This method is readily applied to other population structures (e.g. F2, recombinant inbred), and is currently being tested on new populations of tomato and other crop species.

Complete RFLP maps and interval analysis permit direct investigation of the number and nature of QTLs underlying a character. Each QTL can be studied as a discrete entity, and its individual and interactive properties measured. RFLP markers also permit the rapid and precise transfer of QTLs into superior crop varieties or genetic stocks for further experiments (see next section). Such stocks will likely prove valuable in demonstrating the morphological and physiological effects of individual QTLs and in revealing the biophysical basis of complex traits.

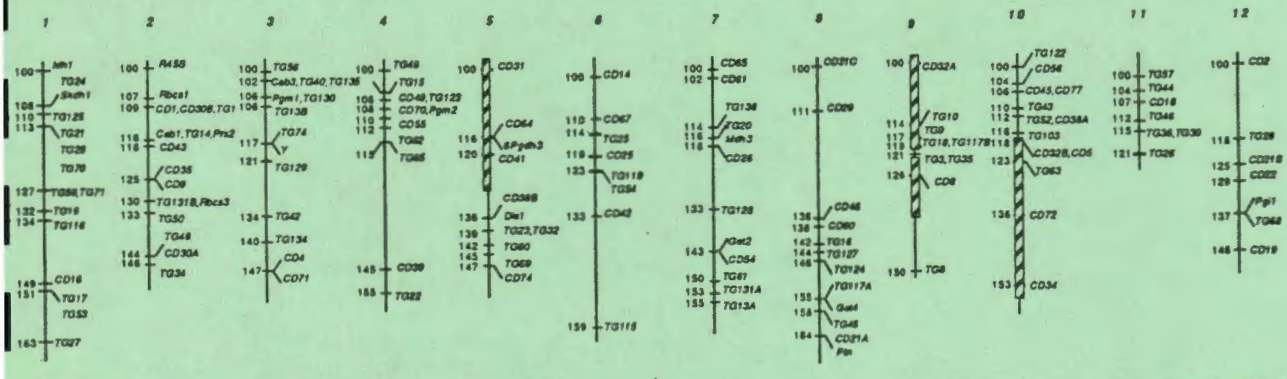


FIGURE 2 Molecular map of potato constructed from a segregating cross between *S. phureja* and a *S. tuberosum*×*S. chacoense* hybrid. The map was based on selected single copy DNA probes from tomato[19]. The use of heterologous probes allows a comparison of the chromosomal locations of orthologous loci in both potato and tomato. Solid lines represent regions of the genome in which the linkage order of the markers was found to be conserved between tomato and potato. Hatched boxes represent intervals within which the order of markers is inverted between the two species. Due to the high degree of conservation of linkage order, the 12 linkage groups in potato were assigned to chromosomes by homeology with the tomato genome. Ninety-eight percent of the 130 tomato clones analyzed hybridized with homologous sequences in potato at a stringency of 0.5×SSC, 65°C; and the relative copy numbers of the sequences were conserved in all but two instances[19].

RFLP-BASED BACKCROSS BREEDING: A DETERMINISTIC APPROACH TO VARIETY DEVELOPMENT

Today's crop varieties are a culmination of many years of direct and indirect selection. Improvements have generally taken place in increments as desirable properties from related varieties, races or species have been introduced. This process requires that a cross be made between the source of the desired trait (the donor parent) and the variety that is being improved (the recurrent parent). Segregating progeny obtained from such a cross are mosaics of chromosomal pieces derived from both parents, and may carry not only desirable, but also undesirable traits from the donor parent.

The technique of backcross breeding was developed as a method to recover the more desirable genotype of the recurrent parent after the introduction of gene(s) from the donor parent. In backcross breeding, the hybrid derived from a cross between the donor and recurrent parents is crossed back to the recurrent parent and the progeny are screened for the target character. The selected individual(s) are crossed again to the recurrent parent and the process repeated. After several cycles, plants are obtained that are nearly identical genetically to the recurrent parent, with the exception that genes for the target character have been added.

The introduction of RFLP techniques promises to overcome major limitations of backcross breeding. If the gene(s) to be transferred are marked by tightly linked RFLP markers, segregating populations of plants can be screened at the seedling stage—before the trait is expressed—for the presence of the gene(s) of interest. Since RFLPs can be used to mark QTLs as well as major genes, there are no limitations to the types of characters that can be manipulated by RFLP-based backcross selection. In tomato, for example, genes that effect slight changes in fruit weight, pH and soluble solids levels have been marked with RFLPs and are amenable to rapid transfer between varieties by the backcross method[22,26–27]. Without linked RFLPs it would be extremely difficult and time-consuming to monitor the flow of such genes in a breeding program.

Graphical genotypes and whole genome selection. Because segregating progeny contain chromosomes that are mosaics of chromosomal pieces derived from parental varieties, it is important to consider the complete genome of individuals, as well as the target genes, in a breeding program. Knowledge of the RFLP genotype at one specific locus yields information about the parental origin of alleles at that particular site in the genome. Knowledge of the RFLP genotypes of many linked loci throughout the entire genome yields an estimate of the exact composition of an individual's chromosomes in terms of its parents'. In other words, information about linked points in the genome permits deduction of a continuous genotype[28].

Figure 4 displays a graphical representation of a genotype (graphical genotype) deduced from RFLP data for a randomly selected individual from an interspecific cross of tomato. Note that it is not only possible to see which portions of each set of homologues are derived from each parent, but also the regions in which crossovers took place.

Using graphical genotypes, plants can be selected that not only contain the gene(s) of interest, but also have the highest probability that the rest of the genome will return to that of the recurrent parent with additional crossing. Based on computer simulations in tomato we estimate that by using RFLP-based whole genome selection the recurrent parent genotype can be reconstructed in only three generations of 30 individuals each (Fig. 5A). By contrast,

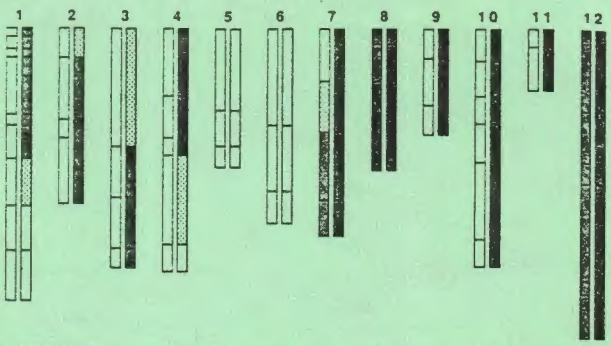


FIGURE 4 Graphical genotype (see text), showing the parental derivation for each part of the genome and the location of all crossover events, calculated for an F2 individual derived from a cross between *L. esculentum* and *L. pennellii*. The genotype for each of 70 RFLP markers, located throughout the genome at a density of approximately one marker every 20 cM, was determined and the numerical genotypic data was converted into a graphical genotype using Hypergene™ software (N.D.Y., and S.D.T., unpub.). White regions indicate segments derived from *L. esculentum*, darkly stippled regions indicate segments derived from *L. pennellii*, lightly stippled regions indicate segments in which crossover events took place.

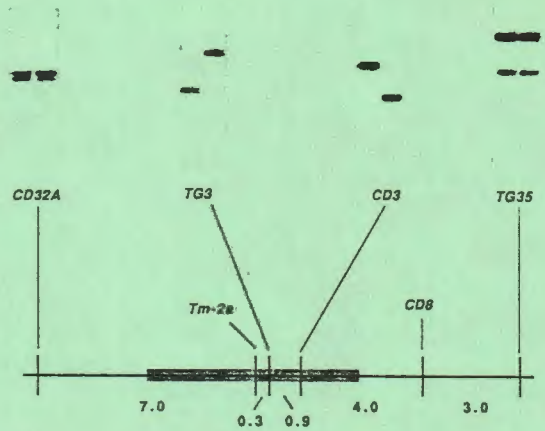


FIGURE 6 Selecting clones near target genes using near-isogenic lines (NILs). Pairs of NILs can be used to rapidly determine whether a genomic clone is located near a target gene, such as *Tm-2a* (resistance to tobacco mosaic virus). The pair of NILs shown (the tomato variety, New Yorker, with and without the *Tm-2a* gene) are identical except for a short introgressed segment (shaded region) surrounding the *Tm-2a* gene, which was introduced by backcross breeding from the tomato relative, *L. peruvianum*. Clones located outside the introgressed segment (*CD32A* and *TG35*) yield identical restriction patterns when probed against Southern blots of the NILs. In contrast, the clones located within the introgressed segment (*TG3* and *CD3*) exhibit hybridizing bands of different molecular weight (RFLPs) for each of the NILs.

this process requires more than six generations with traditional backcross breeding.

A persistent problem in plant breeding is the linkage o desirable genes to undesirable genes. Even after 20 back crosses, one expects to find a sizeable piece (10 cM or ma units) of the donor chromosome still linked to the gen being selected[29]. In most plant genomes 10 cM is enoug DNA to contain hundreds of genes. Therefore, backcross ing results in the transfer, not only of the gene(s) o interest, but also of additional linked genes. This phenom enon has been termed linkage drag[30] and can often resu in a new variety modified for characters other than thos originally targeted.

RFLP-based selection can eliminate, or at least significantly mitigate, linkage drag. Figure 5B shows the average size of the linked segments around a locus being selected in two different backcross schemes. In a traditional backcross program, the linked segments usually remain large for many generations not because recombination has not occurred in these regions, but because there is no effective way to identify recombinant individuals. In classical breeding it is usually only by chance that such recombinants are occasionally selected which contribute to a reduction in the size of the donor segment. With high density RFLP maps it is possible to directly select individuals that have experienced recombination near the gene of interest. In approximately 150 backcross plants there is a 95% chance that at least one plant will have experienced a crossover within 1 cM on one side or the other of the gene being selected. RFLP markers allow unequivocal identification of these individuals[31]. With one additional backcross generation of 300 plants, there would be a 95% chance of a crossover within 1 cM of the other side of the gene, generating a segment surrounding the target gene of less than 2 cM. This would have been accomplished in two generations with RFLP selection, while it would have required, on average, 100 generations without RFLPs (Fig. 5B). It should be apparent that the ability to select for desirable recombinants in a region of interest is a function of the number of markers mapped in that region, as well as the number of plants assayed. As plant RFLP maps become more saturated, the efficiency of selecting recombinants will increase.

Combining backcross breeding with RFLP analysis may make it possible to quickly improve or adjust existing varieties with respect to specific characters. Any gene that can be detected by an RFLP probe (major genes or QTLs) can be rapidly and effectively transferred into other varieties. When a sufficiently large number of genes are so identified, it may be possible to make fine adjustments to important agronomic properties of cultivated varieties, such as maturity date, plant height or quality factors (e.g. fruit pH, solids, grain quality, etc.). Since little additional effort is required to screen with multiple RFLP probes, one could consider adding many genes simultaneously to a variety. Batteries of disease resistance genes could be added in a few generations, as opposed to the many generations required with traditional breeding. The ability to rapidly adjust existing varieties should allow breeders to more quickly respond to market demands, as well as unexpected environmental pressures, such as the appearance of new pathogens.

NEW OPPORTUNITIES FOR UTILIZATION OF EXOTIC GERMPLASM

Primitive races and wild species contain a tremendous amount of genetic variation[32–34]. In fact, the range of variation for a given character is often much greater in exotic germplasm than among cultivated varieties since cultivars are often derived from only one or a few accessions of ancestral species and have been subjected to centuries of selection by man[32].

For some traits, the only source of desirable genes is wild germplasm. In tomato, for example, resistance genes to *Meloidogyne incognita* (root-knot nematode) and tobacco mosaic virus were both introduced from the related species, *L. peruvianum*. Resistance to many other diseases, as well as tolerance to several environmental stresses, are found in other related species, but have not yet been transferred. Similar potential exists in exotic germplasm for most crop plants[34]. In wheat, for example, wild germplasm has been the source of genes for resistance to many different pathogens, including stem yellow rust, powdery mildew and leaf rust. Wild relatives of wheat also possess genes (yet to be transferred) for tolerance to a variety of environmental stresses[33,34].

Breeders have long recognized the potential value of exotic germplasm for variety improvement, but have generally been reluctant to utilize this valuable resource. Reasons for this resistance have been that crosses between wild species and cultivated varieties can be difficult to make and the hybrids are frequently partially or completely sterile. In recent years, tissue and cell culture techniques have circumvented some of these problems[35]. However, even when fertile hybrids are obtained, there is no guarantee that the genes for the character of interest can be successfully moved into cultivated varieties. The most straightforward method of transferring genes from exotic germplasm into cultivated varieties is through backcross breeding. As already discussed, traditional backcrossing is not only time-consuming, but also ineffective in transferring complex genetic characters or those with low heritability. Undesirable effects of linkage drag also are especially pronounced when breeding with exotic germplasm.

Because RFLP markers can be used so effectively to select for individuals with little unwanted donor DNA, they can be used not only to locate the gene(s) controlling a character of interest from exotic germplasm, but also to expedite the transfer of small amounts of foreign chromosomal DNA containing the desired genes into commercial varieties. RFLP methods are now being explored as means of transferring desirable genes from wild species into wheat, rice, tomato and potato[36,37].

In the long run, using RFLPs to access genes from exotic germplasm may prove to be the most significant contribution of this technology to plant breeding. By introducing new genes into the gene pool of cultivated species, RFLP analysis may make it possible to improve characters previously refractory to improvement due to lack of sufficient genetic variation.

DETERMINING GENOME HOMOLOGIES AMONG CROP SPECIES

Plant genetics has grown up largely around important crop species. Because most crop species cannot be hybridized with one another, the genetics of each species has developed into more or less independent, non-overlapping disciplines, each with its own set of devoted researchers and, in some instances, specialized journals.

Despite their seeming independence, a number of crop plants share taxonomic affinities. For example, cabbage, turnips and rape seed all belong to the genus *Brassica*[38], and maize and sorghum are members of the same tribe in the family *Gramineae*[39]. The nightshade family *(Solanaceae)* includes potato, tomato, and pepper. All of these solanaceous crop species share the same basic chromosome number (x=12) yet they cannot be hybridized and, until recently, little was known about chromosome homologies among their genomes. If chromosome content and gene orders were highly conserved among related species, one might entertain the possibility of substituting single chromosomes or chromosomal segments (through somatic hybridization) to combine crop attributes or access genes that might not be available in the normal crossing range of a species.

Using a common set of clones for RFLP mapping, the degree to which the chromosome content and gene order have been conserved in tomato, potato and pepper has recently been determined[14,19]. Single copy clones (cDNA and genomic) from tomato were mapped in segregating populations of each species. Nucleotide sequences of the tomato clones used are sufficiently conserved to hybridize

with homologous sequences in both potato and pepper and comparative linkage maps could therefore be constructed. A comparison of the arrangement of orthologous loci in tomato and potato reveal that no detectable interchanges between chromosomes have occurred. Nine of the chromosomes were found to be entirely homosequential (no detectable rearrangements within chromosomes) and the other 3 chromosomes have apparently experienced paracentric inversions involving one breakpoint near the centromere (Fig. 2). The high degree of linkage conservation between tomato and potato suggests that substitution of chromosomes or chromosome segments from one species to the other might be feasible and raises the possibility of homologous crossing over in conserved regions. In contrast, the RFLP map of pepper is highly rearranged compared with tomato and potato[13]. The prospects for chromosome substitutions between tomato and pepper thus seem less promising.

MAP-BASED GENE CLONING

Methods are well-established for isolating and cloning a gene if the product of that gene is known[40], but unfortunately, the products of many important plant genes are unknown. Disease resistance genes are a pertinent example. Many such genes have been identified by classical genetics, however, the mechanisms by which they act are largely unknown and, thus far, none of their gene products has been definitively identified. Until the genes for disease resistance are isolated and studied with the tools of molecular biology, it seems unlikely that we will fully understand their mode of action. Moreover, if such genes could be isolated and cloned, it might be possible to introduce them, via transformation, into other crop species for which resistance genes are not available.

One approach to cloning plant genes without knowledge of gene product is transposon tagging. This approach has proven useful in cloning genes in species in which transposons are well-characterized[41], and attempts to introduce transposons into plants in which endemic transposons have not been characterized are progressing[42]. With transposon tagging, however, a large number of transposon-mutagenized individuals must be screened (as many as 100,000) and this may limit the technique to genes controlling major visible phenotypes or characters

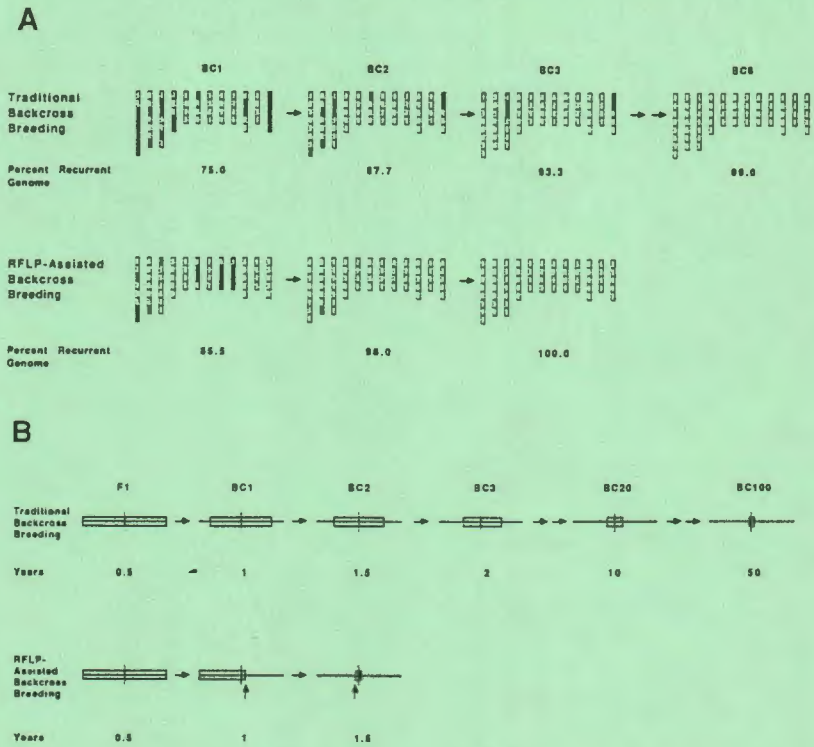


FIGURE 5 Comparison of traditional and RFLP-assisted backcross breeding. (A) Rate of return to recurrent parent genotype in regions of genome unlinked to gene(s) being introgressed. (Top) Traditional backcross breeding. Graphical genotypes were generated for randomly selected individuals from various backcross generations derived from a single BC1 individual by computer simulation. Only one homologue of each of the 12 tomato chromosomes is shown (the other homologue is derived exclusively from the recurrent parent). Darkly stippled regions indicate donor genome segments, lightly stippled regions indicate segments in which cross-overs occurred, and white regions indicate recurrent genome segments. Each interval is 20 cM in length. The numbers beneath each graphical genotype indicate the percent of the genome derived from the recurrent parent. The average number of generations required to return to the recurrent genome, as estimated from 20 independent simulations, was 6.5±1.7 generations. (Bottom) Graphical genotypes of individuals from a simulated RFLP-assisted backcross breeding program showing return to the recurrent parent in only three generations. In each backcross generation, 30 progeny were generated and the best (in terms of percent recurrent parent genome) was used as the parent for the next backcross generations. Graphical genotypes, simulations, and calculations were performed with HyperGene™ software (N.D.Y., and S.D.T., unpub.). (B) Expected linkage drag around a selected gene held heterozygous during backcrossing. (Top) Traditional backcross breeding[29]. (Bottom) RFLP-assisted selection for plants carrying chromosomes with recombination near selected gene[31]. RFLP markers tightly linked to the gene of interest are used to identify individuals with crossovers within 1 cM on one side of the selected gene (vertical arrow) in backcross generation 1 (BC 1). These recombinant individuals are then backcrossed to the recurrent parent and other tightly linked RFLP markers are used to select recombinants within 1 cM on the other side of the gene (vertical arrow) in BC2. The expected number of years to obtain a given level of linkage drag (for a typical crop with a generation time of 0.5 years) is shown below.

that can be screened for in cell cultures. Many traits, such as resistance to root pathogens, are difficult to score or contribute only a fraction of the total phenotype (e.g. QTLs). The genes controlling such traits will likely require alternative gene isolation and cloning strategies.

Map-based cloning offers an independent method for gene isolation. Often referred to as reverse genetics, this approach to cloning genes is based on physical linkage to a cloned RFLP marker(s)[43]. Once tightly-linked RFLP markers are identified, one must "walk" by overlapping clones[44], "hop" by linked libraries[45] or move along the chromosome in some other way from the cloned RFLP to the gene of interest. Reverse genetics has already been used to clone several genes involved in hereditary diseases of humans[42]. There are, however, two major difficulties with the approach of reverse genetics. First, there must be some method for identifying the cloned segment(s) of DNA that actually contain the gene of interest; and second, while a gene of interest and an RFLP may be tightly linked genetically, they may still be physically very far apart. For example in tomato, one cM is roughly equivalent to 550 kilobase pairs of DNA.

DNA transformation techniques are available for several crop plants. The possibility of inserting DNA back into plants provides a functional assay for gene activity and, with the development of large-scale transformation techniques, many clones can potentially be screened for genetic activity[46]. In this regard, determining which DNA segments contain a gene of interest is less of an issue in plants than in other higher eukaryotes.

Moving from a tightly-linked RFLP marker to a gene of interest can be simplified by identifying many DNA markers in the region very near the target gene, thereby minimizing the distance that must be traversed along the genome, as well as providing orientation. In plants a method exists for rapidly detecting markers that are very tightly linked to many different agriculturally important genes. The procedure involves simultaneously testing several genomic clones together as hybridization probes against Southern blots consisting of pairs of near-isogenic lines (NILs) that differ only in the presence or absence of the target gene and a small region of flanking DNA (Fig. 6)[21]. Using this appraoch, it is feasible to screen thousands of clones for those that are near a gene of interest. Moreover, many pairs of NILs have been developed by backcrossing in several different crop species[21], so this technique will be widely applicable. In tomato this strategy has been successfully employed to select clones that are tightly linked to several disease resistance genes[21,36].

SUMMARY AND CONCLUSIONS

Plant breeders have been responsible for impressive gains in crop productivity. More than 50% of increased world food output over the past 50 years can be attributed to improved varieties[47]. Few people doubt that biotechnology will have a significant impact on the future of crop improvement. Some aspects of biotechnology, such as engineering of artificial disease and insect resistance genes[48–50], will augment work by plant breeders and produce characters that previously did not exist. RFLP techniques, on the other hand, will likely be integrated into existing plant breeding programs allowing researchers to access, transfer and combine genes at a rate and with a precision previously not possible. Because of its direct impact on breeding methodology, RFLP probing will likely be one of the first techniques of biotechnology to be incorporated into existing commercial and government breeding programs and may thus be one of the first to have an impact on plant agriculture. Complex or polygenic traits can now be reduced to their individual genetic components, and researchers will have the opportunity to study their individual and combined effects in selected genetic backgrounds. Genetic transfer of polygenes, as well as major genes, will be more readily effected by monitoring linked RFLP markers. As a result, crop varieties may be produced with combinations of characters previously too difficult or time-consuming to produce otherwise.

It is also likely that RFLP mapping will expedite the acquisition of important genes from wild species, which may increase yield, resistance or adaptability to extreme environments. Genetic ties may be established between related crop species through construction of RFLP maps based on common sets of clones. A precedent has already been set by the development of homoeologous maps for tomato, potato and pepper.

Finally, plant geneticists and molecular biologists will increasingly use RFLP maps to locate and potentially clone genes not readily isolated by other procedures. Likely to be included on this list will be genes whose gene products are unknown, like those for disease resistance, as well as genes with small individual effects such as those underlying polygenic traits.

Acknowledgments

Research supported in part by grants from the National Science Foundation, the U.S. Department of Agriculture Competitive Research Grants Organization and the Cornell Biotechnology Program. We would like to thank L. Earle, M. Ganal, and M. Sorrells for reviewing the manuscript and to S. David for typing.

References

1. Allard, R. W. 1960. Principles of Plant Breeding. John Wiley & Sons, Inc., New York.
2. Mendel, G. 1865. Versuche über Pflanzen-Hybriden. Verh. Naturf.-Ver. Brunn IV:3–47 (English translation published as a pamphlet by Harvard University Press, Cambridge, MA, 1925).
3. Emerson, R. A., Beadle, G. W., and Fraser, A. C. 1935. A summary of linkage studies in maize. Cornell Univ. Agr. Exp. Sta. Memoir 180.
4. MacArthur, J. W. 1934. Linkage groups in tomato, J. Genet. 29:123–133.
5. Sax, K. 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552–560.
6. Thoday, J. M. 1961. Location of polygenes. Nature 191:368–370.
7. Tanksley, S. D. and Orton, T. J. 1983. Isozymes in Plant Genetics and Breeding, Parts 1A and 1B. Elsevier, Amsterdam.
8. Tanksley, S. D. 1983. Molecular markers in plant breeding. Plant Mol. Biol. Rep. 1:3–8.
9. Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314–331.
10. Wyman, A. R. and White, R. 1980. A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. U.S.A. 77:6754–6758.
11. Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T. P., Bowden, D. W., Smith, D. R., Lander, E. S., Botstein, D., Akots, G., Rediker, K. S., Gravius, T., Brown, V. A., Rising, M. B., Parkers, C., Powers, J. A., Watt, D. E., Kauffman, E. R., Bricker, A., Phipps, R., Muller-Kahle, H., Fulton, T. R., Ng, S., Schumm, J. W., Braman, J. C., Knowlton, R. G., Barker, D. F., Crooks, S. M., Lincoln, S. E., Daly, M. J., and Abrahamson, J. 1987. A genetic linkage map of the human genome. Cell 51:319–337.
12. Tanksley, S., Miller, J., Paterson, A., and Bernatzky, R. 1988. Molecular mapping of plant chromosomes, p. 157–173. *In:* Chromosome Structure and Function. Gustafson, J. F., and Appels, R., (Eds.). Plenum, New York.
13. Helentjaris, T., King, G., Slocum, M., Siedenstrang, C., and Wegman, S. 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol. Biol. 5:109–118.
14. Tanksley, S. D., Bernatzky, R., Lapitan, N. L., and Prince, J. P. 1988. Conservation of gene repertoire but not gene order in pepper and tomato. Proc. Natl. Acad. Sci. U.S.A. 85:6419–6423.
15. McCouch, S. R., Kochert, G., Yu, Z. H., Wang, Z. Y., Khush, G. S., Coffman, W. R., Tanksley, S. T., 1988. Molecular mapping of rice chromosomes. Theor. Appl. Genet. In press.
16. Chang, C., Bowman, J. C., DeJohn, A. W., Lander, E. S., and Meverowitz, E. S. 1988. Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. 85:6856–6860.
17. Helentjaris, T. 1987. A genetic linkage map for maize based on RFLPs. Trends in Genet. 3:217–221.
18. Landry, B. S., Kesseli, R. V., Farrara, B., and Michelmore, R. W. 1987. A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and mor-

phological markers. Genetics **116**:331–337.
19. Bonierbale, M. W., Plaisted, R. L., and Tanksley, S. D. 1988. RFL maps based on a common set of clones reveal modes of chromosom evolution in potato and tomato. Genetics **120**:1095–1103.
20. Zeven, A. C., and van Harten, A. M. 1979. (Eds.) Proceedings of th Conference. Broadening the Genetic Base of Crops. Pudoc, Wagenin gen, Netherlands.
21. Young, N. D., Zamir, D., Ganal, M. W., and Tanksley, S. D. 1988. U of isogenic lines and simultaneous probing to identify DNA marke tightly linked to the *Tm-2a* gene in tomato. Genetics **120**:579–585.
22. Paterson, A. H., Lander, E. S., Hewitt, J. D., Peterson, S., Lincol S. E., and Tanksley, S. D. 1988. Resolution of quantitative traits in Mendelian factors by using a complete RFLP linkage map. Natu **335**:721–726.
23. Murray, M. Agrigenetics Research Corporation (personal communic tion).
24. Falconer, D. S. 1960. Introduction to Quantitative Genetics. Rona Press Co., New York.
25. Lander, E. S., and Botstein, D. 1989. Mapping Mendelian facto underlying quantitative traits using RFLP linkage maps. Genet **121**:185–199.
26. Osborne, T. C., Alexander, D. C., and Fobes, J. F. 1987. Identificatio of restriction fragment length polymorphisms linked to genes contro ling soluble solids content in tomato fruit. Theor. Appl. Gene **73**:350–356.
27. Tanksley, S. D., and Hewitt, J. 1988. Use of molecular markers breeding for soluble solids content in tomato—a re-examinatio Theor. Appl. Genet. **75**:811–823.
28. Young, N. D. and Tanksley, S. D. 1988. Restriction fragment leng polymophism maps and the concept of graphical genotypes. Theo Appl. Genet. In press.
29. Stam, P., and Zeven, A. C. 1981. The theoretical proportion of donor genome in near-isogenic lines of self-fertilizers bred by bac crossing. Euphytica **30**:227–238.
30. Zeven, A. C., Knott, D. R., and Johnson, R. 1983. Investigation linkage drag in near isogenic lines of wheat by testing for seedli reaction to races of stem rust, leaf rust and yellow rust. Euphyt **32**:319–327.
31. Young, N. D. and Tanksley, S. D. 1989. RFLP analysis of the size chromosomal segments retained around the *Tm-2* locus of toma during backcross breeding. Theor. Appl. Genet. In press.
32. Hawks, J. G. 1977. The importance of wild germplasm in pla breeding. Euphytica **26**:615–621.
33. Goodman, R. M., Hauptli, H., Crosswav, A., and Knaut, V. C. 198 Gene transfer in crop improvement. Science **236**:48–54.
34. Stalker, H. T. 1980. Utilization of wild species for crop improveme Adv. Agron. **33**:111–147.
35. Gleba, Y. Y. and Sytnik, K. M. 1984. Protoplast Fusion and Gene Engineering in Plants. Springer-Verlag, Berlin.
36. Gale, M. (personal communication).
37. Tanksley, S. D., *et al.* unpublished data.
38. McNaughton, I. H. 1976. Turnip and relatives. p. 45–48. *In:* Evo tion of Crop Plants. Simmonds, H. W. (Ed.) Longman, New York.
39. Smith, J. P., Jr. 1977. Vascular Plant Families. Mad River Pre Eureka, CA.
40. Maniatis, T., Frisch, E. F., and Sambrook, J. 1982. Molecular Cloni Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. Federoff, N. V., Furtek, D. B., and Nelson, O. E., Jr. 1984. Cloning the bronze locus in maize by a simple and generalizable proced using the transposable controlling element *Activator (Ac)*. Proc. N Acad. Sci. U.S.A. **81**:3825–3829.
42. Baker, B., Schell, J., Lörz, H., and Federoff, N. 1986. Transposition the maize controlling element "Activator" in tobacco. Proc. Natl. Ac Sci. U.S.A. **83**:4844–4848.
43. Orkin, S. H. 1986. Reverse genetics and human disease. Cell **47**:8- 850.
44. Steinmetz, M., Minard, K., Harvath, S., McNicholas, J., Srenlinger Wake, C., Long, E., Mach, B., and Hood, L. 1981. A molecular map the immune response region of the major histocompatibility comp of the mouse. Nature **300**:35–42.
45. Poutska, A., Pohl, T., Barlow, D. P., Zehetner, G., Craig, A., Michi F., Ehrich, E., Frischauf, A.-M., and Lehrach, H. 1986. Molecu approaches to mammalian genetics. Cold Spring Harbor Symposi on Quantitative Biology **51**:131–139.
46. Klee, H. J., Havford, M. B., and Rogers, S. B. 1988. Gene rescu plants: A model for "shotgun" cloning by retransformation. **210**:2 287.
47. Fehr, W. R. (Ed.) Genetic Contributions of Yield Gains of Five M Crop Plants. Crop Science Society of America, Madison, WI.
48. Abel, P., Nelson, R. S., De, B., Hoffman, N., Rogers, S. G., Fra R. T., and Beachy, R. N. 1986. Delay of disease developmen transgenic plants that express the tobacco mosaic virus coat pro gene. Science **232**:738–743.
49. Loesch-Fries, L., Merlo, D., Zinnen, T., Burhop, L., Hill, K., Kr D., Jarvis, N., Nelson, S., and Halk, E. 1987. Expression of al mosaic virus RNA 4 in transgenic plants confers virus resista EMBO J. **6**:1845–1851.
50. Vaeck, M. et al. 1987. Transgenic plants protected from insect at Nature **328**:33–37.
51. Schroeder, W. T., Provvidenti, R., and Robinson, R. W. 1967. Inc tion temperature and virus strains important in evaluating to genotypes for tobacco mosaic virus reactions. Tomato Genet. C Rep. **17**:47–48.

Reprint Series
22 August 1997, Volume 277, pp. 1063–1066

SCIENCE

Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild

Steven D. Tanksley and Susan R. McCouch

Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild

Steven D. Tanksley and Susan R. McCouch

Nearly a century has been spent collecting and preserving genetic diversity in plants. Germplasm banks—living seed collections that serve as repositories of genetic variation—have been established as a source of genes for improving agricultural crops. Genetic linkage maps have made it possible to study the chromosomal locations of genes for improving yield and other complex traits important to agriculture. The tools of genome research may finally unleash the genetic potential of our wild and cultivated germplasm resources for the benefit of society.

As we head into the next millennium, the world faces a greater demand on agricultural output than at any time in history. Despite efforts to curb birthrates, the Earth's human population is expected to rise to 8.9 billion by the year 2030, corresponding to more than a 50% increase from the current population of 5.7 billion (*1*). In the past, we have met the demand for increased agricultural productivity by a combination of genetic improvements, greater farming inputs (fertilizers, pesticides, and water), and cultivation of more land. With dwindling freshwater reserves and petroleum resources (on which fertilizers and pesticides are based) and increased problems caused by agricultural pollution, we can hardly expect to increase or even maintain our current levels of agricultural inputs. Similarly, much existing farmland is falling victim to urban expansion, and it is unlikely that new farmland will become available in the near future. That leaves the genetic improvement of crops as the most viable approach by which food production can attempt to keep pace with the anticipated growth of the human population. For the genetic approach to succeed, we must harness the wealth of genetic variation provided by nature and currently warehoused in our seed repositories. Until now we have been only modestly successful in utilizing these resources for plant improvement. New findings from genome research indicate that there is tremendous genetic potential locked up in seed banks that can be released only by shifting the paradigm from searching for phenotypes to searching for superior genes with the aid of molecular linkage maps.

The Narrow Genetic Base of Crop Plants

Today, modern agriculture—and, for that matter, human existence—is dependent on the cultivation of a few highly productive crop species. These food crops were first domesticated from wild species about 10,000 years ago during the transition from nomadic hunter-gatherers to life in agrarian societies. Considering that flowering plants first evolved over 150 million years ago, crop plants as we know them have existed for the mere blink of an evolutionary eye.

Although the exact series of steps by which plants were domesticated is unknown, it is likely that strong selection pressure exerted by humans on the genetic diversity found in the wild resulted in rapid and radical changes in plant species (*2*). Certain traits, such as nonshattering of seeds, compact growth habit, or loss of germination inhibition, would have been selected by early agriculturists (*3*). Selective propagation of lines containing these favorable mutations would have resulted in a progressive narrowing of the genetic base of subsequent populations (Fig. 1).

Following domestication, the genetic variation in crop plants has continued to be reduced by another force—modern plant breeding. Over the past century, the development and successful application of plant-breeding methodologies has produced the high-yielding crop varieties on which modern agriculture is based. Yet, ironically, it is the plant-breeding process itself that threatens the genetic base on which breeding depends. Because new varieties are usually derived from crosses among genetically related modern varieties, genetically more variable, but less productive, primitive ancestors are excluded. Soybeans and wheat are good examples of crops with very narrow genetic bases. Virtually all modern U.S. soybean varieties can be traced back to a dozen strains from a small area in northeastern China, and the majority of hard red winter wheat varieties in the United States originated from just two lines imported from Poland and Russia (*4*, *5*).

The limited genetic diversity of crops renders them more vulnerable to disease and insect epidemics and jeopardizes the potential for sustained genetic improvement over the long term (*5*). This risk was brought sharply into focus in 1970 with the outbreak of Southern corn leaf blight. This disease drastically reduced corn yields in the United States and was attributed to extensive use of a single genetic male sterility factor that, unfortunately, was genetically linked to disease susceptibility (*6*).

World Germplasm Resources

Over six decades ago, Vavilov (*2*) first called attention to the potential of crop relatives as a source of genes for improving agriculture. It was this promise that motivated the establishment of gene banks, living seed collections focused primarily on races and species that are closely related to crop plants (collectively referred to as "exotics"). In addition, a report commissioned by the National Academy of Sciences in response to the Southern corn leaf blight disaster recommended placing much greater emphasis on collecting and preserving the genetic diversity in crop species, especially that present in the remaining populations of wild ancestors (*7*). One result of that report was the formation of the International Plant Genetic Resources Institute, which was charged with overseeing worldwide efforts in plant collection and preservation (*8*).

Worldwide, there are more than 700

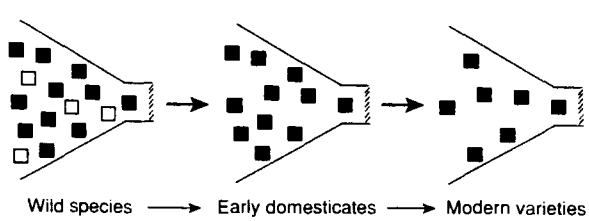


Fig. 1. Genetic bottlenecks imposed on crop plants during domestication and through modern plant-breeding practices. Boxes represent allelic variations of genes originally found in the wild, but gradually lost through domestication and breeding. Such lost alleles can be recovered only by going back to the wild ancestors of our crop species.

The authors are in the Department of Plant Breeding, 252 Emerson Hall, Cornell University, Ithaca, NY 14853, USA.

documented seed collections holding an estimated 2.5 million entries including many exotics (Table 1) (8). The United States alone spends about $20 million per year on germplasm acquisition and preservation, a testament to the importance of these collections (Fig. 2). While it is important to collect and preserve genetic variation in seed banks, these activities are not sufficient to ensure the future productivity of agriculture. The establishment and maintenance of seed banks must be coupled with the ability to actively utilize the materials in those collections. In this regard, we have fallen far short of expectations. While there is a general belief that genes useful for improving crops are contained in these seed banks, the reality is that plant scientists have been largely unsuccessful in finding and extracting such genes. The vast majority of the accessions in gene banks make no contribution to modern varieties, particularly with respect to complex traits such as yield and nutritional quality. As a result, crop improvement is still practiced on a narrow genetic base. DNA-based genetic comparisons of modern cultivars and ancestral wild types found in germplasm banks reflect this situation (Fig. 3).

As we look toward the future, we must reflect upon the rationale for collecting and maintaining germplasm and reexamine the ways in which we are using these genetic resources. Recent applications of genome mapping suggest that the genetic diversity stored in germplasm banks can be utilized with a much higher level of efficiency than previously imagined. However, if we are to discover the true potential of the seeds carefully locked away in gene banks, we must change the way we think about our germplasm resources and be willing to approach the utilization of these living collections in new and innovative ways.

Strategies for the Evaluation and Utilization of Germplasm

The old paradigm: Looking for the phenotype. The traditional approach to the utilization of exotic germplasm is to screen entries from a gene bank for a clearly defined character, recognizable in the phenotype (physical appearance). Once a line with the desired characteristic has been identified, it is crossed with an elite cultivar in order to introduce the genes from the exotic donor into the cultivated type. This approach works well when the trait of interest is controlled by one or a few genes. For example, wild germplasm has been used with great success in breeding for simply inherited resistance to diseases and insects (8). The resistance characters selected by means of this approach are almost always conditioned by single, dominant genes that can be readily transferred into elite cultivars through established breeding procedures. Generally, resistance screening requires inoculating hundreds or thousands of entries with a given pathogen or insect. One or more highly resistant lines are identified and used as donors of the trait in subsequent crossing experiments. Though effective for certain specific characters, only a small proportion of the genetic variation inherent in exotic germplasm will ever be exploited for crop improvement as a result of this strategy.

Most traits important to agriculture, such as yield, are conditioned not by single genes, but by many genes. The yield of wild and unadapted accessions found in germplasm banks is invariably much lower than that of modern elite cultivars. For example, in rice, an accession of the wild ancestor *Oryza rufipogon* yields only a fraction of that of a modern Chinese hybrid (Fig. 4). Crop domestication and plant breeding have been extremely successful at increasing the frequency of beneficial alleles for yield at many loci. As a result, breeders have persisted in making crosses among closely related, high-yielding varieties, unable to rationalize a search for yield-enhancing genes in low-yielding ancestral types. Yet, considering how many genes are likely to influence yield, it is unlikely that modern cultivars have the best alleles for yield at all yield-related loci. Many beneficial alleles have undoubtedly been left behind because of the bottlenecks imposed by domestication coupled with years of modern breeding and selection within adapted gene pools. Thus, although wild and exotic germplasm is perceived to be a poor bet for the improvement of most traits

Table 1. Estimated number of seed bank entries worldwide for selected crops [reprinted from (8)].

Crop	Entries	Collections of 200+	Percent wild species
Wheat	410,000	37	60
Rice	215,000	29	10
Maize	100,000	34	15
Soybean	100,000	28	30
Potato	42,000	28	40
Tomato	32,000	28	70
Cotton	30,000	12	20

Fig. 2. Cryopreservation of seed stocks at USDA–Agricultural Research Service (ARS) National Seed Storage Laboratory, Fort Collins, Colorado. [Photo courtesy of USDA-ARS Information Services]

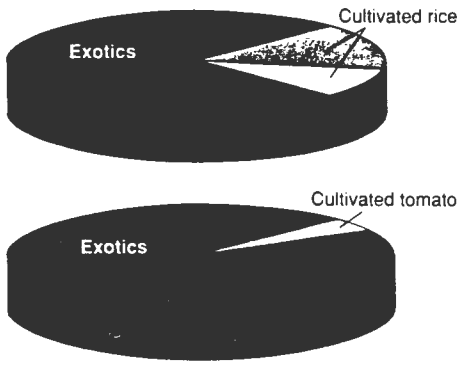


Fig. 3. Pie diagrams depicting proportion of genetic variation (as measured through DNA polymorphisms) found in cultivated rice [*Oryza sativa* spp. indica (blue) and japonica (yellow)] (9) and cultivated tomato (*L. esculentum*) (10) versus their wild relatives (exotics, green). Exotics include only those lines that cross readily with their domesticated counterparts.

Fig. 4. (Left) Wild rice species *O. rufipogon* from Malaysia. [Photo courtesy of C. Martinez, Centro Internacional de Agricultura Tropical] (Right) Modern rice variety from China. Although wild species are low yielding, they contain genes that can significantly increase the yield of modern rice varieties and provide much needed enrichment of the domestic gene pool.

based on phenotypic examination, it is quite possible that some favorable genes (alleles) lie buried amidst the thousands of accessions maintained in gene banks which, if they could be found, might be of great value to crop improvement. Implementing strategies for finding those genes requires a major shift in the paradigm for using our genetic resources.

The new paradigm: Looking for the genes. Genetic linkage maps based on molecular markers have now been developed for most major crop species, as well as for a number of minor species (http://probe.nalusda.gov). These maps have been important in a number of applications in plant science including the localization of genetic loci that condition agronomically important traits, positional gene cloning, comparative mapping, and marker-assisted selection in breeding (*12*). Perhaps their most profound impact has been in the study of complex, or quantitatively inherited, traits. Molecular linkage maps have made it possible to identify, map, and study the effects of the individual loci that control a quantitatively inherited trait (quantitative trait loci, or QTLs) (*13*). Studies from plants (mainly crop plants) have led to several interesting discoveries. Before the QTL era of quantitative genetics, it was assumed that complex traits were determined by a large number of genes of relatively small and equal effect (*14*). However, QTL analysis has revealed that while most complex traits are controlled by a number of loci, the effects of those loci are not equal. Often a substantial portion of the genetic variation in a population can be explained by a few QTLs of moderately large effects. The second significant discovery was that the phenotype of a plant is, at best, only a modest predictor of its genetic potential. For example, if one line of rice is high yielding and another low yielding, one might assume that the high-yielding type possesses most, if not all, of the genes for high yield and that the low-yielding parent has little or nothing to offer in this regard. However, when populations derived from such crosses are examined with molecular markers and the loci controlling yield are identified, a much different picture emerges. While the high-yielding line often does contain a great number of positive alleles at the loci associated with yield, there are almost always some loci for which the inferior parent contributes a superior allele (*11*, *15*).

The implications of these findings are profound with respect to germplasm utilization. They suggest that using phenotypic evaluation to determine the breeding value of an accession is likely to be misleading, especially with respect to quantitative traits. Thus, we have been screening germplasm in a way that fails to expose its potential. A corollary is that exotic germplasm is a likely source of new and valuable genes capable of increasing yield and other complex traits important to agriculture and that molecular linkage maps will enable us to find them. The paradigm needs to shift away from selecting potential parents on the basis of phenotype to evaluating them directly for the presence of useful genes. The tools that make such an analysis possible are molecular maps and the integrative power of QTL analysis.

Exploitation of Genetic Resources Based on Molecular Maps

Rice and tomato are representatives of the two major classifications of flowering plants. A monocot and member of the grass family, rice is one of a large number of staple cereal crops including wheat, barley, oat, maize, sorghum, and millet. Tomato is a dicot, a classification it shares with a varied assortment of other crops such as potato, cassava, soybeans, sugarbeet, cotton, lettuce, and sunflower.

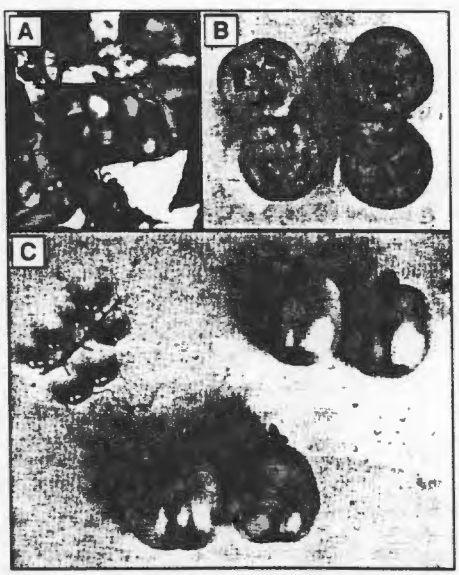


Fig. 5. (**A**) Wild tomato species *L. hirsutum* from Peru that produces small, inedible fruit that does not turn red upon ripening. (**B**) (Left) Fruit from modern processing tomato cultivar E6203. (Right) Fruit from nearly isogenic line (NIL) into which QTL for increased red pigment has been transferred from *L. hirsutum* by the advanced backcross QTL method (*17*). (**C**) (Top left) Fruit from *L. pimpinellifolium* from Peru that produces small berries typical of most fruit-bearing wild species. (Top right) Fruit from modern processing tomato cultivar E6203. (Bottom center) Fruit from NIL into which QTL for increased fruit size has been transferred from *L. pimpinellifolium* by the advanced backcross QTL method. Fruit of this NIL are significantly larger (~10%) than the original E6203 variety (*19*).

A gene-based approach to screening exotic germplasm has been tested in these two species. The strategy involves the use of molecular linkage maps and a breeding technique, referred as to the advanced backcross QTL method, that allows a subset of alleles from the wild or exotic plant to be examined in the genetic background of an elite cultivar (*16*). The molecular linkage map is used to identify the chromosomal position of "wild" alleles that have been transmitted into the progeny, to determine which of the wild species introgressions are associated with superior performance of lines, and to purify lines so that they contain only a specific "wild QTL" in an elite genetic background. The prediction is that the modified elite lines will perform even better than the original elite cultivar.

In tomato, where the most extensive experiments have been conducted, lines have been created that contain specific QTLs from the wild species *Lycopersicon hirsutum* and that outperform the original elite variety by 48, 22, and 33% for yield, soluble solids content, and fruit color, respectively (*17*). Moreover, the performance of these lines has been confirmed under different growing conditions around the world (*17*). The magnitude of these improvements is substantial if we consider that the normal yearly improvement for these traits achieved through traditional breeding is less than 1% per year (*18*). The dramatic improvement in red fruit color (attributable to the pigment lycopene) is especially remarkable considering that the wild tomato lacks an active enzyme for the last step in the pathway and cannot synthesize lycopene (fruits remain green even when ripe) (Fig. 5A). Apparently, the wild tomato contains genes (alleles) that can enhance earlier steps in the biosynthetic pathway leading to lycopene which, when combined with an active form of the gene for lycopene synthesis from cultivated tomato, leads to even higher levels of pigment production in the interspecific offspring (Fig. 5B). Similarly, fruit size has been increased in cultivated tomato lines by the introduction of genes, identified through molecular mapping from the small-fruited ancestor *L. pimpinellifolium* (*19*) (Fig. 5C). These results serve to underscore the point that exotic germplasm often contains genes that are capable of improving traits important to humans. This could not have been predicted from phenotypic evaluation alone, nor could the useful genes be readily discovered without the aid of QTL mapping.

The superior lines described above were created in a two-step process. The QTLs were discovered in an advanced backcross generation (BC2) and, on the basis of that information, new lines were selected con-

taining specific QTL alleles from the wild species. The odds of such favorable genetic combinations occurring by chance and without marker-assisted selection are extremely low.

Thus far in tomato, the genomes of four wild species have been screened by means of the advanced backcross QTL method. While some putatively allelic QTLs have been identified in comparisons among these species, more than 50% of the beneficial QTLs discovered in each experiment appear to be unique to the species tested (*20*). This suggests that continued sampling of wild germplasm is likely to be rewarded with new gene discoveries. It also raises the possibility of creating new crop varieties with a mixture of beneficial genes coming from several different wild ancestors—genotypes that can be efficiently created through molecular breeding.

In rice, conventional hybrid varieties developed at the National Hybrid Rice Research Center (NHRRC) in China outyield the best Chinese inbreds and are among the most productive rice varieties in the world. Experiments were recently conducted to determine whether genes from the low-yielding wild ancestor *O. rufipogon* could significantly enhance yields of one of China's most productive rice hybrids. When the advanced backcross method was used to examine alleles from the wild species in the genetic background of the elite Chinese hybrid, two QTLs were identified that each increased yield ~17% compared with the original hybrid. Moreover, the wild alleles identified in this study appear to be free of many of the negative effects often associated with yield enhancement in cultivated germplasm (*21*). Similar results have been observed in three other experiments conducted independently in Korea and Colombia (*21*).

The results from both tomato and rice indicate that exotic germplasm does contain many new and useful genes that can significantly enhance agricultural production, even for complex traits like yield. They also demonstrate that the phenotype of exotic germplasm is a poor predictor of its genetic potential and that, until we begin applying molecular mapping techniques like the advanced backcross QTL method that allow us to identify genes, it is unlikely that we will realize the benefits to agriculture of most genetic resources.

Strategies for Sampling Exotic Germplasm

Considering the large number of exotic accessions in seed banks and the limited amount of time and resources available, how do we decide which accessions to sample in order to maximize the chance of finding new and useful genes? Realistically, this question can be answered only after further studies have been conducted involving molecular maps and exotic germplasm. However, one principle seems likely to emerge: The sampling of exotic germplasm should emphasize the genetic composition rather than appearance of exotic accessions. With the use of DNA profiles, the genetic uniqueness of each accession in a seed bank, relative to all other accessions, can be determined and quantified (*22*). Accessions with DNA profiles most distinct from that of modern germplasm are likely to contain the greatest number of novel alleles. It is in these accessions that one is most likely to uncover the largest number of unique and potentially agronomically useful alleles. The DNA profiling necessary to make such sampling decisions has already been accomplished or is underway in most crops (*22*).

In tomato and rice, exotic germplasm accessions for use with the advanced backcross QTL method have been selected on the basis of genetic uniqueness with the following two criteria: (i) the exotic accession had to possess a significant number of unique DNA polymorphisms (throughout the genome) relative to modern cultivars and (ii) each new exotic had to be genetically dissimilar (on the basis of DNA profiling) to previously sampled exotics. This strategy has resulted in the sampling of accessions representing the broadest possible spectrum of wild species and races present in the seed banks. With each sampling of a new exotic accession, a high proportion (~50%) of new and useful QTL alleles has been identified (*20*).

As the results accumulate from molecular mapping studies in crop plants, patterns are likely to emerge concerning the chromosomal positions of key loci controlling yield and other important traits. Once this information becomes available, one can imagine sampling exotic germplasm for targeted areas of the genome through backcross breeding and the generation of allelic series in nearly isogenic lines that could then be evaluated agronomically. Such targeted studies can be conducted more rapidly and at less expense than whole-genome studies. Eventually, the genes underlying yield-enhancing QTLs will be cloned, and at that time we can imagine that methods like the polymerase chain reaction will allow allelic variants at these loci to be selectively amplified and sequenced from large numbers of exotic accessions. We can expect that such gene-based searches will ultimately be extended beyond the realm of sexually compatible species so that homologous genes from other genera and families can be tested for optimum performance by a transgenic approach.

Molecular Breeding and Germplasm Utilization in the Future

We are fortunate to be living at a time when genetic engineering holds much promise for modifying crop performance. However, most of the advances thus far in genetic engineering have been directed toward traits other than yield, largely because of the complexity of this trait. The wide repertoire of genetic variants created and selected by nature over hundreds of millions of years is contained in our germplasm banks in the form of exotic accessions. More than 50 years ago, Vavilov predicted the value to agriculture of collecting and maintaining the wild relatives of crop plants in gene banks. Owing to the advent of molecular mapping and the ability to scan the genomes of wild species for new and useful genes, we may now be in a position to unlock the genetic potential of these germplasm resources.

REFERENCES AND NOTES

1. L. R. Brown, *State of the World* (Norton, New York, 1994).
2. N. I. Vavilov, in *The New Systematics*, J. Huxley, Ed. (Clarendon, Oxford, 1940), pp. 549–566.
3. J. R. Harlan, *Crops and Man* (American Society of Agronomy, Madison, WI, 1975).
4. D. N. Duvick, *Ann. N.Y. Acad. Sci.* **287**, 86 (1977).
5. J. R. Harlan, in *Plant Genetic Resources: A Conservation Imperative*, C. W. Yeatman, D Kafton, G. Wilkes, Eds. (Westview, Boulder, CO, 1987), pp. 111–129.
6. A. J. Ullstrop, *Agric. Handb.* **199**, 1 (1978).
7. Anonymous, *Genetic Vulnerability of Major Crops* (National Academy of Sciences, Washington, DC, 1972).
8. D. L. Plunknett *et al.*, *Gene Banks and the World's Food* (Princeton Univ. Press, Princeton, NJ, 1987).
9. Z. Y. Wang, G. Second, S. D. Tanksley, *Theor. Appl. Genet.* **83**, 565 (1992).
10. J. C. Miller and S. D. Tanksley, *ibid.* **80**, 437 (1990).
11. J. Xiao *et al.*, *Nature* **384**, 223 (1996).
12. R. L. Phillips and I. K. Vasil, *DNA-based Markers in Plants* (Kluwer, Boston, 1994); A. H. Paterson, *Genome Mapping in Plants* (Academic Press, Austin, TX, 1996).
13. S. D. Tanksley, *Annu. Rev. Genet.* **27**, 205 (1993).
14. R. Falconer, *Introduction to Quantitative Genetics* (Ronald, New York, 1993).
15. M. C. DeVicente and S. D. Tanksley, *Genetics* **134**, 585 (1993); Y. Eshed and D. Zamir, *Theor. Appl. Genet.* **88**, 891 (1994).
16. S. D. Tanksley and J. C. Nelson, *Theor. Appl. Genet.* **92**, 191 (1996).
17. D. Bernacchi *et al.*, in preparation.
18. S. Grandillo *et al.*, in preparation.
19. S. D. Tanksley *et al.*, *Theor. Appl. Genet.* **92**, 213 (1996).
20. T. F. Fulton *et al.*, *ibid.*, in press.
21. S. R. McCouch and J. Xiao, personal communication.
22. S. M. Brown and S. Kresovich, in *Genome Mapping in Plants*, A. H. Paterson, Ed. (Landes, New York, 1996), pp. 85–93; S. Smith and T. Helentjaris, *ibid.*, pp. 95–110; A. Karp *et al.*, *Nature Biotechnol.* **15**, 625 (1997).

23. We thank A. Frary and E. Earle for many helpful comments. Supported by grants from the U.S. Department of Agriculture (USDA) National Research Initiative–Plant Genome Program, the Binational Agricultural Research and Development Fund, and The Rockefeller Foundation.

where 80–90% of the ozone is lost each spring (September and early October). In the absence of temperature and volcanic aerosol variations, which affect heterogeneous chemical ozone depletion[4,5], the rate of ozone loss in September (*d* in the figure) should vary solely with halogen levels. However, this rate has an interannual variability of about ±5 DU day^{-1}, or 15% of the maximum value expected in about 2000. This variability may be related to transport variations associated with the quasi-biennial oscillation (QBO)[6]. It may also be related to the polar vortex being displaced from the Pole, which can result in South Pole air parcels being subjected to sunlight and enhanced photochemistry earlier than normal.

The amount of ozone remaining in the 12–20-km altitude interval on 15 September (S15) has shown a consistent downward trend during the past 10 years (*c* in the figure), with a variability of roughly 7 DU, or 10% of the minimum value expected in about 2000. The smaller variabilities of S15 and the ozone-loss rate are probably related to the relative stability of the vortex during the mid-September period. A comparison with balloon measurements made at the South Pole in 1971 (ref. 7) is shown in the figure, indicating that the beginning of the reduction in S15 is consistent with that of total ozone, possibly as early as 1970. The latter possibility was also identified in the daily October minimum total ozone at Halley, Antarctica[8].

Owing to their lower variability, the ozone-loss rate and S15 are likely to be suitable early ozone-healing indicators. Through continued monitoring of the ozone profile at the South Pole, it is estimated that detection of the turnaround should be possible by the year 2008, about 10 years earlier than seems possible by monitoring total ozone alone. This estimate obviously assumes that the downward trend in the total equivalent-chlorine content of the atmosphere continues. If there were a large volcanic eruption in future, for example, the resultant aerosol particles would augment polar stratospheric cloud-related heterogeneous chemistry and affect Antarctic ozone in the lower stratosphere[5]. But these particles would probably not increase ozone loss variability more than the Pinatubo eruption in 1992 and 1993. In addition, future downward stratospheric temperature trends during winter and spring, as have occurred over Halley only in October and November[8], would increase the occurrence of polar stratospheric clouds and prolong recovery. Even so, this is not expected to be a major factor in the next 10 years.

D. J. Hofmann

NOAA Climate Monitoring and Diagnostics Laboratory,
325 Broadway,
Boulder, Colorado 80303, USA
e-mail: dhofmann@cmdl.noaa.gov

1. *Scientific Assessment of Ozone Depletion: 1994* World Meteorological Organization, Global Ozone Research and Monitoring Project Report No. 37 (Geneva, 1995).
2. Montzka, S. A. *et al. Science* **272**, 1318–1322 (1996).
3. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
4. Solomon, S. *Nature* **347**, 347–354 (1990).
5. Hofmann, D. J. *et al. Geophys. Res. Lett.* **22**, 2493–2496 (1995).
6. Hofmann, D. J. *Nature* **382**, 129 (1996).
7. Oltmans, S. J., Hofmann, D. J., Komhyr, W. D. & Lathrop, J. A. in *Proc. Quadrennial Ozone Symp. 1992* NASA CP-3266, 578–581 (1994).
8. Jones, A. E. & Shanklin, J. D. *Nature* **376**, 409–411 (1995).

Genes from wild rice improve yield

SIR — Although the wild relatives of crop species are often inferior to modern cultivars, recent evidence suggests that they may contain genes capable of improving both the yield and quality of modern varieties[1]. However, these favourable genes are often masked by the effects of other deleterious genes. Molecular-genetic maps can be used to exploit the genetic potential of wild species for the improvement of yield and quality in modern plant cultivars[2].

We have tested this strategy in rice by screening a single accession of the wild species, *Oryza rufipogon*, a weedy relative of cultivated rice, for genes capable of increasing the yield of one of China's highest-yielding cultivars. *O. rufipogon*, like other wild relatives of rice, has never before been exploited for the genetic improvement of rice yields.

The world's annual rice production will have to increase 70% by the year 2030 to keep up with the demands of a growing population[3]. In the past, increases in production of this staple food were largely achieved by expanding the area under cultivation, increasing fertilizer inputs and using chemical pest control, but none of these options is realistic today. Future productivity increases will require raising the genetic yield potential of the species, but the yield potential of modern rice varieties has remained stationary for many years.

V/64 is one of the top-performing F_1 hybrid rice varieties in China and is derived from a cross between the two inbred lines V20A and Ce64. We used a Malaysian accession of *O. rufipogon* (IRGC 105491) as the male in a cross to V20A, a cytoplasmic male sterile line, to produce an interspecific F_1 hybrid which expressed strong vegetative vigour. We crossed the F_1 with V20B (maintainer line of V20A, having the same nuclear genome as V20A) to generate 52 BC_1 plants which were grown in the field in the summer of 1993.

The best 10 BC_1 plants, selected for desirable plant type, maturity and fertility, were backcrossed a second time to V20B to generate more than 3,000 BC_2 plants. From these, we selected a subset of 300 and crossed them with Ce64 to generate 300 BC_2 test-cross families. Each of these families should have the genetic constitution of the original V/64 hybrid variety, with the exception of small substitutions of chromosome segments from *O. rufipogon* (on average, each BC_2 test-cross line contained 5% *O. rufipogon*-derived DNA).

We evaluated these BC_2 test-cross families, together with the V/64 hybrid, V20B and *O. rufipogon* parents, for yield in summer 1994 at the China National Hybrid Rice Research and Development Center, where most of China's hybrid rice varieties are developed. We planted three-row plots with 11 plants per row in a randomized complete block design with two replications, and evaluated grain yield per plot.

The figure shows the distribution of grain yield observed for the BC_2 test-cross families and the parental controls. *O. rufipogon* plots were among the lowest performers, yielding 55% less than the V/64 hybrid control and possessing fewer grains per plant and a smaller 1,000-grain weight. This was expected given that *O. rufipogon* is not a cultivated species.

Most of the BC_2 test-cross families also

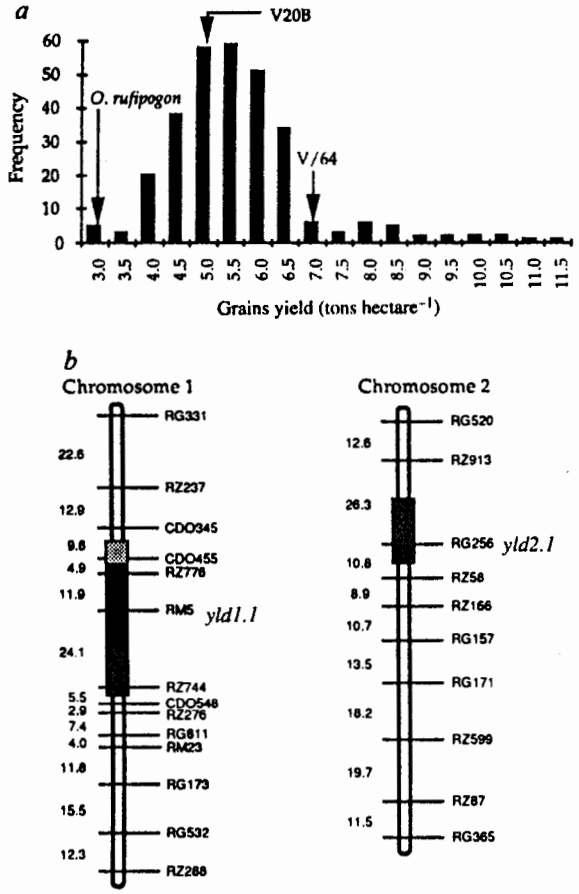
CHARACTERISTICS OF POSITIVE WILD QTL ALLELES ASSOCIATED WITH YIELD

				V/C class		($^1/_2$V/C + $^1/_2$R/C) class		R/C class		
QTL	Marker	Chr	*P*	PM	*N*	PM	*N*	PM	Allele effect	% of V/64
yld1.1	RM5	1	0.0036	5.77	251	6.38	45	6.99	1.22	18.26
yld2.1	RG256	2	0.0058	5.76	247	6.33	49	6.90	1.14	17.07

V/C and R/C refer to genotypic classes defined by single markers: V/C is the heterozygote (V20A/Ce64) and R/C the interspecific heterozygote (*O. rufipogon*/Ce64) in the BC_2 test-cross population. PM is the phenotypic mean of a class. The PM of R/C is 2 × PM of ($^1/_2$V/C + $^1/_2$R/C) − PM of V/C, assuming regular segregation within families. *P* is the probability that the marker genotype was not associated with yield. *N*, number of families within a genotypic class; Chr, chromosome number.

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and ten references.



a, Frequency distribution of yield in an interspecific (*O. sativa*/*O. rufipogon*) BC_2 test-cross population, with phenotypes of *O. rufipogon*, V20B and V/64 (V20A × Ce64) indicated by arrows; *b*, chromosome maps (marker order and map distances based on published rice molecular-genetic map[4]) showing locations of putative yield-enhancing genes from *O. rufipogon*, *yld1.1* and *yld2.1*. *O. rufipogon* allele associated with increased yield at $P < 0.005$ (black shading); $P < 0.010$ (dark grey shading); and $P < 0.050$ (light grey shading).

yielded less than V/64, but a small percentage of the plots yielded as much as 50% greater than the elite hybrid. Overall, 15% of the BC_2 test-cross families outperformed V/64 with respect to yield, 14% with respect to grains per plant, and 56% with respect to 1,000-grain weight. Thirteen (4.3%) of the BC_2 test-cross families outyielded V/64 by at least 30%. These results suggest that genes coming from *O. rufipogon* can increase yield of an elite rice variety, even though *O. rufipogon* itself is inferior to cultivated rice varieties.

If *O. rufipogon* genes are contributing to the higher yield of certain BC_2 test-cross families, it should be possible to detect the presence of the segments of wild chromosomes containing such genes (quantitative trait loci or QTL) and to demonstrate a significant correlation between their presence and higher yield.

A high-density molecular-genetic map developed for rice[4] has previously been used to map and characterize QTL conditioning heterosis[5], biotic and abiotic stress tolerance and various agronomic traits[6,7]. In an effort to identify regions of the *O. rufipogon* genome that might be contributing to a yield increase, 100 informative restriction-fragment length polymorphisms and 20 microsatellite markers (S. R. McC. *et al.*, unpublished data), covering the entire rice genome at intervals of roughly 12 cM, were assayed on each of the 300 BC_2 test-cross families.

We conducted QTL mapping on BC_2 test-cross data by regression of field performance on marker genotype using standard analysis of variance procedures and assuming regular segregation of wild and cultivated alleles within test-cross families. In most cases, introgression of *O. rufipogon* alleles had either no significant effect on yield or were inferior to the cultivated alleles. However, *O. rufipogon* alleles at marker loci RM5 on chromosome 1 and RG256 on chromosome 2 were associated with enhanced yield ($P < 0.006$; see table on previous page), and were designated *yld1.1* and *yld2.1* (see figure). The phenotypic advantage of the lines carrying *O. rufipogon* alleles at these loci was estimated to be 1.2 and 1.1 tonnes per hectare (figures converted to field scale using conventional plant densities), respectively, which corresponds to an 18 and 17% increase over V/64 (see table).

The alleles *yld1.1* and *yld2.1* were both associated with a significant increase in grains per plant ($P < 0.005$), but had no detectable effect on 1,000-grain weight, plant height or growth duration.

Rice was originally domesticated by humans from wild stands of native plants[8]. Only a small portion of the genetic variation found in nature was captured in this domestication process[9], and it is this limited genetic base that forms the foundation of all of our modern cultivars[10]. Many of the ancestors of cultivated crop species still exist in the wild or have been collected and maintained in germplasm banks. Although these wild species are valued as a unique source of genetic variation, they usually have low yields. Nevertheless, our results indicate that one of the closest wild relatives of cultivated rice, *O. rufipogon*, despite its overall inferior appearance, contains genes that can substantially increase the yield of rice. The strategy we describe here is being used to search for yield-enhancing genes in other wild rice species, and could be extended to the genetic improvement of other crop species.

Jinhua Xiao
Silvana Grandillo
Sang Nag Ahn
Susan R. McCouch
Steven D. Tanksley*
Department of Plant Breeding & Biometry, 252 Emerson Hall, Cornell University, Ithaca, New York 14853, USA
e-mail: sdt4@cornell.edu
Jiming Li
Longping Yuan
China National Hybrid Rice Research and Development Center, Changsha, Hunan 410125, People's Republic of China

1. Tanksley, S. D. *et al. Theor. Appl. Genet.* **92**, 213–224 (1996).
2. Tanksley, S. D. & Nelson, J. C. *Theor. Appl. Genet.* **92**, 191–203 (1996).
3. IRRI *1993–1995 IRRI Rice Almanac* (Manila, Philippines, 1993).
4. Causse, M. A. *et al. Genetics* **138**, 1251–1274 (1994).
5. Xiao, J., Li, J., Yuan, L. & Tanksley, S. D. *Genetics* **140**, 745–754 (1995).
6. McCouch, S. R. & Doerge, R. W. *Trends Genet.* **11**, 482–487 (1995).
7. Xu, Y. *et al. Sci. China* **38**, 422–428 (1995).
8. Oka, H. I. *Origin of Cultivated Rice* (Japan Sci. Soc., 1988).
9. Wang, Z. Y., Second, G. & Tanksley, S. D. *Theor. Appl. Genet.* **83**, 565–581 (1992).
10. Ladizinsky, G. *Econ. Bot.* **39**, 191–199 (1985).

*To whom correspondence should be addressed.

Origin of oscillons

SIR — The term 'oscillons', recently described in *Nature*[1,2] and featured on the cover, was earlier used by us[3] to represent strongly nonlinear electrostatic oscillations on a plasma boundary. These peculiar oscillations of surface fields are described by exact solutions of the cold electron fluid and Maxwell equations together with the boundary conditions. They have a unique spatial pattern and their frequency satisfies a simple algebraic dispersion relation involving the oscillon radius and peak amplitude.

Clearly, these oscillons are physically different from those recently discussed in *Nature*. But in all cases, strongly nonlinear anomalous oscillations of surface patterns are involved, and thus there is no fundamental conflict in the terminology.

L. Stenflo
Department of Plasma Physics, Umeå University, S-90187 Umeå, Sweden
M. Y. Yu
Institute for Theoretical Physics I, Ruhr University, D-44780 Bochum, Germany

1. Fineberg, J. *Nature* **382**, 763–764 (1996).
2. Umbanhowar, P. B., Melo, F. & Swinney, H. L. *Nature* **382**, 793–796 (1996).
3. Stenflo, L. & Yu, M. Y. *Phys. Fluids B* **1**, 1543–1544 (1989).

Corn Genome Pops Out of the Pack

Congress is poised to launch a corn genome project, but plant geneticists want to make sure other, related cereal grains aren't ignored

IRVINE, CALIFORNIA—In the next few weeks, key members of the U.S. Congress are planning to plant seed money into appropriations bills to launch a new genome project. This ambitious effort, focused on corn, or maize, the quintessential American food crop, could mean tens of millions of dollars for crop genetics research in the next few years.

The prospect of such an initiative has grabbed the attention of plant scientists, who hope it could do for crop genetics what the multibillion-dollar Human Genome Project is beginning to do for human genetics. By helping to unravel the multitude of genetic mysteries hidden in corn's crunchy kernels, the project could aid in understanding and combating common diseases of grain crops. It could also provide a big boost for efforts to engineer plants to improve grain yields and resist drought, pests, salt, and other environmental insults. Such advances are critical for a world population expected to double by 2050, says Robert Herdt, director of agricultural sciences at the Rockefeller Foundation. "Four species provide 60% of all human food: wheat, rice, maize, and potatoes," says Herdt. "And we don't have good strategies for increasing the productivity of plants."

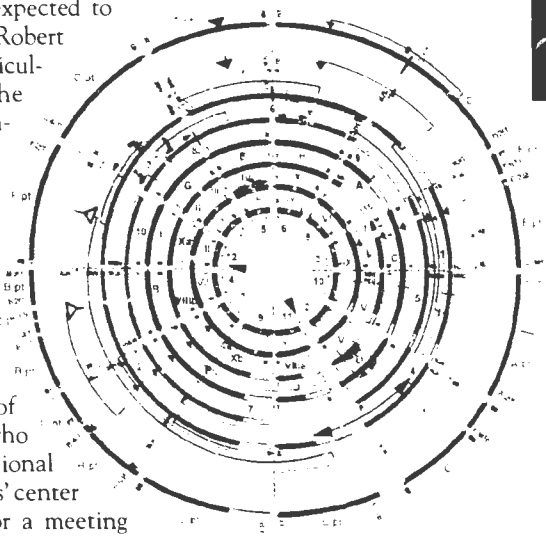
Herdt was one of 50 plant scientists who gathered at the National Academy of Sciences' center here 2 to 5 June for a meeting billed as "Protecting Our Food Supply: The Value of Plant Genome Initiatives." Although there's widespread enthusiasm for a plant genome project focused on a crop, the meeting revolved around the question of how much emphasis should be put on corn—which has a huge, complex genome. Several researchers believe the project should intensively study rice, too, which has a much simpler genome, and several other species from the genetically similar grasses family.

Unlike many meetings in which scientists dreamily discuss prospects in their field, the talks at this one were integral to the policy-making process. One of the meeting's co-organizers, Ronald L. Phillips of the University of Minnesota, St. Paul, is also chief scientist for the competitive grants program at the U.S. Department of Agriculture (USDA) and chair of a task force preparing a report to Congress about how the government should proceed on such an initiative. "This is the only meeting that I knew was going to be important from the outset," said Michael Freeling, a University of California, Berkeley, geneticist and the meeting's co-organizer, at the opening session.

Ear ye, ear ye

As much as scientists would like to see a crop genome project that studies several grasses, organizing it around corn makes a great deal of political sense. Corn is a potent economic force, providing much of the feed for the country's livestock, basic ingredients for everything from drugs to ethanol, and up to $8 billion in exports. And the industry is represented by a formidable political lobby, the National Corn Growers Association.

S. MCCOUCH



Old friends. Rice, wheat, and corn—key grains on a list of grass genomes to be sequenced—are believed to have diverged some 60 million years ago.

Oats

Maize
Sorghum
Sugarcane
Foxtail millet
Rice

SOURCE: M. GALE

Indeed, the idea for a publicly funded corn genome project began to take root late in 1995, when the growers' association put its muscle behind it. The sales pitch includes a 70-page business plan and a slick video promoting a "national corn genome initiative." They have won the backing of Senator Christopher "Kit" Bond (R–MO), who chairs the subcommittee that funds the National Science Foundation (NSF). "One of the reasons we're here [in Irvine] is because of the corn growers and the political momentum they created," said corn geneticist Joachim Messing of Rutgers University in Piscataway, New Jersey.

Bond says he is keeping an open mind about the project and that he welcomes input from scientists. "Give us a game plan," Bond told presidential science adviser Jack Gibbons and NSF director Neal Lane at a 22 April hearing on NSF's 1998 budget request. Bond asked Gibbons to assemble a panel to come up with such a plan, and the White House responded by creating the task force that Phillips chairs. An interim report is due this month, with a final report by December.

Several hundred academic researchers in the United States are studying corn genetics, estimated Ed Coe of the University of Missouri, Columbia, and three companies have projects under way to identify corn genes. But there is scant coordination between any of these efforts, and much of the data from the private companies are not widely available (see sidebar). The solution, according to the corn growers, is a federally funded, $143 million research program that would stitch together these varied efforts.

Given a chance to put science in the driver's seat, plant geneticists are trying to block out the key issues. The most fundamental question is the same one that faced researchers who launched the Human Genome Project a decade ago: What level of detail is needed? One group would like to fish out from corn and other model plants just the sequences of genomic DNA most likely to code for genes. That's typically a small part of any genome. Another camp says that sequencing the entire genome is the only way to find all the genes and understand their relation to each other. "This is 'déjà vu all over again,' " says David Cox, who co-directs a center working on the Human Genome Project at Stanford University. The lesson from the human experience, he says, is that "you need both."

However, the big problem with sequencing the entire corn genome is just that—it's big. Corn has about 3 billion pairs of bases (the building blocks of DNA), which makes it comparable in size to the human genome.

Please Pass the Data

IRVINE, CALIFORNIA—Deciphering the genetics of the mustard plant won't by itself meet the world's increasing demand for food. But plant biologist Christopher Somerville thinks that it can teach his colleagues a lot about sharing as they embark on a grain genome project (see main text).

Somerville is part of a coordinated, international effort to decode, or sequence, all of the DNA in the genome of *Arabidopsis* (*Science*, 4 October 1996, p. 30). But a slide he presented at a recent meeting on food crops (see below) makes clear that the extent of collaboration has been uneven. While several groups were sharing sequence information fully, he says, Japanese and European researchers had yet to put any sequences into public databases. "From the beginning, we've had a lot of international cooperation, and this is not in that spirit," says Somerville, who heads the Carnegie Institution of Washington's plant research branch in Stanford, California.

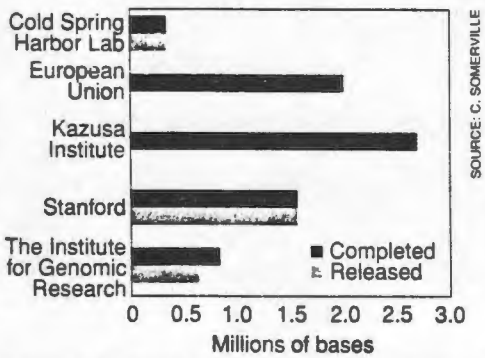
Leaders of both the Japanese and European *Arabidopsis* projects acknowledge the shortfall, but they say there's good reason. Michael Bevan of Britain's John Innes Centre, who heads the international consortium on the plant that Somerville and others contribute to, notes that European researchers, unlike their U.S. counterparts, don't release data until they have verified its accuracy. "The rapid release of highly accurate, annotated sequence is a goal we all aim to achieve," says Bevan. Satoshi Tabata, who heads the Japanese project at the Kazusa DNA Research Institute in Chiba, Japan, says money has been a big obstacle to the posting of data. Tabata says the project will begin releasing data next month and "will keep releasing data without delay after that."

Regardless of which view prevails, the *Arabidopsis* experience illustrates the obstacles to any effort to coordinate the sharing of genomic data. And even when scientists profess fidelity to the idea of sharing, the interests of industry and nationalism can be overwhelming. Michael Gale, a plant molecular biologist also at the Innes Centre, is much worried by what he says is recent pressure from the European Union (EU) to give industry first crack at any genome data. "The EU wants to protect its databases," says Gale. "It's something we should all fight very vigorously."

Rice genome researchers have long complained that Japan's 7-year-old rice genome project has been slow in sharing data (*Science*, 18 November 1994, p. 1187). While tensions have eased as the Japanese researchers have made their data and materials more widely available, similar concerns are being raised about the availability of data from a Chinese project. In particular, says Susan McCouch of Cornell University, Chinese researchers have had little interaction with foreign colleagues. "We have almost no information out of the Chinese project," said McCouch. "No one I know has ever seen any of that data."

Hong Guofan, director of the National Center for Gene Research at the Chinese Academy of Sciences in Shanghai, told *Science* that he expects the situation to improve shortly. An Internet site will offer data "as soon as the relevant financial arrangement has been settled." Hong noted that Chinese rice researchers also submitted an abstract of their work at a meeting in South Carolina last October and have a related paper in press. McCouch replies that Hong canceled his talk at the South Carolina meeting and again at a meeting held in San Diego this January. "As far as I am concerned, an abstract is not 'sharing data,' " says McCouch.

Other genomic data about grains are being held close to the vest because industry is directly funding the work. Three U.S. companies—Pioneer Hi-Bred, Monsanto, and DuPont—have corn genome projects under way. The highest profile belongs to Pioneer's project, a 3-year, $16 million deal with Human Genome Sciences of Rockville, Maryland, to pluck out pieces of corn genes and assemble them in a database.



Unequal portions. *Arabidopsis* collaborators follow different practices about sharing databases.

Setting the rules for access to such databases can be tricky. Pioneer's Steven Briggs says the data, while not in the public domain, are available. "We've granted everyone's request for access," said Briggs about the project, which began in January 1996. Briggs urged the U.S. government to negotiate with industry to gain access to private databases.

James Cook of the U.S. Department of Agriculture had some advice for the feds, too: Be tough with collaborators who withhold information. "If our partners are not holding up their end, pull the plug," said Cook. For their part, most scientists prefer the carrot to the stick. "A data-release war [would be] a disaster for the genome project," says Rob Martienssen, who co-leads the *Arabidopsis* sequencing project at Cold Spring Harbor Laboratory in New York. "We can only encourage them and lead by example." **–J.C.**

It also has an abundance of repeated sequences, which probably contain worthless information. Rice, on the other hand, is the smallest of the crop grasses (see table), with only 430 million base pairs. Moreover, rice has a great deal in common with corn, wheat, oats, barley, and other members of the grasses family, says rice researcher Susan McCouch of Cornell University. "Rice is the closest thing to the ancestral version of the grass genome," says McCouch. "Yet it still embodies the essential set of genes for grasses."

Researchers have already found broad similarities between the genes of different members of the grass family (see circle diagram). "It's no longer OK for me to be a wheat geneticist and for you to be a rice geneticist," says Michael Gale, a plant molecular biologist from the John Innes Centre in Norwich, U.K. "We all have to be cereal geneticists." But he says "it's still an open question" how many genes the different grasses actually share.

Obtaining the entire rice sequence would clarify how much "synteny"—similar genes that appear in similar locations of the genome—exists between the grasses. But how to get the most bang for the bucks that researchers hope Congress will devote to the project is another question. "If we get $100 million, do we suck it all into the rice genome?" asked Jeff Bennetzen, a corn geneticist at Purdue University. "For me, whole-genome sequencing is a low priority." Timothy Helentjaris of Pioneer Hi-Bred emphasized that rice has little political muscle. "If you got down to details and said 70% [of the budget] is for rice sequencing, you'd set off red flags," said Helentjaris.

Cornucopia project

By the meeting's end, participants had cobbled together a plan that seemed to satisfy the majority. The first was a proposal for an international rice genome sequencing effort, with the United States putting up half of the money and inviting China and Japan, both of whom are

funding rice genome projects, to join. The scientists also suggested building up a database of short sequences, called "expressed sequence tags" (ESTs), that can be used to identify expressed genes. They recommended sequencing 500,000 ESTs for corn and 100,000 each for rice, wheat, oats, barley, and sorghum. The group also called for computer databases to share data as they are generated and stock centers where researchers can freely receive the clones used to study the various plants.

The plan won plaudits from government officials eager to avoid a congressional mandate. "I'm very enthusiastic about what I've heard at this meeting," said Mary Clutter, head of the biology directorate at the NSF. "That is, focus on the science and let us build a program" to present to Congress. Clutter would like several agencies to participate in a project led by the USDA, with NSF funding a steering committee that would draw up a request for the 1999 fiscal year that begins on 1 October 1998.

Species	Genome size (millions of bases)	Predicted # genes
Mycoplasma genitalium	0.58	482
Haemophilus influenzae	1.83	1,727
Escherichia coli	4.72	4,307
Saccharomyces cerevisiae	12.50	6,000
Caenorhabditis elegans	100	13,100
Arabidopsis thaliana	150	20,000
Oryza sativa (rice)	430	30,000
Sorghum bicolor (sorghum)	760	30,000
Zea mays (corn)	2,000	30,000
Homo sapiens	3,000	100,000
Triticum aestivum (wheat)	16,000	30,000

SOURCE: SUSAN MCCOUCH/CORNELL

Gnats and giants. Grain genomes range in size, but are much larger than the nonhuman species being sequenced.

That's not soon enough for Kellye Eversole, a lobbyist for the corn growers at the meeting. "We don't want to spend another year on planning," says Eversole. "We want to see this get off the ground in 6 to 7 months." But James McLaren of Inverizon International, which drew up the business plan for the corn growers, signaled a willingness to be flexible about the scope of the project. "If you all tell me the best way to improve corn is to sequence rice, I'll support you," said McLaren. "But you'd better be right, because [the corn growers] are standing out front."

Congress also seems eager to get started. A staffer in Bond's office who asked not to be named told *Science* that legislators plan to designate $10 million for the effort in two separate parts of the spending bill for the agriculture department. Another earmark might appear in the appropriations bill that funds NSF. But Clutter takes issue with that approach. "Earmarking ... is anathema to the Administration," said Clutter. "It means taking away money from something planned."

Indeed, says Cliff Gabriel of the White House's Office of Science and Technology Policy, starting a genome project means curbing or ending an existing program. And although he didn't propose any candidates for the chopping block, he told the group that the Administration supports a grain initiative. "The time is right to do something," he said.

–Jon Cohen

GENOMICS

Alzheimer's Maverick Moves to Industry

Since 1994, the British drug company Glaxo Wellcome has been buying bits and pieces of U.S. biotech firms as part of a push into genetics. On 17 June, the company announced a surprising choice to direct its growing genetics empire: Allen Roses of Duke University, a prominent neuroscientist and controversial Alzheimer's disease researcher. Roses will run this $47 million directorate from Glaxo's U.S. headquarters in Research Triangle Park, North Carolina.

LES TODD/DUKE UNIVERSITY

Glaxo-bound. Roses will head genetics program.

Roses, an outspoken researcher whose ideas about the genetics of Alzheimer's have drawn a mixed reception from his peers, has been at Duke for 27 years and was named director of the university's Center for Human Genetics in 1996. He says the main reason he took the job with Glaxo is that "We are at a point now in the understanding of Alzheimer's disease [at Duke] that we are targeting" therapeutic products. "Universities don't make drugs and governments don't make drugs," Roses says, but "Glaxo Wellcome does." Glaxo Wellcome has funded Roses's work at Duke, and he says his research program will "be accelerated by my being inside" the company. Glaxo has agreed to allow Roses to continue some research at Duke as an adjunct professor.

As director of Glaxo's international genetics program, Roses will command a program based in labs in three countries (the United Srates, Britain, and Switzerland), comprising 150 researchers. According to Glaxo, the staff is expected to double over the next 18 months, as new departments are created to "ensure that genetics plays its part not only in drug discovery but also in development and in the commercialization of medicines." Roses's job will be to forge a coherent strategy, linking combinatorial chemistry at Affymax Research Institute of Palo Alto, California (purchased by Glaxo in 1995), gene expression research at Incyte Pharmaceuticals Inc. of Palo Alto (as of last month, a partner of Glaxo's), and clinical genetics studies at Spectra Biomedical of Menlo Park, California (purchased by Glaxo this month).

Roses says one of the reasons the company chose him is that he's not a fence straddler. Indeed, he notes, some of his peers have called him a "street fighter." For example, he recently spoke out at a Senate subcommittee hearing about what he called lack of vision in the public biomedical funding agencies. He says his grant requests to the National Institutes of Health received poor ratings from "narrowly focused scientists" with "dogmatic belief systems." His lab would have closed, he added, had it not received funding from Glaxo Wellcome.

Roses may be best known for showing that a protein involved in cholesterol transport (apolipoprotein E) is a factor in Alzheimer's disease. Roses and his colleagues also linked genes that encode variants of the protein (the apoE genes) to varying degrees of risk for Alzheimer's disease. Alison Goate, an Alzheimer's researcher at Washington University in St. Louis, says that while most researchers would agree that the gene known as apoE 4 is "the single most important risk factor" for Alzheimer's disease in the under-70 population, some of Roses's other conclusions are not widely accepted. Most controversial, Goate says, is a theory of Roses and his Duke colleague Warren Strittmatter that "good" versions of the apoE gene (E2 and E3) produce a protein that helps maintain healthy nerve cells, while the "bad" variant (E4) fails to do so, leading to Alzheimer's disease (*Science*, 19 November 1993, p. 1210). Because some Alzheimer's patients do not have the apoE 4 gene, and some people who have the gene do not have the disease, many researchers doubt that a test for apoE 4 would have value in predicting whether a healthy person will get the disease.

While Roses may seem an iconoclast to some, his colleague Peter St. George-Hyslop of the University of Toronto says he's really "not all that outrageous ... he likes to play that angle." Goate agrees: "He thrives on controversy." As for Roses's move to Glaxo, St. George-Hyslop comments: "It's good for them, bad for academic science."

–Eliot Marshall

Polymorphism revealed by simple sequence repeats

Wayne Powell, Gordon C. Machray and Jim Provan

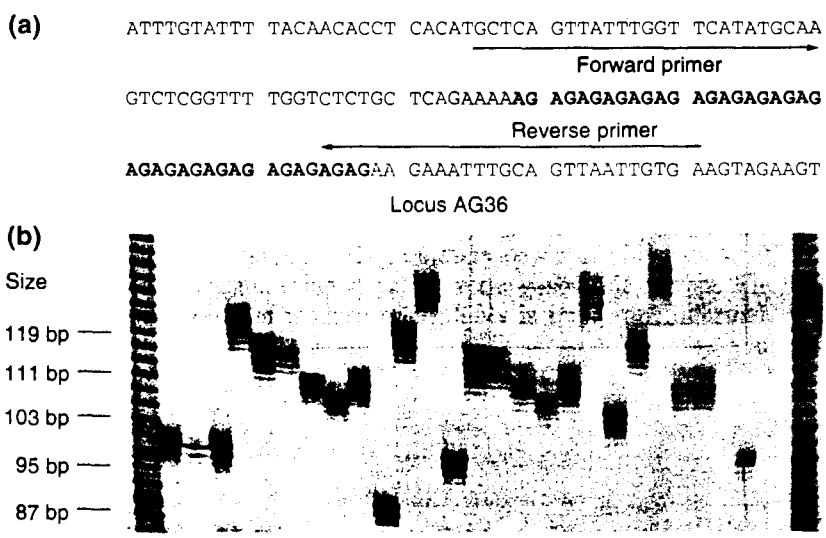
Simple sequence repeats (SSRs) are a group of repetitive DNA sequences that represent a significant portion of higher eukaryote genomes. They can serve as highly informative genetic markers, and in conjunction with the use of polymerase chain reaction (PCR) technology enable the detection of length variation. This novel means of detecting polymorphism targets highly variable regions of the genome, and has revolutionized human and mammalian research. It is now poised to have a significant impact in plant science.

Simple sequence repeats (also termed microsatellites) are stretches of DNA, consisting of tandemly repeating mono-, di-, tri-, tetra- or penta-nucleotide units, that are arranged throughout the genomes of most eukaryotic species. The existence of dinucleotide repeats – poly(C-A).poly(G-T) (i.e. an alternating sequence of cytosine and adenine, with, on the opposite strand of the DNA molecule, alternating guanine and thymine) – was first documented almost 15 years ago by Hamada and colleagues[1]. Subsequent studies by Tautz and Renz have confirmed both the abundance and ubiquity of microsatellites in eukaryotes[2]. A systematic study of the occurrence of human microsatellite sequences in the GenBank database by Weber and May revealed variation in the number of C-A repeats at seven of the eight loci examined[3]. In addition, these authors developed a general approach for detecting polymorphic (i.e. different length) microsatellites: polymorphism is revealed by PCR-amplification from total genomic DNA, using two unique primers, composed of short lengths of nucleotides, that flank and hence define the microsatellite locus. Amplification products obtained from different individuals can be resolved on gels to reveal this polymorphism.

The uniqueness and value of microsatellites arises from their multiallelic nature, codominant transmission, ease of detection by PCR, relative abundance, extensive genome coverage and requirement for only a small amount of starting DNA (Fig. 1). Also, these 'second generation' markers may be distributed between labs as primer sequences, thus providing a common language for collaborative research and acting as universal genetic mapping reagents[4]. Although the rationale behind the development of microsatellites was based on human and mammalian biology, significant opportunities exist for the use of microsatellite technology in plant science research. In this review, we outline our present understanding of microsatellite distribution in plants and examine the prospects for their deployment in the plant sciences.

Plant microsatellite isolation and characterization

The isolation and cloning of plant microsatellites was first performed with tropical tree species[5]. Poly(A-C) and



(c)

Advantages	Disadvantages
Relatively abundant with uniform genome coverage	Cloning and sequencing necessary
Hypervariable (i.e. have a high information content)	High resolution gels may be needed
Codominant, robust and reproducible	
PCR-based	
Each locus defined by primer sequence: facilitates exchange	
Represent sequence-tagged sites	

Fig. 1. Revealing polymorphism using microsatellites. (a) Soybean (*Glycine max*) library-derived microsatellite $(AG)_{20}$ at locus AG36. The repeat motif is shown in bold; forward and reverse primers for the polymerase chain reaction (PCR) are indicated. (b) Autoradiogram of a polyacrylamide gel separation of ^{33}P-labelled PCR amplification products. Genomic DNA from a range of soybean and *G. soja* samples was used as a template for amplification. The outer lanes contain size-marker DNA. (c) Advantages and disadvantages of using microsatellites to detect polymorphism.

 1360 - 1385/96/$15.00 PII S1360 - 1385(96)10025-X

TIGS ?

poly(A-G) dinucleotide repeats were detected at a frequency of 5×10^3 to 3×10^5 per genome, with A-G repeat motifs being more frequent than A-C. Although there are much less data on plant SSRs than for animal species, some important trends and differences are already apparent[6–8]. Overall, there is a tenfold reduction (Fig. 2) in the number of dinucleotide repeats detected in plant sequences compared with humans – but the sequence information available for humans is much greater. It has been estimated that on average a repeat longer than 20 bp in length occurs every 33 kb in plant nuclear genomes compared with every 6 kb in mammals[8]. In plant genomes, the A-T repeat motif predominates, whereas in humans an A-C or T-G repeat is most common (Fig. 2); these appear to be general features distinguishing all plant and animal genomes.

The relatively low frequency of microsatellites in plant genomes presents technical problems for the large-scale isolation of SSRs. Furthermore, A-T dinucleotides, which are the most abundant type of SSR in plants, are difficult to isolate from libraries because they are palindromic (i.e. the sequence is the same when one strand is read left to right or the other is read right to left). Standard methods for the isolation of SSRs involve:

- The creation of a small insert genomic library.
- Library screening by hybridization.
- DNA sequencing of positive clones.
- Primer design and locus-specific PCR analysis.
- Identification of polymorphisms.

Technical details of these procedures are provided by Rafalski *et al.*[9]. In order to improve efficiency, SSR-enriched libraries have been developed using various methods (Table 1), with selection either before or after library construction. Precloning enrichment methods involve the fragmentation of genomic DNA by sonication or endonuclease digestion, and the creation of DNA fragments with defined sequences at both ends. This can be achieved by ligation of adaptors, or through the use of degenerate (nonspecific) primers and PCR analysis. Microsatellite-containing fragments are enriched by hybridization to a short length of biotinylated SSR and subsequent selection on streptavidin-coated magnetic beads. Alternatively, direct DNA affinity chromatography may be deployed in the selection step. Further protocols for genomic DNA selection involve triple-helix formation using a capture oligonucleotide that can also be biotinylated. Triplex formation (some methods for which feature enhancement by RecA protein) has also been used for microsatellite enrichment after cloning of genomic DNA into vectors. A further postcloning selection method generates single-stranded template from cloned DNA for second-strand synthesis primed by an oligonucleotide repeat, with subsequent cloning of the reconstituted, double-stranded product. These methods have met with some success in detecting microsatellites, but both approaches to enrichment require careful evaluation to ensure adequate representation, independence of clones and the absence of PCR-induced amplification artefacts.

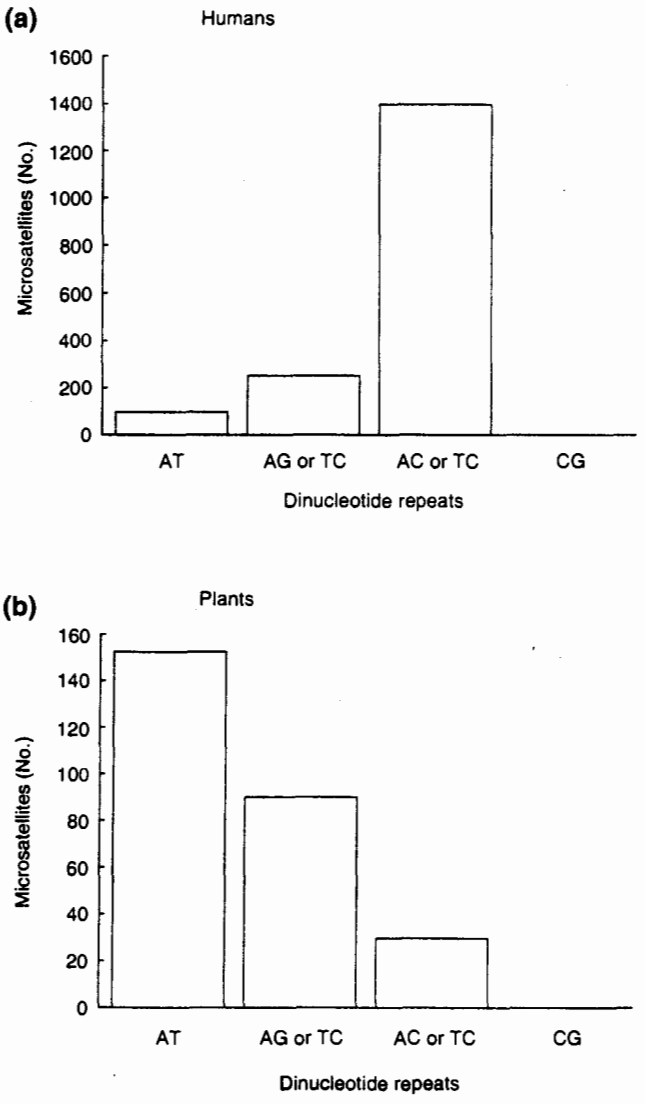


Fig. 2. Dinucleotide microsatellite numbers and composition in humans (a) and plants (b). *Data derived from the primate and plant subsections of the GenBank database.*

Applications

Agricultural genetics and plant breeding

The utility of microsatellites arises from two main factors: their high information content (which is a feature of the number and frequency of alleles detected); and ease of genotyping. The ability to distinguish between closely related individuals is particularly important for many crop species, which tend to have a narrow genetic base. To date, an evaluation of the amount of diversity detected with microsatellites has revealed more polymorphism compared with other assay procedures. Table 2 summarizes results from several recent investigations designed to examine levels of allelic variability detected with microsatellites in six major crop plants. With the exception of the tomato genus (*Lycopersicon*)[19], where only two out of ten microsatellites detected polymorphism in *L. esculentum* and *L. pennelli*, microsatellites proved to be highly discriminatory. This is best illustrated from data obtained with the soybean (*Glycine*) genus, in which levels of polymorphism detected with either restriction fragment length polymorphisms (RFLPs) or microsatellites have been compared[30]. More than 80% of the microsatellites used were able to detect three or more alleles, while fewer than 15% of the RFLP probes used proved to be as effective. It is important to bear in mind that many of the plant-based studies with microsatellites have been conducted using germplasm selected to maximize the levels of diversity detected. Two

Table 1. Microsatellite enrichment methods

Method	Description	Application	Refs
DNA affinity chromatography	Fragments generated by sonication or endonuclease digestion of genomic DNA are denatured, captured by hybridization to poly(U-G)-Sepharose or nylon-bound poly(G-T), eluted and cloned. Options exist to amplify by the polymerase chain reaction (PCR) using ligated adaptors.	Genomic DNA	10
Duplex-hybridization-based formation	Fragments generated by sonication or endonuclease digestion are denatured and hybridized to a biotinylated oligonucleotide with a repeat motif complementary to the target sequence. The duplex is captured on streptavidin beads. Fragments may be cloned or products may be selectively amplified by PCR, using primers against sequences added to the starting material by ligation or degenerate PCR.	Genomic or cloned DNA	11, 12, 13
Triplex formation	Triple helix formation using duplex DNA (genomic fragments or cloned DNA inserts) and selective oligonucleotides. Potential for RecA protein enhancement, and direct or biotin-mediated capture.	Genomic or cloned DNA	14, 15 16, 17
Oligo-primed second strand synthesis	Single-stranded template generated from cloned DNA in *dut ung* strain of *Escherichia coli*; second strand synthesis primed using selective repeat oligonucleotide; and transformation into standard *E. coli* for recloning of selected duplexes.	Cloned DNA	18

exceptions to this are from recent work on soybean (*G. max*)[30] and maize (*Zea mays*)[29], in which the germplasm being analyzed was of direct relevance to breeders and represented the genetic base of the species being studied. Further studies of this type are required for a comprehensive and realistic assessment of the usefulness of microsatellites.

Perhaps not surprisingly, the first application of microsatellites in plants has been in cultivar identification, where microsatellites have been used to genotype, unequivocally, such diverse material as grapevine (*Vitis vinifera*)[31] and soybean[21]. It is anticipated that the need to protect proprietary germplasm will increase in the future, and microsatellites will have an important role in securing plant variety rights.

A further application of microsatellites is in the determination of 'hybridity'. For this purpose, the codominant nature of microsatellites is particularly important, and allows the allelic contribution of each parent to be detected in sexual or somatic hybrids. One of the limitations to the wider application of somatic cell hybridization techniques in plant biology is a method for clearly identifying nuclear somatic hybrids. Recently, microsatellites have been used to characterize and determine the hybridity of intraspecific somatic hybrids of potato (*Solanum tuberosum*)[32]. In addition to highlighting the use of microsatellites in identifying heterokaryons, this study also demonstrates the somatic stability of potato microsatellites in plants that have been regenerated from single-cell protoplasts.

The high discriminatory power of SSRs is also important for analyzing variation in the gene pools of crops. The genetic base of modern crop plants is important information that

Table 2. Levels of variability detected in six major agricultural crops with microsatellites

Crop	Alleles (No.)	Diversity indices*	Refs
Tomato (*Lycopersicon esculentum*)	1–5	Unavailable	19
Soybean (*Glycine max and G. soja*)	11–26	0.71–0.95	20, 21, 22
Rice (*Oryza sativa*)	5–11	0.64–0.90	23, 24
Barley (*Hordeum vulgare*)	3–37	0.46–0.93	25, 26
Wheat (*Triticum aestivum*)	2–7	0.29–0.79	27
Maize (*Zea mays*)	2–11	0.40–0.89	28, 29

*Diversity index is $1-\Sigma p_i^2$, where p_i is the allele frequency of the ith allele.

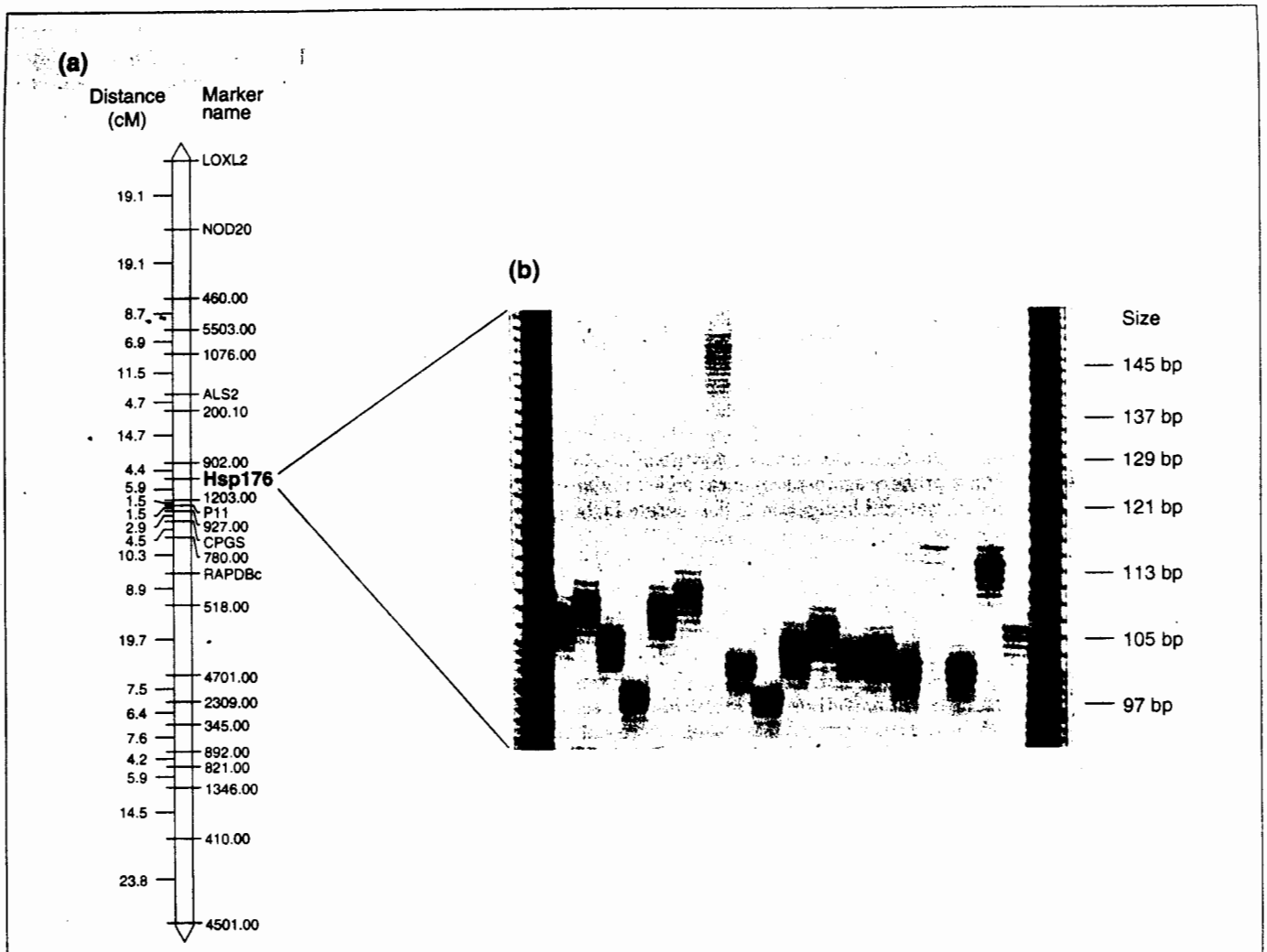


Fig. 3. (a) Microsatellite locus (HSP176) linked [0.5 centimorgans (cM)] to a gene conferring resistance to soybean (*Glycine max*) mosaic virus (*Rsv*) on linkage group 1 (Ref. 20). (b) Primers flanking the poly(A-T) repeat motif were used to amplify soybean genomic DNA using the polymerase chain reaction (PCR). The autoradiogram of polyacrylamide-gel-separated, ^{33}P-labelled (PCR) amplification products is shown. The outer lanes contain size-marker DNA. A total of 13 alleles was detected in a sample of soybean and *G. soja* accessions.

needs to be quantified in relation to progenitor landraces. A recent study with rice (*Oryza sativa*) established that 28% of the allelic variability was lost during the process of cultivar development from landraces[24]. Further studies with SSRs would improve our understanding of the domestication process involved in the evolution of crop plants. The data generated by this analysis would also provide useful criteria for enriching the genepool of crop plants, and determine how efficient plant breeders have been in accessing pre-existing forms of variation.

The creation of high resolution linkage maps based exclusively on SSRs may be an unrealistic goal for most crop plants. However, the creation of microsatellite index maps is an achievable objective, where microsatellites can provide reference ('anchor') points for specific regions of the genome. Thirty microsatellites have already been assigned to the five *Arabidopsis* linkage groups[33], providing sequence-tagged sites (STSs) for relating physical and recombinational genetic maps. Microsatellite markers have also been integrated into the soybean genetic linkage map[34]. Based on the analysis of 40 microsatellite loci, good overall genome coverage was obtained, although evidence of some clustering of microsatellites was also reported.

One of the main reasons for developing microsatellites is their potential use as diagnostic markers for important traits in plant breeding programmes. The first example of a microsatellite linked to a plant disease resistance gene has recently been published[35]. An $(A\text{-}T)_{15}$ repeat located within a soybean heat-shock-protein gene is 0.5 centimorgans (cM) from *Rsv*, which is a gene conferring resistance to soybean mosaic virus (Fig. 3). Several resistance genes, including those to peanut mottle virus (*Rpv*), *Phytophthora* (*Rps3*) and Javanese root-knot nematode, are clustered in this region of the soybean genome. Linkage to a diagnostic microsatellite marker would help allow the transfer of these resistance genes into elite soybean cultivars.

For some crop species, particularly those that inbreed, genetic improvements have been achieved using the pedigree selection scheme. Detailed records of progeny–parent relationships exist from using this procedure. A retrospective analysis of this germplasm using genetic markers would allow the transfer of alleles between generations to

be assessed. Furthermore, the consequence of selection for these alleles on plant phenotype can be inferred by relating the chromosomal location of marker loci to agronomic and biochemical data available for each genotype involved in the analysis. The choice of DNA marker system to be deployed in such studies is crucial. Microsatellites have many advantages for tracing pedigrees, because they represent single loci and avoid the problems associated with multiple banding patterns obtained with other marker systems. Furthermore, the multiallelism of microsatellites means that comparable allelic variability can be detected reliably across a wide range of germplasm of relevance to plant breeders. Pedigree-based studies will also provide much-needed information on the germline stability of microsatellites during breeding generations. Graphically based software is available[36], which will allow molecular-marker-based data to be integrated with ancestral relationships and performance information.

Estimating genetic relatedness and microsatellite locus conservation

The unique mechanism responsible for generating high levels of microsatellite-allele variability – by 'replication slippage' – may reduce the usefulness of microsatellites for assessing genetic similarity between genotypes. Microsatellites have been used to generate information on subpopulation relationships in humans[37], but studies of this type are less advanced in plants. Despite the multiallelic nature of soybean microsatellites, they have been used successfully to describe phenetic relationships (i.e. reflecting phenotypic characters) between accessions that correspond to their geographical origins[38]. Only a small number of microsatellites (approximately 30) may be required to obtain high-resolution estimates of relatedness. However, it is important to distinguish between the use of microsatellites to estimate phenetic as opposed to phylogenetic relationships. In the latter case, the two-way nature of the mutation process may be particularly troublesome. However, microsatellites do contain some phylogenetic information[39], and alternative statistical measures of genetic distance have been derived that may account for the mutational processes operating at microsatellite loci[39].

The ability to use the same SSR primers in different plant species will depend on the extent to which primer sites flanking SSRs are conserved between related taxa, and the stability of the SSR over evolutionary time. Some studies have considered cross-species amplification, with primer-sequence conservation being demonstrated in rice[23], grapevine[31] and *Citrus* spp.[40]. Cross-species amplification of microsatellite alleles with the same primers would increase the value of such markers in plant science.

Natural plant communities

An important issue in conservation biology is the need to understand what determines patterns of genetic variation in natural plant populations. Most species that occupy environmentally sensitive habitats in both temperate and tropical regions are outbreeders. For these species, information on the extent of gene flow, genetic differentiation and levels of inbreeding (particularly in fragmented populations) has important practical relevance for the conservation of plant genetic resources. Microsatellites have great potential in this context: their multiallelic, codominant nature allows individuals to be uniquely genotyped, so that gene flow and paternity can be established.

The resolving power of microsatellites has recently been demonstrated in a tropical tree species (*Pithecellobium elegans*)[41]. Four polymorphic microsatellites were evaluated in two populations of *P. elegans* collected in Costa Rica. A comparison of microsatellites with isozyme data conducted on the same material revealed that the number of alleles and expected heterozygosity (diversity) of microsatellites was more than double that obtained with six polymorphic-isozyme loci. Microsatellites are being used in this species to undertake a hierarchical analysis of paternity at the levels of the tree and individual fruit pod. This approach may provide much-needed information on the pollination biology of tree species. A further example of such an approach is the use of microsatellites to measure pollen-mediated gene transfer in an insect-pollinated, leguminous tropical tree – *Gliricidia sepium*[42]. This study was based on a restricted, endangered population located in Guatemala, and revealed that most pollen movement occurred within a radius of 75m of the mother tree. However, 3% of observed events occurred over a greater distance, with one example of gene flow over a distance of 289m being recorded. These direct estimates of gene flow are sufficiently high to influence the evolution of small populations of insect-pollinated tropical tree species. One area of concern is the occurrence of null alleles (where there is no PCR-amplification product), which may be a problem with the use of microsatellites for highly outbred, heterozygous species.

Chloroplast microsatellites

The analysis of animal mitochondrial DNA has revolutionized studies of genomic variation and gene flow in natural populations. The main reasons for this include the mainly uniparental inheritance and absence of recombination in mitochondria, providing a series of nonrecombining characters that allow a clear definition of maternal lineages. A comparison of the patterns of variability detected with biparentally (nuclear) and uniparentally (organellar) transmitted markers can also provide complementary information for population and evolutionary biologists. In contrast to animal mitochondrial DNA, chloroplast DNA variation has been used relatively little by plant population biologists[43]. The occurrence of polymorphic mononucleotide

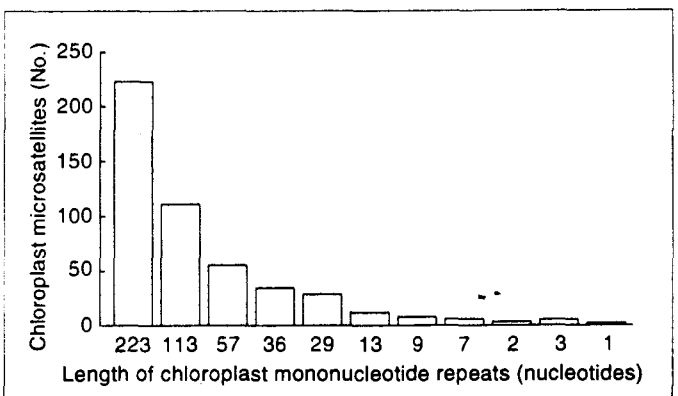


Fig. 4. Length distribution of mononucleotide repeat microsatellites in chloroplasts. The numbers of microsatellites observed in fully sequenced chloroplast genomes and other published chloroplast sequences are classified based on their length (10–20 nucleotides).

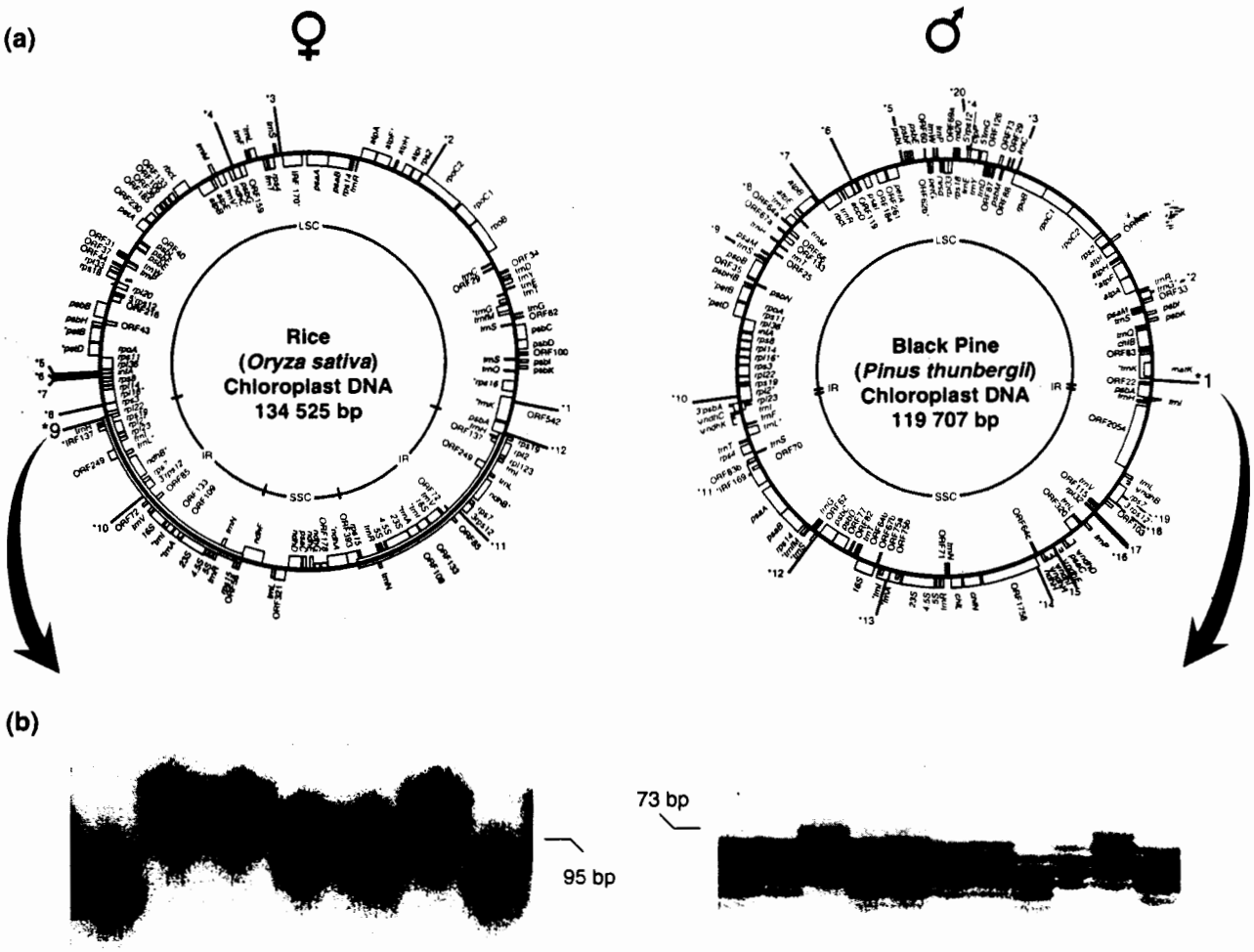


Fig. 5. Seed- and pollen-specific markers. Microsatellite locations in maternally inherited (rice) and paternally inherited (pine) chloroplasts are shown in (a) (*reproduced, with permission, from Ref. 45*). (b) Examples of intraspecific polymorphism detected with chloroplast simple sequence repeats: primers flanking mononucleotide repeats in the rice (*9) and pine (*1) chloroplast genomes were used to amplify (polymerase chain reaction) DNA from these two species and the autoradiogram of the polyacrylamide-gel-separated, ^{33}P-labelled amplification products is shown. (c) Features of maternally and paternally inherited microsatellites.

repeats in the chloroplast genomes of higher plants has recently been reported[44], providing alternative approaches to detect intraspecific chloroplast polymorphism.

The results of computer database searches to detect chloroplast SSRs (cpSSRs) have been presented[44] and updated by the inclusion of data derived from the recently sequenced maize chloroplast genome. Single nucleotide repeats predominate, with A.T mononucleotide repeats (i.e. repeating adenine on one strand of the DNA and repeating thymine on the other strand) almost exclusively being detected. The number of A.T repeats observed is higher than expected and decreases exponentially when the length of the repeats increases (Fig. 4). Dinucleotide poly(A-T) and poly(T-A) repeats were only found in the chloroplast of liverworts (*Marchantia polymorpha*), maize, the nonphotosynthetic green plant *Epifagus virginiana* and pea (*Pisum sativum*). A total of 500 cpSSRs was identified with repeat motifs greater than ten repeat units. This development provides new opportunities for high resolution analysis of chloroplast variability at the intraspecific level. The chloroplast genomes of five species [rice, tobacco, black pine (*Pinus thunbergii*), liverwort and maize] have now been fully sequenced. The distribution of mononucleotide chloroplast repeats for rice and pine, together with examples of polymorphism detected in cpSSRs, is shown in Fig. 5. These two species exhibit different modes of chloroplast

transmission and hence cpSSRs provide seed- and pollen-specific markers. The detection of length variation at multiple, physically linked sites may be used to provide haplotype data and so uniquely genotype individual plants. Intraspecific polymorphisms caused by variation in the length of chloroplast microsatellites will provide new tools with which to examine the relative contribution of seed and pollen movement to gene flow and hence the genetic structure of plant populations. Furthermore, both nuclear and chloroplast microsatellites may be analyzed simultaneously to provide a combined assay for two independently transmitted genomes[45]. This will provide further opportunities to examine chloroplast–nuclear interactions. Polyploidy is a common feature in the evolution of flowering plants. The use of uniparentally (chloroplast) and biparentally (nuclear) inherited microsatellites may help unravel the origin of polyploid complexes that have been recalcitrant to traditional forms of analysis.

Conclusions and future prospects

Microsatellite markers are highly informative and provide an efficient, accurate means of detecting genetic variation in any organism from which DNA can be extracted. Automated, fluorescence-based detection methods[46] and capillary electrophoresis for separating and sizing microsatellite alleles will help in the practical application of this technique. Further research is also required to characterize trinucleotide and higher order microsatellite repeats and understand the key factors responsible for generating polymorphism in complex plant genomes. For example, more information is required to assess the evolution of microsatellite length variation and to determine the efficiency of different repeat motifs in detecting polymorphism. Mutational processes operating at microsatellite loci will also need to be better defined to allow rigorous statistical analysis of data from experiments designed to investigate population differentiation. Access to primer sequences through a public database would maximize the use of microsatellites and benefit a variety of research disciplines in plant biology.

Acknowledgements

The authors are funded by the Agriculture, Environment and Fisheries Dept of the Scottish Office.

References

1 Hamada, S.R. *et al.* (1982) A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes, *Proc. Natl Acad. Sci. USA* 79, 6465–6469
2 Tautz, D. and Renz, M. (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes, *Nucleic Acids Res.* 12, 4127–4138
3 Weber, J. and May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction, *Am. J. Genet.* 44, 388–396
4 Silver, L.M. (1992) Bouncing off microsatellites, *Nat. Genet.* 2, 8–9
5 Conduit, R. and Hubbell, S.P. (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes, *Genome* 34, 66–71
6 Morgante, M. and Olivieri, A.M. (1993) PCR amplified microsatellites as markers in plant genetics, *Plant J.* 3, 175–182
7 Lagercrantz, U., Ellegren, H. and Andersson, L. (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates, *Nucleic Acids Res.* 21, 1111–1115
8 Wang, Z. *et al.* (1994) Survey of plant short tandem DNA repeats, *Theor. Appl. Genet.* 88, 1–6
9 Rafalski, J.A. *et al.* (1996) Generating and using DNA markers in plants, in *Analysis of Non-mammalian Genomes – A Practical Guide* (Birren, E. and Lai, E., eds), Academic Press
10 Brenig, B. and Brem, G. (1991) DNA cloning of sequence tagged microsatellite sites by DNA affinity chromatography, *Nucleic Acids Res.* 19, 5441
11 Karagyozov, L., Kolcheva, I.D. and Chapman, V.M. (1993) Construction of random small-insert genomic libraries highly enriched for simple sequence repeats, *Nucleic Acids Res.* 21, 3911–3912
12 Lyall, J.E.W. *et al.* (1993) A method for creating chromosome-specific plasmid libraries enriched in clones containing $(CA)_n$, *Nucleic Acids Res.* 21, 4641–4642
13 Kijas, J.M.H. *et al.* (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles, *Biotechniques* 16, 656–662
14 Ito, T., Smith, C.L. and Cantor, C.R. (1992) Sequence-specific DNA purification by triplex affinity capture, *Proc. Natl Acad. Sci. USA* 89, 495–498
15 Rigas, B. *et al.* (1986) Rapid plasmid library screening using RecA-coated biotinylated probes, *Proc. Natl Acad. Sci. USA* 83, 9591–9595
16 Kiyama, R., Nishikawa, N. and Oishi, M. (1994) Enrichment of human DNAs that flank poly(dA)·poly(dT) tracts by triplex DNA formation, *J. Mol. Biol.* 237, 193–200
17 Nishikawa, N., Oishi, M. and Kiyama, R. (1995) Construction of a human genomic library of clones containing poly(dGdA)·poly(dTdC) tracts by Mg^{2+}-dependent triplex affinity capture – DNA polymorphism associated with the tracts, *J. Biol. Chem.* 270, 9258–9264
18 Ostrander, E.A. *et al.* (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences, *Proc. Natl Acad. Sci. USA* 89, 3419–3423
19 Broun, P. and Tanksley, S.D. (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome, *Mol. Gen. Genet.* 250, 39–49
20 Morgante, M. *et al.* (1995) Genetic mapping and variability of seven soybean simple sequence repeat loci, *Genome* 37, 763–769
21 Rongwen, J. *et al.* (1995) The use of microsatellite DNA markers for soybean genotype identification, *Theor. Appl. Genet.* 90, 43–48
22 Maughan, P.J., Saghai-Maroof, M.A. and Buss, G.R. (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean, *Genome* 38, 715–723
23 Wu, K-S. and Tanksley, S.D. (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice, *Mol. Gen. Genet.* 241, 225–235
24 Yang, G.P. *et al.* (1994) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice, *Mol. Gen. Genet.* 245, 187–194
25 Saghai-Maroof, M.A. *et al.* (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics, *Proc. Natl Acad. Sci. USA* 91, 5466–5470
26 Becker, J. and Heun, M. (1995) Barley microsatellites: allele variation and mapping, *Plant Mol. Biol.* 2, 835–845
27 Plaschke, J., Ganal, M.W. and Röder, M.S. (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers, *Theor. Appl. Genet.* 91, 1001–1007
28 Taramino, G. and Tingey, S. Simple sequence repeats for germplasm analysis and mapping in maize, *Genome* (in press)
29 Senior, M.L. and Heun, M. (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer, *Genome* 36, 884–889
30 Powell, W. *et al.* The utility of RFLP, RAPD, AFLP and SSRP (microsatellite) markers for germplasm analysis, *Mol. Breeding* (in press)
31 Thomas, M.R. and Scott, N.S. (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs), *Theor. Appl. Genet.* 86, 985–990
32 Provan, J. *et al.* Analysis of intra-specific somatic hybrids of potato (*Solanum tuberosum*) using simple sequence repeats, *Plant Cell Rep.* (in press)

33 Bell, C.J. and Ecker, J.R. (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*, *Genomics* 19, 137–144
34 Akkaya, M.S. *et al.* (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map, *Crop Sci.* 35, 1439–1445
35 Yu, Y.G. *et al.* (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance, *Phytopathology* 84, 60–64
36 Boutin, S.R. *et al.* (1995) Marker based pedigrees and graphical genotypes generated by supergene software, *Crop Sci.* 35, 1703–1707
37 Bowcock, A.M. *et al.* (1994) High resolution of human evolutionary trees with polymorphic microsatellites, *Nature* 368, 455–457
38 Powell, W. *et al.* Genepool variation in genus *Glycine* subgenus *soja* revealed by polymorphic nuclear and chloroplast microsatellites, *Genetics* (in press)
39 Goldstein, D.B. *et al.* (1995) An evaluation of genetic distances for use with microsatellite loci, *Genetics* 139, 463–471
40 Kijas, J.M.H., Fowler, J.C.S. and Thomas, M.R. (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species, *Genome* 38, 349–355
41 Chase, M., Kesseli, R. and Bawa, K. (1996) Microsatellite markers for population and conservation genetics of tropical trees, *Am. J. Bot.* 83, 51–57
42 Dawson, I.K. *et al.* Simple sequence repeats provide an exact indicator of pollen-mediated gene flow in the leguminous tropical tree species *Gliricidia sepium*, *Mol. Ecol.* (in press)
43 McCauley, D.E. (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants, *Trends Ecol. Evol.* 10, 198–202
44 Powell, W. *et al.* (1995) Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines, *Proc. Natl Acad. Sci. USA* 92, 7759–7763
45 Powell, W. *et al.* (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome, *Curr. Biol.* 5, 1023–1029
46 Kresovich, S. *et al.* (1995) Abundance and characterisation of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed), *Theor. Appl. Genet.* 91, 206–211

Wayne Powell, Gordon Machray and Jim Provan are at the Dept of Cell and Molecular Genetics, Scottish Crop Research Institute, Invergowrie, Dundee, UK DD2 5DA.

Dynamic microtubules: implications for cell wall patterns

Carol Wymer and Clive Lloyd

The shape of plant cells is thought to depend largely on the relationship between cortical microtubules (MTs) and cellulose microfibrils in the cell wall. As templates for the deposition of newly polymerized microfibrils, MTs indirectly influence the direction of cell expansion, and this parallelism between MTs and microfibrils is therefore a central aspect of plant morphogenesis. However, microfibrils can change their alignment with each new layer that is deposited in the cell wall, and it was unclear until recently how MTs could undergo corresponding reorientations. Now, microinjection studies have begun to show how MTs rapidly interconvert between complex three-dimensional configurations, allowing models for the construction and behaviour of the MT array to be revised.

Cells in files tend to elongate considerably more than they widen. This is because cellulose microfibrils are deposited in an oriented manner, usually around the transverse axis of elongating cells. Since cellulose has a high tensile strength, it resists radial cell expansion (increase in girth) while adjacent transverse microfibrils can still separate to allow longitudinal expansion (increase in length). This 'hoop reinforced' arrangement of cellulose was first observed with polarization optics, and was confirmed by electron microscopy (EM) (see Ref. 1). Electron microscopic studies then revealed that cortical microtubules run parallel to those microfibrils closest to the plasma membrane[2] and so the idea persisted that MTs were also hooped.

Most cell walls are very complex, however, consisting of multiple layers in which the orientation of the cellulose microfibrils can change from layer to layer and in response to hormones. This places severe demands on the organization and behaviour of any supposed cytoplasmic template. The 'multi-net hypothesis' stated that cellulose microfibrils were always polymerized transversely around the cell surface, forming a flat helix – but that as the cell continued to expand, each new layer was displaced outwards such that the initially transverse microfibril network became passively stretched, giving steeper helical angles[3]. However, subsequent EM research showed that cellulose is not always deposited with a transverse alignment, and that MTs are not always transverse. An alternative idea (the 'ordered subunit model'[4]) was that the layer-to-layer shift in the microfibril angle was not caused by reorientation after polymerization, but was a record of what happened to each new layer as it was deposited at the cell surface: regular changes in helical angle, from layer to layer, produced a helicoidal wall.

Not all walls are helicoidal, but for some investigators it was the undoubted occurrence of microfibrils in nontransverse and varying angles that strained the credibility of the idea of microtubules as templates for wall patterning. Box 1 illustrates some observed cellulose patterns, giving an indication of the flexibility required of a template. In some models, a cytoplasmic template was dispensed with altogether – the alternation of microfibril alignment from layer to layer was

 1360 - 1385/96/$15.00 PII S1360 - 1385(96)10024-8

Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines

J. ANTONI RAFALSKI AND SCOTT V. TINGEY

The science of mapping genetic traits, including those of agronomic interest, is well established and many genetic marker systems are available. However, the application of genetic diagnostics in plant breeding is in its infancy. The sample throughput and cost requirements are very different from those of medical DNA diagnostics. It will be necessary to automate the DNA isolation process, DNA amplification-based allele identification and data handling. Here, we discuss recent progress in the development of molecular technology for plant breeding.

Many of the problems associated with breeding programs based on the phenotypic estimation of an agronomic trait, such as major environmental effects or quantitative inheritance, can be eliminated by the use of DNA-based diagnostics. These approaches are expected to be particularly advantageous in breeding programs that involve multigenic or quantitative traits, which are notoriously difficult to deal with when relying on a phenotypic assay alone[1].

The recognition that DNA markers can be extremely useful in plant breeding[2,3] came with the development of RFLP-based genetic mapping in plants (see footnote* for a list of the abbreviations used in this review). Markers that flank a gene determining a trait of agronomic interest can be used to track the trait in genetic crosses. In another application, are molecular markers that are uniformly distributed throughout the genome can be used to estimate the genetic contribution of each parent to each member of a segregating population. Individuals whose genome composition most resembles the recurrent (i.e. backcross) parent genome may be selected for the next cross; this will accelerate the introgression of traits from genetically distant germplasm sources[4].

The term 'molecular breeding' is used to describe plant breeding programs that are supported by the use of DNA-based diagnostics. The exact way in which genetic diagnostics is applied in plant breeding will be influenced by many considerations, both biological and economic, but the availability of cost-effective diagnostic technology that is compatible with plant breeding applications is a major factor.

Genetic marker systems

The realization that DNA sequence polymorphisms between individuals can be used for genetic mapping[5] changed the practice of genetics. Recently, several new marker systems that promise to meet the requirements of an automated genetic diagnostic assay have become available (Table 1). Many of these assays are based on DNA amplification.

Restriction fragment length polymorphisms

RFLP markers remain extremely useful in research applications. Such markers are codominant (heterozygotes can be distinguished from either homozygote) and provide complete genetic information at a single locus. The amount of DNA required for RFLP analyses is relatively large (5–10 μg), but a single Southern blot may be re-probed many times over a period of years, making this technology more efficient. Multiple Southern blots, corresponding to hundreds of individuals, can be probed simultaneously and many highly saturated RFLP maps have been produced in this way.

*Abbreviations: AP-PCR, arbitrarily-primed PCR; AS-PCR, allele-specific PCR; CAPS, cleaved amplified polymorphic sequences; DAF, DNA amplification fingerprinting; DGGE, denaturing gradient gel electrophoresis; PCR, polymerase chain reaction; QTL, quantitative trait loci; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; SCAR, sequence-characterized amplified regions; SSCP, single-strand conformational polymorphism; SSR, simple sequence repeats.

New genetic markers or genes can easily be located within the context of an existing RFLP map, however, with few exceptions, there is little information available on allele sizes and the distribution of markers in the germplasm. Databases containing such information would be very useful to the plant breeding community. The logistical problems associated with documenting and distributing large numbers of clones have slowed the dissemination of RFLP probes, although a mechanism for dispersing probes has been put in place by the *Arabidopsis* and maize research communities.

An alternative to one of the disadvantages of RFLP markers, the need for radioactive probes, has come with the availability of sensitive nonradioactive detection systems[6]. Automation of RFLP mapping is difficult, and it may be more practical to turn to one of the DNA-amplification based marker systems to provide an automated genotype assay.

Random amplification of polymorphic DNA

Technology for the amplification of discrete loci with single, random-sequence, oligonucleotide primers (RAPD[7], AP-PCR[8], DAF[9]; see Box 1) became popular because of its simplicity and ease of use in a modestly equipped laboratory.

The RAPD amplification reaction is performed on a genomic DNA template and primed by an arbitrary oligonucleotide primer, resulting in the amplification of several discrete DNA products. These are usually separated on agarose gels and visualized by ethidium bromide staining. Each amplification product is derived from a region of the genome that contains two short DNA segments with some homology to the primer; these segments must be present on opposite DNA strands, and be sufficiently close to each other to allow DNA amplification to occur. Using short primers and low annealing temperatures ensures that several sites, randomly distributed in the genome, give rise to amplification products. The polymorphisms between individuals result from sequence differences in one or

Table 1. Properties of systems for generating genetic markers

	RFLP	RAPD	Microsatellites	CAPS
Principle	Endonuclease restriction Southern blotting Hybridization	DNA amplification with random primers	PCR of simple sequence repeats	Endonuclease restriction of PCR products
Type of polymorphism	Single base changes Insertions Deletions	Single base changes Insertions Deletions	Changes in length of repeats	Single base changes Insertions Deletions
Genomic abundance	High	Very high	Medium	High
Level of polymorphism	Medium	Medium	High	Medium
Dominance	Codominant	Dominant	Codominant	Codominant
Amount of DNA required	2–10 μg	10–25 ng	50–100 ng	50–100 ng
Sequence information required?	No	No	Yes	Yes
Radioactive detection required?	Yes/No	No	No	No
Development costs	Medium	Low	High	Medium/High
Start-up costs	Medium/High	Low	High	High

both of the primer binding sites, and are visible as the presence or absence of a particular RAPD band. Such polymorphisms behave as dominant genetic markers. Analysis of RAPD markers is attractive for automated breeding applications because it requires only small amounts of DNA (15–25 ng), a nonradioactive assay that can be performed in several hours, and a simple experimental set-up.

RAPD technology provided researchers with a quick and efficient screen for DNA sequence-based polymorphisms at a very large number of loci. Sets of short primers (usually 10mers) suitable for RAPD amplification are available commercially or can easily be synthesized. Apart from a thermocycler and an agarose gel assembly, no special equipment is required. Difficulties with assay reproducibility may be overcome by eliminating variation in DNA concentration, and taking care to ensure consistent reaction conditions and thermal profile during amplification.

RAPD markers are dominant (profiles are scored for the presence or absence of a single allele), and mapping experiments have to be designed to take this into account. In both mapping and diagnostic applications, the phase of the diagnostic markers (coupling or repulsion to the trait) must be chosen so as to maximize the usefulness of the information to the plant breeder.

Michelmore[10] showed recently that RAPD bands linked to a trait of interest could be easily identified by using two pooled DNA samples: one from individuals that express the trait, the other from individuals that do not. Any polymorphism between the two pools should be linked to the trait. Markers thus identified could be confirmed by mapping in a segregating population. This technique, named bulk segregant analysis, is now widely used for mapping simple traits with RAPD markers.

Once a marker linked to a trait of interest is found, it is relatively easy to turn the RAPD assay into a more reproducible PCR-type assay based on secondary DNA sequence, by use of allele-specific PCR[11] (AS-PCR), allele-specific ligation[12,13] or a sequence-characterized amplified region (SCAR) assay[14] (see Box 1 for brief explanations of these techniques). Additional sequence polymorphisms in RAPD bands may also be detected using restriction enzyme digestion[15]. Other techniques that scan for DNA sequence divergence, like the single-strand conformational polymorphism assay[16] (SSCP) or denaturing gradient gel electrophoresis[17] (DGGE) (Box 1) can be applied to individual amplification products derived from RAPD bands. Several reviews and detailed technical discussions of RAPD technology have appeared recently[15,18–21].

Microsatellite repeat polymorphisms

Microsatellite repeats (also known as simple sequence repeats), in particular the dinucleotide repeats $(AC)_n$, $(AG)_n$ and $(AT)_n$, have recently been shown to be abundant and highly polymorphic in eukaryotic genomes[22–24]. In vertebrates, $(AC)_n$ repeats are very prevalent and informative, and have been used to construct genetic maps of the human, rat and mouse genomes[25–28]. In plants, Akkaya *et al.*[29] and

Box 1. Explanation of techniques used to identify polymorphic genetic loci

RAPD, AP-PCR, DAF All refer to DNA amplification using single random primers. All three techniques share the same principle, but there are some differences in the experimental details. RAPD bands are usually detected on agarose gels using ethidium bromide staining[7]. AP-PCR products are frequently analysed on acrylamide gels and detected by autoradiography[8]. DAF products are analysed on acrylamide gels and detected by silver staining[9].

Allele-specific PCR (AS-PCR)[11] Refers to amplification of specific alleles, or DNA sequence variants, at the same locus. Specificity is achieved by designing one or both PCR primers so that they partially overlap the site of sequence difference between the amplified alleles. Many variants of this technique have been described, under different names.

Allele-specific ligation[12] A technique permitting discrimination of two alleles at a locus by providing two short synthetic oligonucleotides that would bind adjacent to each other on an amplified DNA fragment, and could be ligated in the presence of DNA ligase. If one of the alleles contains a mutation overlapped by the 3' end of one oligonucleotide, its ligation to the oligonucleotide bound 3' to it would be prevented. To deduce the identity of the unknown allele, differentially labeled oligonucleotide pairs may be designed for each allele, and their ligation efficiency compared in the presence of the unknown allele.

Sequence-characterized amplified regions (SCARs)[14] A name given to the genetic markers derived by specific PCR amplification of individual RAPD bands, by using longer (24mer) specific primers based on the DNA sequence of the RAPD bands of interest. SCARs may be dominant-polymorphic (see AS-PCR), or may show size polymorphism (and codominance). Polymorphism may also be identified by restriction digestions before or after amplification (see CAPS, described in the text). A more pleasing acronym for these markers would be STARs (sequence-tagged amplified regions).

Single-strand conformational polymorphism (SSCP)[16] Relies on secondary and tertiary structure differences between denatured and rapidly cooled amplified DNA fragments that differ slightly in their DNA sequence. Different SSCP alleles are resolved on non-denaturing acrylamide gels, usually at low temperatures. The ability to resolve alleles depends on the conditions of electrophoresis.

Denaturing gradient gel electrophoresis (DDGE)[17] Resolves partially denatured double-stranded DNA in precisely defined conditions of temperature and denaturant concentration. Different alleles may denature to various extents under such conditions, and migrate differently on DDGE acrylamide gels.

Morgante and Olivieri[30] found that simple sequence repeats are very polymorphic and that AT repeats are more numerous than AC repeats, at least among the DNA sequences in genome databases. For example in soybean breeding germplasm, RFLP loci usually have no more than two alleles, but AT microsatellites have up to eight alleles (M. Morgante *et al.*, pers. commun.). At present, we do not know enough about the relative abundance of AT, AC and AG repeats in plants to assess their comparative value as genetic markers.

Simple sequence repeats are analysed by PCR amplification of a short genomic region containing the repeated sequence, and size estimation of the repeat length by gel separation. Only a small quantity of DNA is required, and it is possible that the automation technology developed to handle RAPDs or PCR-amplified samples could be applied directly to simple repeat analyses. Agarose gels can be used for band separation[30], but resolution of all alleles may require the use of acrylamide gels. When the sequences of primers specific for a given microsatellite marker are published, this technology is made available to the whole scientific community without any need to distribute clones or other materials. However, a considerable amount of effort and expense is involved in identifying enough microsatellite markers to create a complete genetic map. Such projects have been undertaken for several vertebrate genomes[25,26,31] and a similar effort to generate freely available microsatellite markers for major crop and model plant systems would surely prove worthwhile. The funding of such projects should be a high priority for plant genome initiatives such as those of the US Department of Agriculture and the European Community. It may not be cost-effective to establish microsatellite maps for minor crops.

Cleaved amplified polymorphic sequence

Another assay has recently been developed that captures some of the advantages of the RFLP assay, while avoiding the disadvantages of Southern blot analysis. In this approach[32,33], named cleaved amplified polymorphic sequence (CAPS), partial DNA sequence information for the locus of interest is used to create a set of PCR primers. These are used to amplify a segment of DNA at the locus using samples from several different individuals. The amplified bands are then digested with a number of restriction enzymes to identify RFLPs among individuals. Because of the limited size of the fragments that can be amplified by PCR, RFLP polymorphisms in the amplified band may be more difficult to identify than 'classical' RFLPs; however, the technique has the advantage that many restriction enzymes may easily be tested on amplified DNA. CAPS could easily be applied in species like maize, for which there is already a wealth of DNA sequence information in the databases. Random sequencing of *Pst*I genomic clones or sequencing of individual RAPD bands could quickly generate a wealth of sequence information for use in this assay.

Comparison of diagnostics in plant breeding and human genetics

The requirements of a diagnostics program for agricultural genetics are quite different from those of a program for a human genetic or forensic diagnostics. In human and forensic applications, one needs to have

extremely high confidence in the accuracy of each assay, but relatively few assays are required. The value of each assay is therefore very high.

By contrast, plant breeding programs require a genetic diagnostic assay that is relatively inexpensive and can be performed on thousands of individuals, but an accuracy of 95% should be adequate. One breeder may make phenotype-based selections from 100 000 or more individual plants in a single breeding season (120 days). To provide genotypic information on just 10% of these individuals at 100 loci would involve 1×10^6 assays. Even if just one locus is assayed (two markers flanking a gene of interest), 20 000 assays would have to be performed over the breeding season.

These criteria make DNA amplification assays very attractive. The high cost of DNA polymerase remains an issue, but it is hoped that the price will fall in the near future. Automation technology, which promises to increase the assay throughput per technician, is indispensable.

The enabling technology

Molecular breeding requires the following technology:

- genetic maps
- genetic markers linked to agronomic traits
- high throughput, automated diagnostic techniques
- a modification in breeding practice that takes full advantage of the information provided by genome diagnostics

Great progress has been made in constructing genetic maps in many plant species, including all major crops[34]. Most of these maps are made using RFLP markers, and many new ones are being assembled using both RAPD and microsatellite markers[25,35].

Identification of molecular markers that are linked to traits of interest is facilitated by the availability of highly saturated genetic maps. Mapping a monogenic trait is simple as long as it can be assayed reliably. Identification of genetic markers linked to genes that contribute to quantitative characteristics (like yield) is more complicated[1]. Many traits of agronomic interest are encoded by several genes, the expression of which is influenced by environmental factors. Very large genetic experiments are required to map these traits with statistical significance. In addition, usually only a subset of the loci of interest may be mapped in any one experimental cross, so that multiple mapping experiments are needed. It is hoped that technological improvements in mapping efficiency will make quantitative trait loci (QTL) mapping experiments yield better data.

As important agronomic traits are mapped, plant breeders will want to use this information in their breeding programs. We believe that RFLP markers are not suitable for large-scale agricultural diagnostic applications. Southern blotting has been successfully automated by D. Cohen and colleagues at Généthon, and diagnostic services based on Southern blot analyses are available. However, it appears to us that delivery of genetic diagnostic services to the plant breeding industry will have to come from automated assays based on DNA amplification, if both cost and throughput requirements are to be met. As this technology becomes widely available, plant breeders will need to modify their breeding practices to make full use of it. In addition to the phenotype-based selection that is used today, genotype-based selections will be made routinely. As yet, relatively few comparative studies of phenotype- versus genotype-based selection are available.

Automation: needs and progress

The following steps in a genetic diagnostic assay can be automated:

- DNA extraction
- DNA quantitation
- assembly of the amplification reaction
- allele analysis
- data readout

We are not aware of a completely integrated diagnostic system that incorporates all these. However, several laboratories have made significant progress in automating some of the stages. Semi-automated extraction of plant DNA can be performed using a commercial apparatus (Applied Biosystems) operating in batch mode. Williams and Rosanio[36] have automated the DNA extraction process using a disposable vessel that is used for both tissue collection and DNA extraction. This process is designed to produce 5–10 μg of DNA per tissue sample, and the output of the machine is one sample per minute in continuous operation.

Several laboratories have automated the assembly of PCR and RAPD reactions by using commercially available pipetting robots (Ref. 37; R. Sederoff, pers. commun.). The reactions may be assembled in microplates, 96-tube arrays or custom-made reaction vessels compatible with the thermocycler[37].

At present, agarose gels are used to visualize the amplification products and to identify allele composition. Gel loading has been automated in several laboratories[42], but gels still need to be handled manually during electrophoresis, visualization and image recording, and an assay that did not involve electrophoresis would be a considerable advantage. DNA capture methods can be used to identify specific alleles present in amplification reactions in 96-well plates without any need for gel separation. In this approach, developed for AS-PCR[11,38], one of the PCR primers is labeled with biotin, and the other with a fluorescent tag or an antigen. The amplification product is captured on an avidin-coated surface, and detected using either fluorescence or an enzyme-coupled anti-fluorescein antibody. Several variants of this method have been developed[39–43]. Such assays could be conveniently automated using commercially available equipment developed for automatic ELISA assays[44]. A number of other assays based on DNA sequence polymorphisms, like the allele-specific ligation[12,13] or nucleotide incorporation assays[45], are also amenable to automation[13].

To take full advantage of laboratory automation, documentation and analysis of results should, as far as possible, also be automated. Database storage of high-resolution digital images means that gels can be documented in an electronic form. Progress has also been made in computer-assisted data read-out from gel images, but to our knowledge no reliable

completely automated system exists as yet. Reading and recording of data from colorimetric or fluorimetric plate assays is readily automated.

Handling and use of data in plant breeding

The usefulness of DNA-based genetic diagnostic information will depend on its accessibility to the plant breeder, and the way in which it is used in conjunction with the results of phenotypic and biochemical assays (Fig. 1). Suitable computer software is being developed (M.K. Hanafey *et al.*, unpublished computer software); in addition to allowing graphic display and analysis of genotypes[40], such software should permit computer simulation of breeding experiments, so that the most efficient use of genetic diagnostic resources can be modeled.

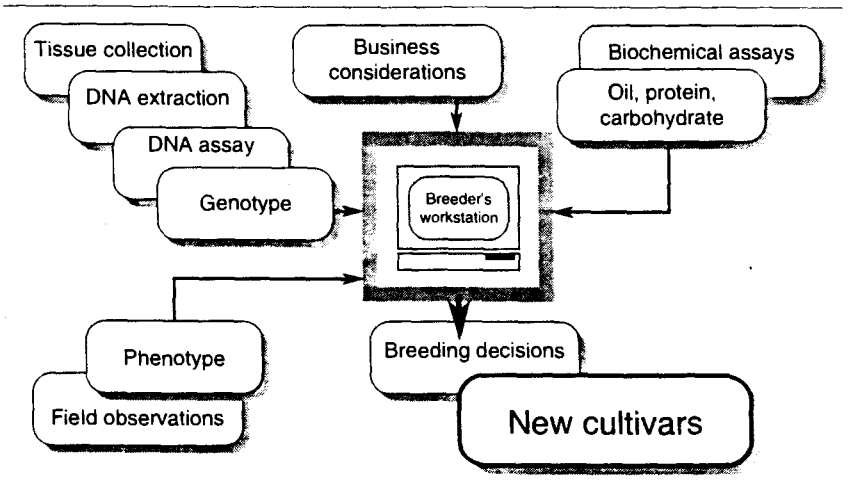


FIG 1

Information flow in marker-assisted plant breeding.

Conclusions and perspectives

Many classes of genetic markers and DNA-based diagnostic techniques are available to plant geneticists and breeders. Their efficient application in plant breeding will depend on the development of cost-effective technologies that automate genetic diagnostic assays in a way similar to the automation of many clinical diagnostic assays. Their use will also depend on computer technology that will display the resulting complex data in a way that is easily understood and accessible to the plant breeder. Considerable progress has been made in this area, but a lot remains to be done. There is a need for a cultural change in the plant breeding community, so as to accept the need for designing breeding schemes that take account of the efficiency of genetic selection. This may be achieved, in part, by in-depth coverage of genetic mapping and diagnostic technology in Plant Breeding curricula.

Acknowledgements

We acknowledge the help and participation of our colleagues at DuPont, Asgrow Seeds and Holden Seeds in developing a molecular breeding program at DuPont. We thank N. Caraher and the anonymous reviewers for helpful suggestions.

References

1 Paterson, A.H. *et al.* (1988) *Nature* 335, 721–726
2 Beckmann, J.S. and Soller, M. (1983) *Theor. Appl. Genet.* 67, 35–43
3 Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989) *BioTechnology* 7, 257–264
4 Young, N.D. and Tanksley, S.D. (1989) *Theor. Appl. Genet.* 77, 353–359
5 Botstein, D., White, R.L., Skolnick, M.H. and Davis, R.W. (1980) *Am. J. Hum. Genet.* 32, 314–331
6 Allefs, J., Salentijn, E., Krens, F.A. and Rouwendal, G. (1990) *Nucleic Acids Res.* 18, 3099–3100
7 Williams, J.G.K. *et al.* (1990) *Nucleic Acids Res.* 18, 6531–6535
8 Welsh, J. and McClelland, M. (1990) *Nucleic Acids Res.* 18, 7213–7218
9 Caetano-Anolles, G., Bassam, B.J. and Gresshoff, P.M. (1991) *Bio/Technology* 9, 553–557
10 Michelmore, R.W., Paran, I. and Kesseli, R.V. (1991) *Proc. Natl Acad. Sci. USA* 88, 9828–9832
11 Wu, D.Y., Ugozzoli, L., Pal, B.K. and Wallace, R.B. (1989) *Proc. Natl Acad. Sci. USA* 86, 2757–2760
12 Landegren, U., Kaiser, R., Sanders, J. and Hood, L. (1988) *Science* 241, 1077–1080
13 Nickerson, D.A. *et al.* (1990) *Proc. Natl Acad. Sci. USA* 87, 8923–8927
14 Paran, I. and Michelmore, R.W. (1993) *Theor. Appl. Genet.* 85, 985–993
15 Williams, J.G.K., Rafalski, J.A. and Tingey, S.V. (1993) *Methods Enzymol.* 218, 704–740
16 Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1989) *Genomics* 5, 874–879
17 Myers, R.M., Maniatis, T. and Lerman, L.S. (1987) *Methods Enzymol.* 155, 501–527
18 Rafalski, J.A., Tingey, S.V. and Williams, J.G.K. (1991) *AgBiotech News Inf.* 3, 645–648
19 Waugh, R. and Powell, W. (1992) *Trends Biotechnol.* 10, 186–191
20 *Applications of RAPD Technology to Plant Breeding* (1992), p. 52, Crop Science Society of America, Minneapolis
21 Tingey, S.V. and del Tufo, J.P. (1993) *Plant Physiol.* 101, 349–352
22 Weber, J. and May, P.E. (1989) *Am. J. Hum. Genet.* 44, 388–396
23 Tautz, D. (1989) *Nucleic Acids Res.* 17, 6463–6471
24 Litt, M. and Luty, J.A. (1989) *Am. J. Hum. Genet.* 44, 397–401
25 Todd, J.A. (1992) *Hum. Mol. Genet.* 1, 663–666
26 Dietrich, W. *et al.* (1992) *Genetics* 131, 423–447
27 Serikawa, T. *et al.* (1992) *Genetics* 131, 701–721
28 Hazan, J. *et al.* (1992) *Genomics* 12, 183–189
29 Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992) *Genetics* 132, 1131–1139
30 Morgante, M. and Olivieri, A.M. (1993) *Plant J.* 3, 175–182
31 Weissenbach, J. (ed.) (1992) *The Généthon Microsatellite Map Catalogue*, p. 122, Généthon
32 Akopyanz, N., Bukanov, N.O., Westblom, T.U. and Berg, D.E. (1992) *Nucleic Acids Res.* 20, 6221–6225
33 Konieczny, A. and Ausubel, F.M. *Plant J.* (in press)

34 O'Brien, S. (ed.) (1993) *Genetic Maps: Locus Maps of Complex Genomes Vol. 6*, (6th edn), Cold Spring Harbor Laboratory Press
35 Reiter, R.S. *et al.* (1992) *Proc. Natl Acad. Sci. USA* 89, 1477–1481
36 Williams, J.G. and Rosanio, L. (1992) *US Pat.* 5, 114, 858
37 Garner, H.R., Armstrong, B. and Lininger, D.M. (1993) *BioTechniques* 14, 112–115
38 Skolnick, M.H. and Wallace, R.B. (1988) *Genomics* 2, 273–279
39 Nagata, Y. *et al.* (1985) *FEBS Lett.* 183, 379–382
40 Syvanen, A.C., Bengtstrom, M., Tenhunen, J. and Soderlund, H. (1988) *Nucleic Acids Res.* 16, 11327–11338
41 Keller, G.H., Huang, D-P., Shih, W-K. and Manak, M.M. (1990) *J. Clin. Microbiol.* 28, 1411–1416
42 Keller, G.H., Huang, D-P. and Manak, M.M. (1991) *J. Clin. Microbiol.* 29, 638–641
43 Adler, K., Buzby, P., Bobrow, M. and Erickson, T. (1992) *DuPont Biotech Update* 7, 160–164
44 Holodniy, M., Winters, M.A. and Merigan, T.C. (1992) *BioTechniques* 12, 36–39
45 Syvanen, A.C. *et al.* (1990) *Genomics* 8, 684–692
46 Young, N.D. and Tanksley, S.D. (1989) *Theor. Appl. Genet.* 77, 95–101

J.A. RAFALSKI AND S.V. TINGEY ARE IN BIOTECHNOLOGY RESEARCH IN THE AGRICULTURAL PRODUCTS DEPARTMENT, E.I. DU PONT DE NEMOURS AND CO., PO BOX 80402, EXPERIMENTAL STATION, WILMINGTON, DE 19880-0402, USA.

Genetic skin diseases caused by mutations in keratin intermediate filaments

PETER M. STEINERT AND SHERRI J. BALE

Keratin intermediate filaments are the major differentiation products of epithelial cells such as the epidermis. The filaments are highly dynamic entities involved in the maintenance of the structural integrity of both the individual cells and the entire tissue. Recent biochemical studies suggest that the keratin proteins overlap each other in several key locations when packed together in filaments. Interestingly, mutations that introduce inappropriate amino acid substitutions in at least some of these overlap regions cause defective keratin filaments that result in at least three classes of autosomal dominant skin disease.

Virtually all eukaryotic cells contain a complex cytoskeleton composed of three different structural proteins: actin-containing microfilaments; tubulin-containing microtubules; and intermediate filaments (IFs). This last class is by far the most complex in terms of the numbers of different constituent protein chains and their modes of expression during development and differentiation. Six distinct types of intermediate filaments have now been recognized. Those expressed in epithelia are the keratin intermediate filaments (KIFs), which are obligatory copolymers of type I and type II keratin protein chains[1–4].

Within the past two years, basic research on the structures of KIFs and their dynamics in living cells has enabled an understanding of how mutations in the constituent keratin proteins cause pathology. Here, we summarize some of these salient points.

Epithelial tissues serve as protective flexible linings throughout the body. Much of this flexibility is provided by a complex and elaborate cytoskeletal network of KIFs that connects with the nuclear surface, courses throughout the cell, and appears to provide continuity into neighboring cells of the epithelium through specialized cell junction points termed desmosomes[5]. Like other IF types, KIFs are highly dynamic structures. Recent data shows that in living cells IF proteins are continually exchanged along the length of the filament, in response to experimental manipulation, stages of the cell cycle, mitosis, and epithelial cell differentiation[4,6–9]. This dynamic behavior is a cycling process, involving phosphorylation by a variety of protein kinases that promote disassembly of the IF into subfilamentous 'soluble' forms, and dephosphorylation by specific phosphatases that promote spontaneous IF reassembly[3,4,10–12].

While the type I and type II keratin families are large, each consisting of about 15 different proteins, most epithelial tissues express only a defined pair of type I and type II keratins, which is characteristic for a given epithelium[13]. For example, simple single-layer epithelial tissues such as ductal linings express keratin K8 (type II) and K18 (type I). In more complex stratified squamous epithelia such as the epidermis, the proliferative basal cells express the K5 (type II) and K14 (type I) pair; epidermal cells committed to terminal differentiation predominantly express the K1 (type II) and K10 (type I) pair, with smaller amounts of K2e or K2p (type II) and K9 (type I) in certain parts of the body[13]. Hair and related 'hard' keratinizing tissues express one or two pairs of other unique keratin proteins.

Packing of molecules in keratin intermediate filaments

All IFs, and KIFs in particular, are built from protein chains with a common organization: a central α-helical rod domain that has a precisely conserved secondary structure, and flanking amino- and carboxy-terminal domains. The exact sequences and properties of these three domains have been used to categorize the six known types of intermediate filaments[1–4]. A model for the type I and type II keratins is shown in Fig. 1a. In this model, the central rod domain is subdivided into four discrete α-helical segments, 1A, 1B, 2A and 2B, which have a repeating heptad amino acid motif that has the potential to form a two-chain coiled-

dosperm development in maize. *Genetics* 107:103–15
39. Lin B-Y. 1982. Association of endosperm reduction with parental imprinting in maize. *Genetics* 100:475–86
40. Lin B-Y, Coe EH Jr. 1986. Monosomy and trisomy induced by the *r-xl* deletion in maize, and associated effects on endosperm development. *Can. J. Genet. Cytol.* 28:831–34
41. Lucchesi J, Rawls JM Jr. 1973. Regulation of gene function: A comparison of enzyme activity levels in relation to gene dosage in diploids and triploids of *Drosophila melanogaster*. *Biochem. Genet.* 9:41–51
42. Mangelsdorf PC, Fraps GS. 1931. A direct quantitative relationship between vitamin A in corn and the number of genes for yellow pigmentation. *Science* 73:241–42
43. McClintock B. 1929. A cytological and genetical study of triploid maize. *Genetics* 14:180–227
44. Muntzing A. 1933. Hybrid incompatibility and the origin of polyploidy. *Hereditas* 18:33–55
45. Nishiyama I, Inomata N. 1966. Embryological studies on cross-incompatibility between 2x and 4x in *Brassica*. *Jpn. J. Genet.* 41:27–42
46. Nishiyama I, Yabuno T. 1978. Causal relationships between the polar nuclei in double fertilization and interspecific cross-incompatibility in *Avena*. *Cytologia* 43:453–66
47. Punyasingh K. 1947. Chromosome numbers in crosses of diploid, triploid and tetraploid maize. *Genetics* 32:541–54
48. Rabinow L, Nguyen-Huynh AT, Birchler JA. 1991. A trans-acting regulatory gene that inversely affects the expression of the *white, brown* and *scarlet* loci in *Drosophila*. *Genetics* 129:463–80
49. Randolph LF. 1932. Some effects of high temperature on polyploidy and other variations in maize. *Proc. Natl. Acad. Sci. USA* 18:222–29
50. Randolph LF. 1935. Cytogenetics of tetraploid maize. *J. Agric. Res.* 50: 591–605
51. Reddy PS, D'Cruz R. 1969. Mechanism of apomixis in *Dichanthium annulatum* (Forssk) Stapf. *Bot. Gaz.* 130: 71–79
52. Redei G. 1964. Crossing experiences with polyploids. *Arabidopsis Inf. Serv.* 1:13
53. Rhoades MM, Dempsey E. 1966. Induction of chromosome doubling at meiosis by the *elongate* gene in maize. *Genetics* 54:505–22
54. Rutishauser A. 1954. Die Entwicklungserregung des Endosperms bei pseudogamen *Ranunculus*-Arten. *Mitt. Naturforsh. Ges. Schaffhausen* 25:1–45
55. Sarkar KR, Coe EH Jr. 1971. Anomalous fertilization in diploid-tetraploid crosses in maize. *Crop Sci.* 11:539–42
56. Satina S, Blakeslee AF, Avery AG. 1938. Chromosome behavior in triploid Datura. III. The seed. *Am. J. Bot.* 25:595–602
57. Wangenheim KH von. 1962. Zur Ursache der Abortion von Samenlagen in Diploid-Polyploid-Kreuzungen. *Z. Vererbungsl.* 93:319–34

Annu. Rev. Genet. 1993. 27:205–233

MAPPING POLYGENES

Steven D. Tanksley
Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853

KEY WORDS: quantitative genetics, quantitative trait loci, gene action, heterosis

CONTENTS

INTRODUCTION

Macromutations (alleles with major phenotypic effects) have greatly simplified the lives of geneticists. Wingless flies, hairless mice, and dwarf plants can all be conditioned by macromutations at single loci. Loci for which macromutations exist are easy to study with the tools of Mendelian genetics since they allow the genotype of a particular locus to be unambiguously inferred from the phenotype of the individual. It is for this reason that, until recently, most of the loci displayed on genetic linkage maps corresponded to loci for which macromutations were available. While macro-

0066-4197/93/1215-0205$02.00

mutations are common in the laboratory, they are much rarer in nature. Most of the macromutations found in genetic stocks seldom occur in natural populations of a species. There is good reason since many, if not most, macromutations are deleterious to the individual and would be weeded out of the population by natural selection.

While the incidence of macromutations sustained in natural populations is relatively low, these populations are by no means deficient in phenotypic or genetic variation. However, the phenotypic variation is usually continuous instead of discrete and conditioned by allelic variation at several (and some times many) genetic loci, each with a relatively small effect. Characters whose phenotypic variation is continuous and determined by the segregation of multiple loci have often been referred to as *quantitative* traits and the inheritance as *polygenic*. The individual loci controlling a quantitative trait are referred to as polygenes or quantitative trait loci (QTL) (see later section for discussion). In nature, most genetic variation in readily observable traits is polygenic.

Lack of discrete phenotypic segregation has all but prevented the use of classical Mendelian techniques for studying polygenes. In the early part of this century a subspecialty of genetics, "quantitative genetics", began to emerge for dealing with quantitative traits. Instead of using discrete phenotypic segregation to study the inheritance of individual loci, quantitative genetics relied upon statistics to describe the characteristics of continuous phenotypic distributions. From these statistics several things could be estimated, including the approximate number of loci affecting the character in a particular mating, the average gene action (e.g. dominance, recessiveness), and the degree to which the various polygenes interact with each other and the environment in determining the phenotype (21, 36). What was not easy to decipher with this approach was the magnitude of effect, inheritance, and gene action of any *specific* locus that was affecting the character. While quantitative genetics has made many important contributions to basic genetics as well as to animal and plant breeding, the inability to describe, study, and ultimately clone individual genes affecting quantitative traits has seriously hampered the study of natural variation at loci for which macromutations do not exist.

Tracking Polygenes with Single Gene Markers

In 1923, Sax reported the association of seed size in beans (a quantitatively inherited character) with seed-coat pigmentation (a discrete monogenic trait) (48). He interpreted this finding as the linkage of the single gene controlling seed color with one or more of the polygenes controlling seed size. Subsequent reports have shown linkage of genes controlling quantitative variation with single gene markers (8, 50, 59, 60). The idea of using single gene markers to systematically characterize and map individual polygenes controlling quantita-

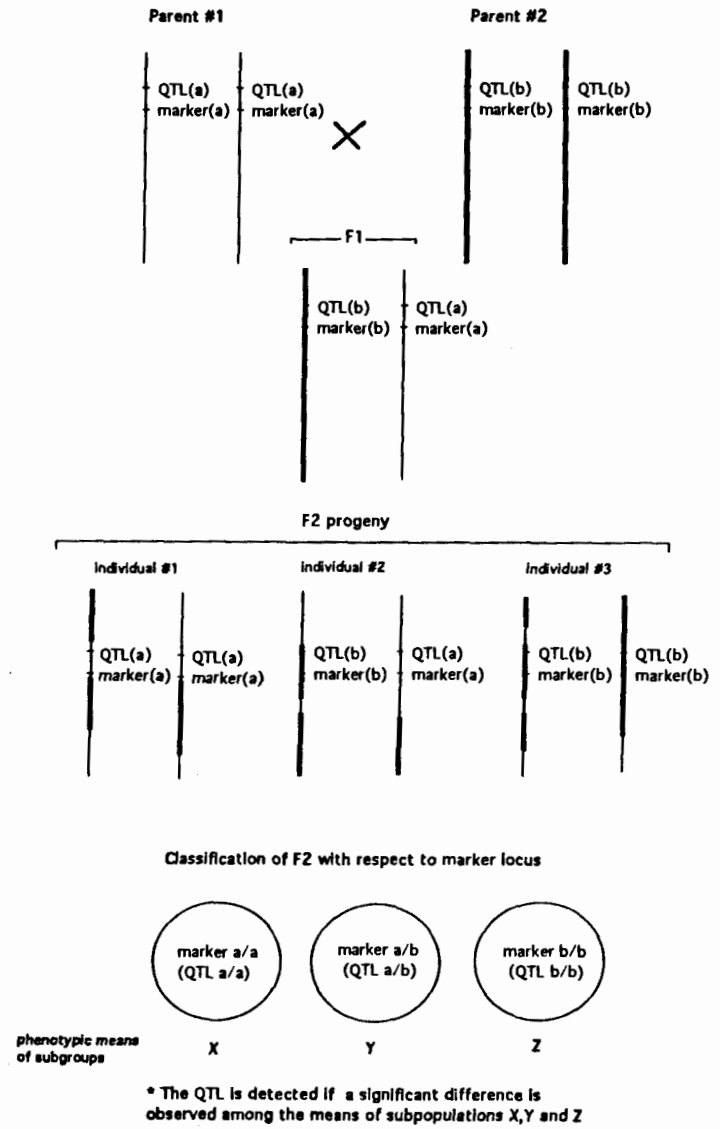


Figure 1 Marker-aided detection of a QTL (quantitative trait locus) in a theoretical F2 population segregating for quantitative trait. Top portion of figure depicts region of chromosomes in parent 1 (thin lines) and parent 2 (thick lines) that contains a marker gene and a linked QTL. F1 hybrid is heterozygous at both marker and QTL. F2 progeny segregate for recombinant chromosomes, but in most instances, genotype of linked marker accurately predicts the genotype for the QTL. Statistical comparisons of the phenotypic means of the subpopulations (comprised of each of the three possible marker genotypes —a/a, a/b, b/b) can permit statistical detection of the linked QTL.

tive traits was put forth by Thoday in 1961 (58). The idea was simple. If the segregation of a single gene marker could be used to detect and estimate the effect of a linked polygene and if single gene markers were scattered throughout the genome of an organism, it should be possible to map and characterize all of the polygenes affecting a character (Figure 1).

Putting Thoday's ideas into practice was difficult and fraught with problems. For most organisms only a few monogenic markers had been mapped, and even for those organisms where such markers were available, most were not suitable for studying quantitative traits. Often the marker genes had a larger effect on the quantitative character than did the linked polygene, thus preventing the detection of the polygene or causing a gross over/underestimate of the polygene's effect. For example, in plants a common monogenic marker is dwarfism. Individuals homozygous for the mutant allele are greatly reduced in stature compared to the wild type. It would be nearly impossible to use a dwarf marker to study the polygenic inheritance of plant height since the effect of the marker gene on height would be much greater than any individual polygene. Other problems with marker genes available at that time were dominance, epistasis, and lack of polymorphism in natural populations (see later section for discussion).

Advent of Molecular Markers in Quantitative Genetics

The discovery that the allelic forms of enzymes (often referred to as allozymes or isozymes) can be separated on electrophoretic gels and detected with histochemical activity stains heralded the era of molecular markers in genetics research (26, 51). No longer was it necessary for a gene to cause a discrete and visible change in the phenotype of an organism in order to study that gene. Enzyme coding genes could be screened for polymorphism in natural populations and mapped genetically using electrophoretic techniques independent of any phenotypic changes (33). By the early 1980s isozyme markers were being employed as a general tool for mapping polygenes and these studies met with considerably more success than previous studies using morphological markers (20, 55, 62, 65). The strategies for using isozyme markers for mapping polygenes were very similar to those proposed by Thoday (58). The difference in success was due to the nature of the markers themselves (see next section).

The next advance in molecular markers came with the introduction of DNA-based genetic markers, the first of which was restriction fragment length polymorphism (RFLP) (7). In the past few years a new generation of DNA-based genetic markers, based on the polymerase chain reaction (e.g. RAPDs and microsatellites), has been developed (63, 67). As with isozymes, allelic variation for DNA-based markers usually has no detectable phenotypic effect. But unlike isozymes, the genetic variation is surveyed directly at the DNA level and thus can reveal more polymorphism.

Attributes of Molecular Markers with Respect to Quantitative Genetics.

The advent of molecular markers has made it feasible to map and characterize the polygenes underlying quantitative traits in natural populations. The key properties that differentiate molecular markers from morphological markers and have permitted the rapid advance of polygene mapping are summarized below.

PHENOTYPIC NEUTRALITY The problem with the marker gene having a larger phenotypic effect than the linked polygene is largely overcome with molecular markers. Alternate alleles at molecular marker loci usually cause no obvious changes in the phenotype of the organism. For example, most allozymes differ by amino acid substitutions that change the migration rate of the protein on a gel, but do not significantly affect the function of the enzyme. For DNA-based markers, most of the allelic variation is in the noncoding portion of the genome. Phenotypically neutral molecular markers not only made it easier to detect linkage between the segregating marker and the polygene, but also provided an unbiased way to estimate the phenotypic effect of each polygene without interference by the marker locus.

POLYMORPHISM As discussed earlier, morphological markers identified by macromutant alleles are rare in natural populations. Without allelic variation there is no segregation, and without segregation no linkage tests can be performed to detect polygenes. The level of polymorphism maintained at any given locus in natural populations is determined by many factors, including population size, mating habit, selection, mutation rate, and migration. Two of these factors, relaxed selection pressure and higher mutation rates (in some types of DNA-based markers), cause allelic variation to be higher at molecular marker loci than at morphological marker loci. In addition, the laboratory techniques used to monitor molecular markers (e.g. gel electrophoresis, restriction enzyme analysis, and polymerase chain reaction amplification) are more sensitive in detecting existing variation (such as missense mutations and small deletions/insertions) at genetic loci than are the phenotypic screening methods used for scoring morphological marker loci. As a result, the proportion of informative (segregating) molecular marker loci can be high in crosses between individuals from the same or different populations.

ABUNDANCE If enough segregating marker genes are scattered throughout an entire genome, it is theoretically possible to detect and characterize all of the polygenes affecting a quantitatively inherited character. However, if some chromosomal regions are devoid of segregating markers, there is no

chance of detecting polygenes located in those areas. The number of useful morphological markers for quantitative genetics was so limited that in most studies only a few markers were used, representing only a small fraction of the genome. The situation improved with isozyme markers, but the number of markers was limited by the number of available enzyme activity stains and in no instances were there enough informative isozyme markers to cover an entire genome (20, 55, 62).

DNA-based markers solved the problem of limited marker abundance. By the late 1980s, complete RFLP linkage maps were available for several organisms and in 1988 the first study was published in which molecular markers, covering an entire genome, were used to map quantitative traits (43). The availability of complete genome maps also opened up the opportunity for new statistical approaches for detecting polygenes (e.g. interval analysis), which will be discussed later. Today, molecular linkage maps covering the entire genome are available for quantitative-trait studies in many organisms including humans (64), mice (12, 17), rats (49), and many plants (11, 44, 56, 61).

CODOMINANCE For loci with codominant alleles there is a one to one relationship between genotype and phenotype (i.e. all possible genotypes can be deduced directly from the phenotype in any generation). This is not true of loci with dominant-recessive alleles, where only homozygous recessive genotypes can be deduced unambiguously from the phenotype. Alleles of most molecular markers are codominant whereas most morphological marker loci segregate dominant-recessive alleles. Thus, the advent of molecular markers has allowed straightforward polygene mapping in virtually any segregating generation (e.g. F2, F3, backcross, recombinant inbreds).

EPISTASIS Epistasis is a form of interaction between nonallelic genes whereby one gene interferes with the phenotypic expression of another gene (4). With morphological marker loci, strong epistatic interactions among loci limit the number of segregating markers that can be unequivocally scored in the same generation. Because molecular marker loci do not normally exhibit epistatic or pleiotropic effects, a virtually limitless number of segregating markers can be used in a single population for mapping polygenes through an entire genome.

STRATEGIES FOR DETECTING POLYGENES

The underlying assumption of using marker loci to detect polygenes is that linkage disequilibrium exists between alleles at the marker locus and alleles of the linked polygene(s). Linkage disequilibrium can be defined as the nonrandom association of alleles at different loci in a population and can be caused by a number of factors including selection and genetic drift. However, in primary segregating generations (e.g. F2, F3, or backcross populations) the predominant cause of linkage disequilibrium is physical linkage of loci and this has formed the basis for classical linkage mapping for the past century. Linkage disequilibrium due to physical linkage of loci is at its highest value in populations derived from controlled matings and as a consequence the ability to map and characterize polygenes using marker loci is also at its highest.

Controlled Matings

Backcross or F2/F3 populations have been used most commonly for detecting linkage between molecular markers and genes controlling quantitative traits (34, 43, 52, 55, 62). In species where severe inbreeding is tolerated, recombinant inbred populations (derived by inbreeding F2 progeny until they become virtually homozygous lines by selfing or sibbing) have also been used (9, 57). While the latter case has less linkage disequilibrium, due to more opportunity for meiotic recombination, it has the advantage of homozygous lines that can be replicated and retested for more accurate measurement of the quantitative trait (9).

There are several statistical procedures for determining whether a polygene is linked to a marker gene and they all share the same basic principle: to partition the population into different genotypic classes based on genotypes at the marker locus and then to use correlative statistics to determine whether the individuals of one genotype differ significantly compared with individuals of other genotype(s) with respect to the trait being measured. If the phenotypes differ significantly, it is interpreted that a gene(s) affecting the trait is linked to the marker locus used to subdivide the population (Figure 1). The procedure is then repeated for additional marker loci throughout the genome to detect as many QTL as possible. Normally, it is not possible to determine whether the effect detected with a marker locus is due to one or more linked genes affecting the trait. For this reason, the term quantitative trait locus (QTL) was coined to describe a region of a chromosome (usually defined by linkage to a marker gene) that has a significant effect on a quantitative trait. Determining whether a QTL is comprised of one or more genes is still one of the most difficult aspects of quantitative genetics (see later section).

SINGLE POINT ANALYSIS The simplest approach for detecting QTL is to analyze the data using one marker at a time (Figures 1, 2). This approach is often referred to as single point analysis or point analysis and does not require a complete molecular marker linkage map. It is for this reason that

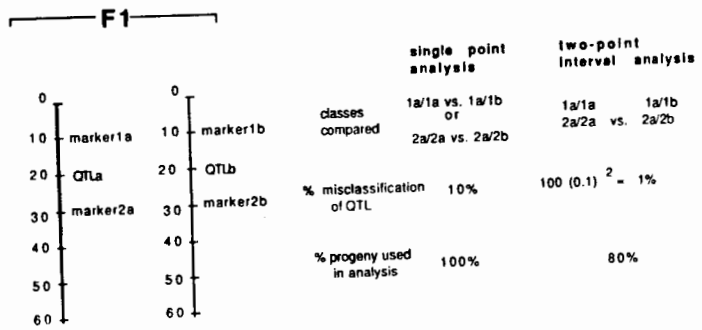


Figure 2 Point analysis versus interval analysis for detecting a QTL in a theoretical backcross population. Left side depicts F1 hybrid genotype for a QTL and two flanking marker loci (numbers on left are centimorgan distances). Note that F1 is heterozygous for both marker loci as well as QTL. For single point analysis (in backcross to parent contributing "a" alleles) statistical comparisons would be made on the *basis* of either the marker 1 genotype *or* marker 2 genotype in the backcross population. In this instance, the QTL genotype will be misclassified approximately 10% of the time since the map distance from the QTL to either marker locus is 10 cM. For interval analysis statistical comparisons are made on the basis of *both* marker loci simultaneously. Misclassification would occur only in the case of a simultaneous crossovers between the QTL and marker 1 and the QTL and marker 2. Misclassification is thus reduced to approximately 1%. Individuals recombinant between marker 1 and marker 2 are not used in the type of interval analysis shown here, therefore the number of individuals analyzed is less (80%) than with single point analysis (100%). In more sophisticated versions of interval analysis, recombinant individuals are also used in the detection and placement of QTL within interval between two markers (31).

point analyses were employed in the first molecular marker/quantitative genetic studies (20, 55, 65, 68). The disadvantages of point analysis are: (*a*) The further a QTL is from the marker gene, the less likely it is to be detected statistically due to crossover events between the marker and QTL that result in misclassification (Figure 2). (*b*) The magnitude of the effect of any detected QTL will normally be underestimated, due also to recombination between the marker locus and QTL. Both problems are minimized when a large number of segregating molecular markers are used, covering the entire genome (usually at intervals less than 15cM). Under these conditions, any potential QTL would be closely linked to at least one molecular marker.

INTERVAL ANALYSIS The availability of molecular linkage maps covering entire genomes has made it possible to overcome some problems with point analysis. To take the fullest advantage of linkage maps for quantitative studies, Lander & Botstein proposed a method called interval analysis (29–31). Instead of analyzing the population one marker at a time, sets of

linked markers are analyzed simultaneously with regard to their effects on quantitative traits (Figure 2). By using linked markers for analysis, it is possible to compensate for recombination between the markers and the QTL, increasing the probability of statistically detecting the QTL and also providing an unbiased estimate of the QTL effect on the character. Interval analysis, using a molecular linkage map of an entire genome, was first demonstrated on an interspecific backcross of tomato (43) and has subsequently been used successfully for several quantitative trait linkage studies (18, 27, 41, 53).

The maximum benefit of interval analysis versus point analysis is realized when linked markers are fairly far apart (> 20cM). Under these conditions there are likely to be many crossovers between the markers and QTL, which can be compensated for with interval analysis. Where the marker density is higher (markers < 15 cM apart) point and interval analysis give nearly identical results (53). When marker loci are very far apart (e.g. > 35 cM), even interval analysis is inefficient in detecting QTL in the interval between the marker loci.

DISTRIBUTIONAL EXTREMES Despite technological improvements in the speed and accuracy with which molecular markers can be assayed, it can still be time consuming and expensive to assay large populations. When the time and expense of assaying molecular markers is significantly greater than measuring the quantitative character of interest on each individual, it is possible to use a modified approach to detect QTL. The approach, also proposed by Lander & Botstein (31), starts with a large segregating population (e.g. F2 or backcross). A quantitative measure of the character of interest is taken on each individual in the population. Marker analysis is performed *only* on individuals in the extreme tails of the distribution (i.e. those with the lowest and highest values for the character) (Figure 3). If the allele frequency at any molecular marker locus differs significantly between the two extreme subpopulations, it is inferred that a QTL controlling the character of interest is located near the marker.

The benefit of distributional extreme analysis is in the savings of time and resources in assaying molecular markers. Given the same number of individuals assayed for molecular markers in total population analysis versus distributional extreme analysis, the statistical power of detecting QTL will be greater for the latter (31). The penalties for distributional extreme analysis are: (*a*) More segregating individuals must be analyzed for the quantitative phenotype to collect enough individuals in the distributional extremes. In many instances this is not a problem. However, in some situations the time and cost of characterizing a large population phenotypically outweighs the advantages; and for some organisms (especially animal species) it is difficult

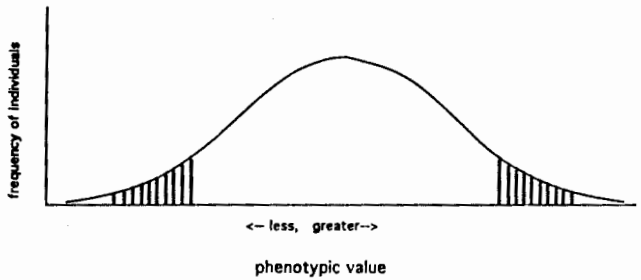


Figure 3 Distributional extreme analysis for detection of QTL. Quantitative characters often assume a continuous distribution (i.e. normal distribution shown in this figure). Individuals with extreme phenotypes (striped portions of distribution) are likely to have a large number of either "positive" or "negative" QTL alleles. Comparison of marker genotypes between individuals in the two extremes can significantly increase the probability of detecting QTL affecting the quantitative trait, but limits the ability to measure the individual effect of each QTL (see text for more detailed discussion).

to generate the large segregating populations required. (*b*) While distribution extreme analysis is more efficient at detecting linkage between marker loci and QTL, it is less efficient in determining individual QTL effects. Individuals in the extremes tend to have either a large number of positive or negative alleles at all QTL, depending on which extreme they represent. There is thus a deficiency of individuals with a mixture of positive and negative alleles in the subpopulations being analyzed, which confounds the ability to individually measure the effects of any specific QTL. (*c*) It is often impractical to use distributional extremes to map more than one quantitative character, since the individuals with extreme phenotypes for one character are not likely to represent the extremes for other characters.

LIMITATIONS ON PROGENY SIZE To detect QTL with molecular markers normally requires analysis of fairly large segregating populations (i.e. > 100 individuals). Although most plants and some animals readily produce offspring in such large numbers, not all species do. Where offspring numbers are limiting, alternative strategies (other than backcross or F2 analysis) must be employed. The types of modified populations employed for QTL mapping are largely a function of the reproductive characteristics of the species under study and the ingenuity of the investigator.

Half-sib analysis Half sibs are defined as the progeny derived from the mating of a single individual to random individuals in the population. If the original individual involved in the mating is heterozygous for both

molecular markers and QTL, then linkage between the molecular markers and the QTL can be detected by analyzing a large half-sib population. Both molecular markers and quantitative traits would be measured on the half sibs and either point or interval analysis would be employed to detect significant association between the alternate alleles donated by the original individual used to create the half sibs. This strategy can be used when the gametes of one sex (usually the female sex) are limiting. Prime candidates for this approach would be farm animals where a single heterozygous male could be mated to a large number of females to create large half-sib populations. Because the phenotype of half sibs is determined only in part (theoretically one half) by genes inherited from the original heterozygous individual, the statistical power of detecting linkage between segregating markers and QTL is decreased significantly, requiring even larger populations.

Random Matings

For most plants and many experimental organisms, it is possible to select any two individuals and make controlled matings between them. Controlled matings have the advantage of allowing the investigator to pick individuals that differ significantly for the character of interest. The greater the phenotypic difference between any two individuals, the more likely one is to detect significant QTL controlling that character in a derived, segregating population. Controlled matings are also advantageous in that they result in maximum linkage disequilibrium (due to physical linkage between loci) for detecting QTL with linked molecular markers. However, for some species, it is either impractical or unethical to make controlled matings. Humans are a prime example. Not only are natural family sizes too small for QTL analyses, but it is necessary to work with the matings that nature has given us. Unfortunately for the geneticists, individuals with the most extreme phenotypes do not always produce offspring. The only practical solution to this problem is to pool quantitative and molecular marker data from a multitude of families and attempt to analyze for QTL linkage using the entire dataset. This creates one of the most difficult problems in polygene mapping—trait heterogeneity.

HETEROGENEOUS TRAITS In random-mating populations, individuals with similar phenotypes do not necessarily have similar genotypes. The phenotype is the result of the combined action of all the genes affecting a character, and different combinations of alleles can create phenotypically similar individuals. The underlying assumption of pooling data from different segregating families is that the phenotypic variation segregating in the families is due to segregation of the same QTL. If this assumption is met,

then pooling data from families will increase the statistical power of detecting QTL. However, if different QTL are segregating in each family, pooling of data will not increase the statistical power and such QTL will likely go undetected. A comprehensive discussion of the problems and potential solutions for mapping polygenes controlling heterogeneous traits is presented by Lander & Botstein (29, 30).

CHARACTERIZATION OF POLYGENIC TRAITS

Number of Polygenes

Classical quantitative genetics theory has yielded several statistical approaches for estimating the number of genes segregating in a population that affect a given morphological character (21, 28, 32, 40). These estimates are based on characteristics of the phenotypic distribution in the segregating population(s) (especially on the phenotypic variance) and often depend on various underlying assumptions, including approximately equal effects of individual polygenes, independent assortment of polygenes, similar gene action (allelic interaction) of polygenes, and minimal epistasis among polygenes. In practice, one or more of these assumptions are normally violated, making the estimates suspect. Furthermore, until recently there was no practical way to independently verify or falsify one's estimates.

Estimating the number of polygenes using molecular marker approaches is straightforward, although not without limitations and biases. The approach is simply to add up the number of QTL detected in a particular study and to use that value as an estimate of the number of segregating polygenes affecting the character in that population. The limitations of this approach are : (*a*) Only those genes with a sufficiently large phenotypic effect to be detected statistically will be counted. Genes with lesser effects will fall below the threshold of detection, depending on the size of the segregating population (Figure 4). The larger the population, the more likely it is to statistically detect genes of lesser effects. Therefore, the molecular marker approach almost always underestimates the number of genes affecting a character and is biased towards the detection of genes of larger phenotypic effects. (*b*) Using typical population sizes (< 500 individuals) two or more polygenes closer together than approximately 20 cM will usually appear as a single QTL (i.e. they cannot be distinguished as separate genes). Again, this is a bias toward underestimating the number of genes.

The advantage of the molecular marker approach to estimate the number of polygenes segregating in a population is that results are quantifiable and testable. They are quantifiable in that the cumulative genetic and phenotypic variance attributable to the QTL detected can be estimated. For example,

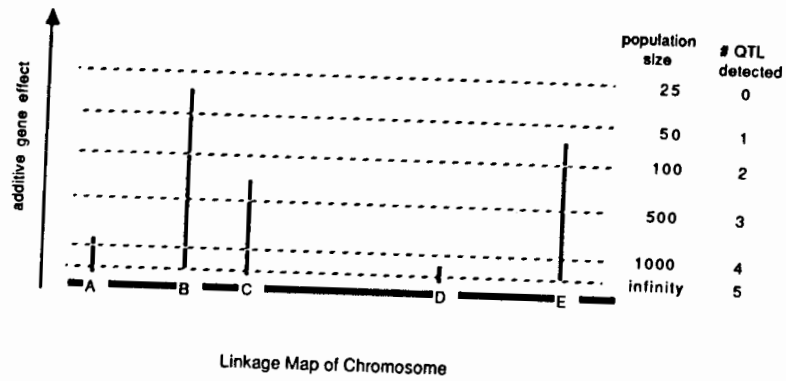


Figure 4 Relationship between the size of a segregating population and the probability of detecting QTL with varying effects. QTL (A,B,C, etc) are shown along X axis which represents a single chromosome. Y axis reflects magnitude of effect of individual QTL and is indicated by vertical height of bar above each QTL. Dotted lines correspond to different population sizes. If the dotted line intersects a vertical bar, the QTL will be detected statistically in that sized population. Note that with small population sizes (e.g. 25 or 50), only the major QTL are detected. To detect the QTL of smaller effects, very large population sizes are required.

if five QTL are detected and together account for 70% of the genetic variance, additional QTL (probably of smaller effects) must account for the remaining 30% of the variance. The results are testable since the experiment can be repeated (even in different environments) to see if the same number and genomic placement of QTL are observed (41, 53). Also, new populations or genetic stocks can be created to study any QTL in further detail, including high-resolution mapping to determine if any particular QTL is composed of more than one gene (see section on high-resolution mapping). Table 1 gives a sample of gene number estimates derived from molecular marker mapping in several organisms for a variety of characters.

Magnitude of Effects

It would be expected that different polygenes should be unequal in the magnitude of effect they exert on a character. However, in the past it was difficult to test this assumption. In the past 10 years enough data have accumulated from marker studies to establish definitively that polygenes do vary widely in their effects and that the assumption of many polygenes with equal effects (required for some of the previous statistical models) is not valid. More surprising has been the finding that in many instances a large proportion of quantitative variation can be explained by the segregation of a few major QTL. It is not uncommon to find individual QTL that can

Table 1 Summary of QTL detected for different characters in a variety of species

Organism	Trait	Population	# QTL	% phenotypic variance max	min	total	Ref.
Tomato	soluble solids	BC1 (237)	4			44	41
		F2/F3 (350	7	28	6.0	44	43
	fruit mass	BC1 (237)	6			58	41
		F2/F3 (350)	13	42	4.0	72	43
	fruit pH	BC1 (237)	5			48	41
		F2/F3 (350	9	28	4.2	34	43
	growth	F2 (432)	5	7	3.2	21	16
	leaflet shape	F2 (432)	9	30	3.1	60	16
	height	F2 (432)	9	8	3.1	42	16
Maize	height	F2 (112)	6			73	5
		F2 (112)	4			53	5
		F2 (144)	3			34	5
		F2 (144)	3			45	5
		F2 (260)	7	35	4.0	67	18
		F2 (187)	11	27	4.0		20
	tiller number	F2 (260)	1			24	18
	leaf length	F2 (260)	7	17	5.0	57	18
	glume hardness	F2 (260)	5	42	6.0	72	18
	grain yield	F2 (1930)	13*	4	0.6		19
		F2 (1776)	18*	5	0.3		19
		F2 (187)	13	17	4.0		20
		F3BC1 (264)	8	15	6	59	53
		F3BC1 (264)	6	21	8	61	53
	number of ears	F2 (1930)	10*	5	0.4		19
		F2 (1776)	9*	4	0.2		19
Common bean	nodule number	F2/F3 (70)	4	17	11	50	39
Mung bean	seed weight	F2 (58)	4	33		50	22
Cowpea	seed weight	F2 (58)	2	37	32.9	53	22
Wheat	pre-harvest sprout	RI (78)	4	14	10.0	37	3
Mice	epilepsy	BC1 (87)	2			50	46
Rats	hypertension	F2 (115)	2	19	11	30	27

*Does not take into account linked molecular markers

account for > 20% of the phenotypic variation in a population (Table 1) and values as high as 42% have been reported for single QTL (18). Since the percent phenotypic variance includes environmental variance, the actual genetic variance attributable to major QTL is even larger (16, 53).

QTL with major effects have been identified for most characters studied, but most by far of QTL reported are those of smaller effects (Figure 5). This result is logical if one considers the possibility that most segregating genes in a population probably have some effect on most characters—albeit very small in most cases. The smallest effect a QTL can have and still be

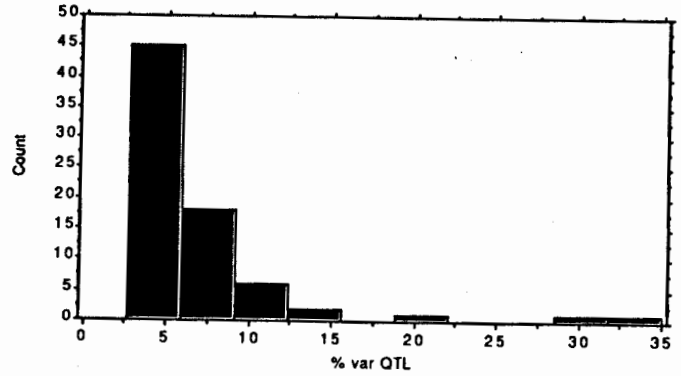


Figure 5 Histogram depicting the percentage of total phenotypic variance attributable to individual QTL derived from a published study in tomato (16). In this study 74 significant QTL were detected for 11 different quantitatively inherited characters. While "major" QTL were detected for most characters (i.e. those accounting for more than 10% of the phenotypic variance), most QTL were of much smaller effect, causing an asymmetrical distribution skewed towards QTL of smaller effects. QTL with individual effects less than 2% were not reported, presumably because the population size was not large enough to statistically detect such QTL.

detected by the marker method depends on a number of factors: (*a*) The map distance from nearest marker to the QTL: the closer a QTL is to a marker, the smaller the effect that QTL can have and still be detected statistically because the effects of QTL closer to the marker are less confounded by recombination events between the marker and the QTL. (*b*) Size of segregating population: the larger the population size, the more likely the effects of lesser QTL will reach statistical significance. (*c*) Heritability of trait: the larger the environmental effect on the character (i.e. lower the heritability), the less likely a QTL will be detected. (*d*) Probability criteria used for declaring a QTL effect significant: higher probability thresholds reduce the chances of spurious QTL being reported, but also reduce the chances of detecting QTL with smaller effects. In maize using an F2 population size of 1700 individuals and a probability threshold of 0.05, QTL contributing as little as 0.3% of the phenotypic variance were reported (20). In experiments with smaller population sizes and higher probability thresholds, QTL that explain less than 3% of the phenotypic variance are not normally detected (Table 1).

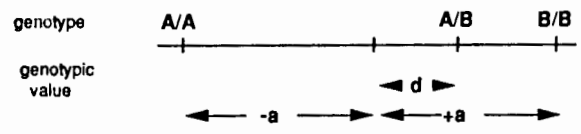
It is difficult to determine how efficient an experiment has been in identifying the QTL responsible for a trait. One measurement of success has been the cumulative phenotypic variance attributable to the combination of all significant QTL. Where complete molecular maps have been used

for quantitative studies, the values have ranged from as high as 95% (18) to less than 10% (16), with the average approximately 30–40% (Table 1).

The bias towards detecting QTL with larger effects means that it is unlikely that one will ever detect, map and characterize all of the polygenes affecting a character in any given segregating population. Theoretically, this may be a problem, but is only a minor practical limitation. Using the marker approach, one is likely to identify and characterize QTL making the largest contribution to the phenotype and it is likely to be these QTL that one would want to further characterize and ultimately to clone. For plant and animal breeders, the major QTL are of greatest interest to manipulate in breeding schemes via association with molecular markers.

Gene Action

In diploids, the two alleles at a genetic locus can interact in a number of ways to produce the phenotype of the individual, often referred to as gene action. In classical genetics, which is normally confined to macromutations, alleles are normally either dominant or recessive. However, in natural populations (and especially in loci underlying quantitative variation) alleles are seldom absolutely recessive or dominant. Instead, the gene action for specific alleles ranges from complete dominance to complete recessiveness to overdominance (heterozygotes exceed either parental homozygote), and everything in between. Quantitative geneticists have devised several parameters to describe such continuous gene action. The dominance/additivity (d/a) statistic is the most common and describes the degree to which the heterozygous genotype resembles the parental homozygotes (Figure 6). A d/a value of 1.0 is considered to be complete dominance and the heterozgote would be identical to the greater parental homozygote (Figure 6). A value of −1.0 is considered complete recessiveness and the heterozygote would



a = additive effect (of a single allele) = (B/B - A/A)/2
d = dominance deviation = A/B - [(A/A + B/B)/2)]
d/a = degree of dominance
(-1 < d/a < 1)

Figure 6 Statistics for individual QTL that can be estimated from marker-aided studies of quantitative traits. The dominance/additivity (d/a) statistic is most commonly used to describe gene action for QTL (see text for discussion).

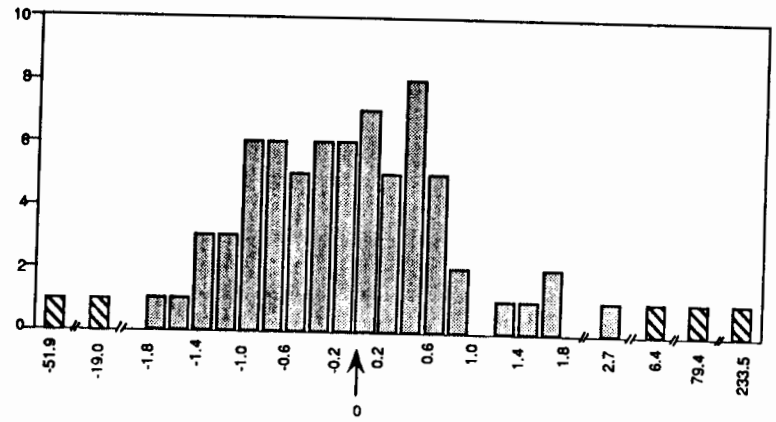


Figure 7 Histogram of individual QTL d/a values derived from a published study in tomato (16). Note the approximate normal distribution of gene action ranging from overdominance (striped bars on right) to underdominance (striped bars on left), with most QTL showing additivity ($1 \leq d/a \leq 1$).

be identical to the lesser homozygote. A value of 0.0 applies to cases where the heterozgote is exactly intermediate between the two homozygotes—a situation referred to as additive gene action. Values greater than 1.0 or less than −1.0 imply overdominance/underdominance.

The histogram in Figure 7 displays the estimated gene action for 74 individual QTL identified in a single mapping study in tomato (16). The continuous distribution of gene action observed in this study is typical of that reported in a variety of other QTL mapping studies (18, 20, 41, 53). This wide range of gene actions for naturally occurring polygenes is in sharp contrast to the dominance/recessiveness exhibited by the majority of alleles utilized in laboratory studies. This can probably be attributed to the fact that most alleles studied in laboratories are derived from either spontaneous or induced mutations, which result in loss of a gene product or production of a defective gene product. Such mutations would be expected to behave in a recessive manner. In nature, such loss-of-function mutations would likely be deleterious or lethal and would be weeded out by natural selection. The type of variation maintained in nature is more likely to be for allelic forms that modify, but do not eliminate, the gene product.

Epistasis

For polygenic traits, the potential number of epistatic interactions (two-locus, three-locus, etc) is enormous. In classical quantitative genetic studies, the genetic variance attributable to epistasis has often been referred to as

genotype-by-genotype interaction (21). In the past it was seldom possible to determine whether such interactions were occurring between specific polygenes or to measure the exact nature of specific interactions. The ability to map and characterize individual QTL via molecular mapping techniques raises the possibility of detecting and characterizing interactions among specific QTL. One simple statistical method for determining two-locus interactions of QTL is through the use of two-way analysis of variance (ANOVA). Once the significant QTL affecting a trait have been detected (using methods discussed previously), any two significant, unlinked QTL can be used as the independent variables and the quantitative character as the dependent variable in a two-way ANOVA. The "interaction" factor calculated in the two-way ANOVA will be an estimate of the interaction between the two QTL in determining the phenotype. A significant interaction factor suggests that the effect rendered by the two QTL together is not simply a sum of their independent effects.

Theoretically, all possible two-way, three-way, and up to *n*-way interaction among QTL can be measured through *n*-way ANOVA or related statistics. Practically, there are several problems (some very serious) with this approach:

1. Segregating population sizes are usually too small for accurate estimates of multi-locus interactions. To measure interaction between two loci, all possible genotypes for those two loci must appear in the population in sufficient frequencies to allow statistical comparisons. Taking the simple case of two independent QTL (A and B) segregating in an F2 population, there would be nine possible two-locus genotypes and the four rarest classes (double homozygotes, e.g. AABB, AAbb, etc) would each occur in an expected 1/16 frequency. If the original population size had been 500 individuals, the expected number of individuals in each of these rare classes would be 31. This number of individuals is a relatively small sample on which to estimate a phenotypic effect and with which to compare that effect with the effects observed in other similarly rare classes. To have a good measure of the phenotypic effect of each class (and thus to gain a more accurate measure of epistasis) would require a much larger population. This is only for the two-way interactions. Three-way interactions would result in 27 genotypic classes in an F2 for which phenotypic effects must be compared and the rarest class would occur at a frequency of 1/64. Segregating populations of thousands of individuals would be required to obtain enough individuals in each class for adequate statistical comparisons.
2. The number of potential multilocus interactions is very large, requiring many statistical tests, some of which will, by chance, reach statistical significance. If N significant QTL were detected for some trait, there would be $N!/2(N-2)!$ possible two-way interactions. In the case of 10 QTL, the number would be 45. Using a significance level of 0.05, one would expect two spurious interactions to be reported. One solution is to raise the significance threshold to avoid reporting spurious interactions. However, as the threshold is increased, it is less and less likely that real interactions will be detected—a problem exacerbated by small population sizes, as discussed above. Higher level interactions are even more problematic statistically. With 10 QTL there would be 240 possible three-way interactions!
3. Interactive effects will be underestimated with simple statistical approaches like *n*-way ANOVA due to recombination between the QTL and the linked molecular marker on which the ANOVAs are based. The problem can be mitigated by using higher density molecular linkage maps for detecting and analyzing QTL such that most QTL are closely linked with at least one molecular marker.

Despite the limitations of estimating epistatic effects among specific loci, an interesting pattern is emerging from molecular marker studies of quantitative traits. In most studies conducted thus far the number of statistically significant QTL interactions has normally been close to the number expected to occur by chance (16, 20, 53, 55). While a few of the specific QTL interactions have had very high probability values (suggesting that they are real and not spurious), most have been near the probability threshold, which would be consistent with them being a statistical artifact. These results would suggest that strong epistatic interactions (i.e. the type observed between some macromutant alleles) are the exception and not the rule for naturally occurring polygenes. These conclusions are supported by the few studies in which individual QTL have been genetically isolated from other QTL in nearly isogenic lines and have been shown to continue producing their same individual effects (16).

Because of the limitations listed above, it seems unlikely that it will be possible to produce detailed descriptions of QTL by QTL interactions using primary segregating populations (e.g. F2, backcross, etc). More likely, it will be necessary for geneticists and breeders to construct nearly isogenic lines with single QTL and combinations of QTL that can be replicated in experimental designs to allow a more precise measurement and description of epistatic interactions.

Environment by QTL Interactions

The phenotype of an individual is conditioned not only by its genotype, but by the interaction of the genotype with the environment. Quantitative

geneticists have long recognized the importance of genotype by environment interactions and have documented numerous cases of such interactions (21, 24, 38). These results suggest that QTL important in one environment may not be as important in determining the phenotype in another environment. Two recent studies involving QTL mapping of the same populations in different environments have provided a test of this prediction.

In maize, seven agronomic traits, including grain yield, were measured in the same population in six different environments in the United States (53). QTL detected in one environment were frequently detected in the other environments, suggesting very little environment by QTL interaction. A related study in tomato examined segregating populations derived from the same original cross in three environments (two in the United States and one in the Middle East) (41). Forty-eight percent of the QTL were detected in at least two of the environments, with the remainder being detected in single environments. The QTL showing the largest effects in one environment were also more likely to be detected in another environment. The two US sites shared more common QTL with each other than they did with the Middle East sites, consistent with major differences in both the climate and horticultural practices.

While it is premature to draw any sweeping conclusions about environment by QTL interactions, the two studies described above suggest that a substantial proportion of QTL affecting a quantitative trait in one environment will be active in other environments and that this is especially true for QTL with major effects. This would be an especially welcome result for plant and animal breeders trying to use linked markers to transfer potentially valuable QTL into modern varieties/breeds in the hope that they will make positive contributions to the phenotype in many diverse environments.

High Resolution Mapping of QTL

In the genetics of diploids, complementation is the acid test for whether or not two genes are allelic. Complementation depends upon the availability of dominant/recessive alleles. Since very few QTL demonstrate such allelic interactions, complementation tests are of little value in determining whether a QTL is comprised of a single gene or several linked genes affecting the character of interest. Geneticists must therefore resort to other strategies for determining the genic composition of a QTL. If the QTL is the result of two or more linked genes, it should be possible to separate those genes by crossing-over. This entails high-resolution mapping of QTL. Unfortunately, high-resolution mapping of QTL in primary generations suffers from the same limitation as does measuring epistasis—segregating population sizes are too small to allow such fine mapping. Fine mapping of QTL involves comparing the means of individuals recombinant for molecular markers in the vicinity of the QTL with individuals that are nonrecombinant. The classes of limiting size will be the recombinant classes. For example, in a backcross population of 500 individuals one would expect approximately 13 individuals in each of the two recombinant classes using two markers that are 5 cM apart. For markers 1 cM apart there would be approximately three individuals in each recombinant class. Current strategies for mapping QTL depend on comparing the means of recombinant and nonrecombinant classes (31, 55). Because of the small proportion of recombinant individuals, the statistical tests are not very powerful and given the practical sizes of segregating populations currently used for QTL studies, mapping resolution of QTL has been limited to approximately 10–20 cMs—inadequate for distinguishing between single gene versus multigene composition of individual QTL.

Paterson et al (42) proposed a solution to this problem in which recombinant individuals are identified in primary generations and selectively multiplied in subsequent generations so that the recombinant classes occur at near equal frequency with the nonrecombinant classes, increasing the power for the statistical comparisons among classes. In the ideal situation, a series of nearly isogenic lines (NILs), differing in recombination in the QTL regions, would be compared for the quantitative trait being mapped. Such comparisons can potentially allow placement of a QTL to a very small interval on the map (Figure 8). NILs have the advantage of being genetically uniform throughout the rest of their genomes, which makes it easier to study the effects of a single QTL. Unfortunately, it is often time consuming to develop NILs. As an alternative, earlier generations can also be used in a similar manner for high-resolution mapping of QTL—albeit with less efficiency (42).

If linked molecular markers are used as a means of selecting for recombinants in the vicinity of QTL for fine mapping, then the number of available molecular markers in the chromosomal region of interest can become a limiting factor. High density molecular linkage maps are being developed for a number of organisms and may provide a solution to this problem. In addition, methods of selectively identifying molecular markers in targeted regions of the genome have been reported and may provide the degree of marker saturation necessary for high-resolution mapping of QTL (23, 35, 37).

Recombination can be used as evidence for multigene composition of a QTL, but cannot be used directly to prove a single gene composition. If a recombinant separates a QTL into two separate QTL, then the multigene hypothesis is supported. Failure to do so could be due to the genes being so tightly linked that the appropriate recombinants were not recovered in the population examined. However, the more recombinants examined in an area, the more likely that multiple genes would be separated. Recovering crossovers at intervals of 0.1 cM (or even 0.01 cM), which might be

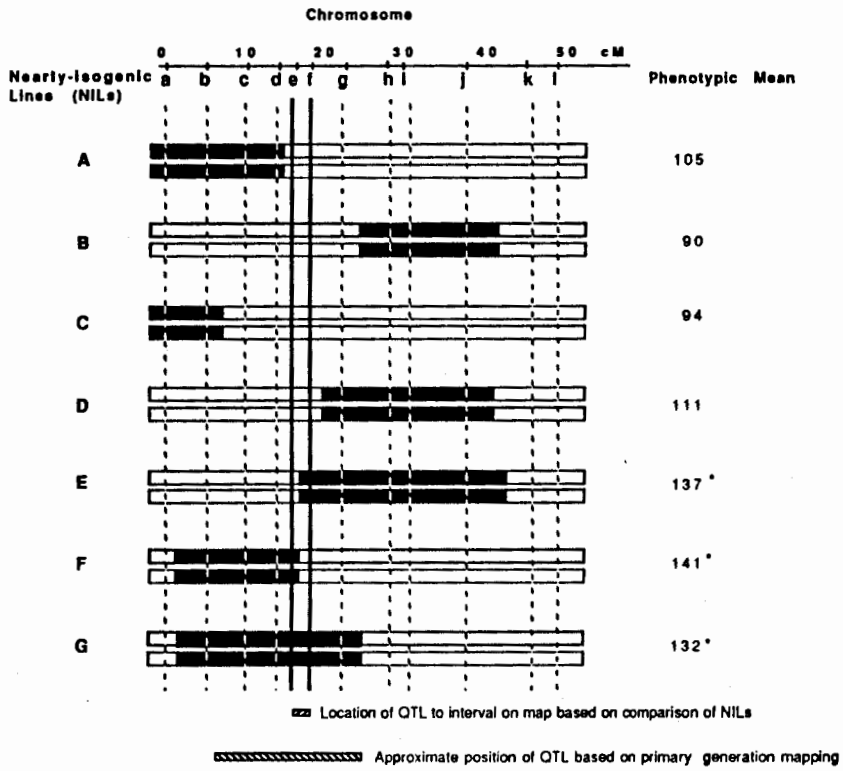


Figure 8 High-resolution mapping of a QTL with nearly isogenic lines (NILs). In primary generations (e.g. F2, BC1, RI), the map position of individual QTL is only roughly approximated (shown by hatched bar at bottom of figure). Once the approximate position of a QTL is known, NILs with crossovers in the vicinity of the QTL can be constructed using linked marker loci (a,b,c, etc). In this theoretical example, seven NILs have been created by backcrossing to the parent depicted by the white chromosomal segments. Dark chromosomal segments come from the other parent. By comparing the phenotypic values of the NILs, the QTL can be mapped more accurately. In this instance, E,F,G have phenotypic values significantly (indicated by asterisks) greater than other NILs indicating that the QTL resides in the interval bounded by marker e and f. If additional markers were available in the e,f interval, a second generation of NILs could be created with crossovers in the interval for even finer mapping of QTL.

necessary for studying QTL composition, can be a time-consuming and labor-intensive task. To have a 95% chance of recovering at least one crossover in a 0.1 cM interval would require approximately 3000 individuals. For some species, it will be difficult to generate such a large number of segregating progeny and even in the species where sufficient numbers can be generated, it is potentially a time-consuming and expensive task to screen

through these numbers with molecular probes. New sampling procedures for selectively isolating recombinants in specific regions of the genome (10), combined with advances in molecular marker technology, are likely to overcome this problem and in the near future, high-resolution mapping of QTL should be realized.

Relationship between Genes Controlling Polygenic and Monogenic Traits

A long-standing question in quantitative genetics is whether the loci controlling qualitative variation (e.g. those with macromutant alleles) are the same as those controlling quantitative variation (e.g QTL). It is apparent from previous work that the alleles for quantitative variation assume a wider range of allelic interactions and have a smaller individual effect than alleles for qualitative variation. However, as pointed out by Robertson (47), it is possible that alleles for qualitative mutants are simply loss-of-function alleles at the same loci underlying quantitative variation. Theoretically, QTL mapping studies can provide a test of this hypothesis. If a gene contributing to quantitative variation is allelic to a gene controlling qualitative variation, then these genes should map to the same locus along the chromosome. For some organisms (e.g. maize, Drosophila), many of the major qualitative loci controlling morphological variation have been mapped with a high degree of precision on genetic maps and these locations should be predictive of the locations of polygenes mapped for the same character.

This hypothesis has been tested in maize. Plant height is a quantitative trait with known qualitative mutants—many of which have been mapped in maize. Beavis and colleagues (5) attempted to test the relationship of qualitative mutants to quantitative variation by mapping QTL for plant height in four maize F2 populations and comparing the map position of those QTL with previously known positions of qualitative variations for the same character. The results showed a general concordance in map positions of QTL and major genes affecting height and is therefore consistent with the hypothesis. Unfortunately, the QTL were located on the map with a low degree of resolution, raising the possibility that the QTL are linked, but not identical to the qualitative loci—a problem acknowledged by the authors. Until QTL are mapped to higher degrees of precision and/or cloned, it will be difficult to prove that the particular QTL actually correspond to known loci defined by macromutant alleles.

Orthologous Polygenes

The advent of technologies for mapping genomes directly at the DNA level has opened the door for comparative genetic mapping among sexually incompatible species. Using a common set of RFLP probes, comparative

linkage maps have now been constructed for tomato-potato-pepper (6, 54, 56), maize-sorghum-rice-wheat (1, 2, 25, 66), and humans-cattle-mice (15, 69). Comparative linkage maps provide a basis for interpreting genetic information among divergent species. For species connected by comparative genetic maps, it should be possible to compare the map positions of QTL for the same or similar characters. Coincidence of map positions would support the hypothesis that loci underlying natural quantitative variation have been conserved during long periods of evolutionary divergence (i.e. they are orthologous genes). In this case, breeders might be able to predict the positions of important QTL (e.g. for growth rates in animals or yield in plants) in one species based on mapping studies from different species.

At present, too few QTL mapping studies have been published for species connected by comparative linkage maps to draw any general conclusions regarding the hypothesis of conserved QTL among divergent species. Perhaps the best evidence for orthologous QTL comes from mung bean and cowpeas (22). In this study the researchers showed that the single most significant QTL for determining seed weight in these two distinct species maps to the same chromosomal locus in both genomes and that the chance occurrence of such coincidental mapping is very unlikely.

Currently, QTL mapping studies are being conducted in rats and mice in the expectation that loci related to congenital diseases are conserved in mammals and that comparative linkage maps will provide a basis for extrapolating the results from studies with rats and mice to humans (14, 27).

Heterosis

Outcrossing (reproduction by mating of two different individuals versus self-fertilization of a single individual) is the commonest mode of reproduction in nature and is enforced by a number of mechanisms, including sex chromosomes in animals (and in some plants) and genetic self-incompatibility in plants. Outcrossing promotes heterozygosity and is often associated with heterosis or hybrid vigor, whereas self-fertilization promotes homozygosity and is associated with inbreeding depression in many organisms.

The genetic basis of heterosis has been debated for many years and is still not resolved. Most geneticists agree that heterosis is at least in part due to dominance (masking of deleterious recessive alleles in heterozygous individuals). More controversial is the role of overdominance in heterosis. Overdominance refers to the situation where individuals heterozygous at a particular locus are superior compared with individuals homozygous for either allele—implying a synergistic interaction between the gene products encoded by the two alleles. In the past, individual loci responsible for hybrid vigor were hard to identify and characterize, making it difficult to resolve the issue of dominance versus overdominance. Recent QTL mapping studies have mapped some of the genes controlling heterosis and are shedding some light on this issue.

The most comprehensive study mapping loci for heterosis was reported by Stuber et al in maize (53). Progeny families derived from a highly heterotic F1 hybrid were evaluated for grain yield and a variety of other traits. The most notable result was that the majority of QTL detected for yield demonstrated overdominant gene action, which would support overdominance as an important factor in heterosis in this species. The limitation to this and virtually all studies involving the mapping of QTL in primary generations is that, since the QTL could not be mapped to exact points in the chromosomes, the QTL demonstrating overdominant gene action may actually be composites of two or more loci with dominant and recessive alleles in coupling which would mimic overdominance—a phenomenon sometimes referred to as pseudo-overdominance (13). High-resolution mapping of individual heterotic QTL should help resolve the issue of overdominance versus pseudo-overdominance since recombination between dominant and recessive loci would be expected with the latter. Ultimately it should be possible to clone heterotic QTL using map-based cloning techniques, and thus allow the molecular basis of this important phenomenon to be determined.

Transgressive Variation

Transgression is defined genetically as the appearance of individuals in segregating populations that fall beyond the parental phenotypes (usually with respect to quantitatively inherited characters) and are often observed in offspring of both intraspecific and interspecific matings. There are several potential causes of transgression including de novo mutation and unmasking of recessive deleterious alleles due to inbreeding (45). However, the cause most often proposed for transgression is accumulation in certain progeny of complementary alleles at multiple loci inherited from the two parents (45).

QTL mapping studies have provided direct evidence for the basis of transgression. Most QTL mapping studies reported thus far have involved analysis of the segregating progeny derived from crossing two individuals significantly different with respect to one or more quantitative traits. One would normally expect the parent with the higher value for a quantitative trait to possess a higher proportion of positive alleles with regard to that character. QTL mapping studies have largely borne out that expectation. At the same time, most QTL mapping studies also report the detection of QTL with allelic effects opposite to those predicted by the phenotype of the parent from which they originated (16, 18, 43). Such "complementary

QTL" would be expected in cases where transgressive segregation is observed in the progeny. What is somewhat surprising is the relatively high frequency with which complementary QTL have been found to occur—especially in interspecific crosses. A good example can be found in tomato.

An F2 population, derived from a cross between two different tomato species, was evaluated for 11 quantitative traits. Transgressive segregation was observed for eight of these traits. Mapping using RFLP markers uncovered 74 QTL that accounted for a large proportion of the genetic variance for the 11 characters, and 36% of those QTL had allelic effects opposite to those predicted by the parental phenotypes (16). In some instances, over half of the QTL detected for any given character were of a complementary nature. These complementary QTL were directly related to the appearance of transgressive individuals in the F2. The finding that different species contain such a high proportion of complementary QTL supports hypotheses that interspecific hybridization can rapidly lead to new races or species with characteristics or adaptations that exceed those of the parental species, due to recombination of a relatively few complementary QTL (16). These results also suggest a strategy for more efficient use of wild plant germplasm to improve domesticated crops (16).

PROGRESS AND PROSPECT

Our ability to map and characterize genes underlying quantitativcly inherited traits has advanced considerably in the past 10 years due to the use of molecular marker techniques. We have learned that, while there are many genes influencing the phenotype of quantitative characters in segregating populations, these genes vary tremendously in both the magnitude and nature of their effects. Experimental results have also shown that single "major" QTL often account for 10–50% of the phenotypic variation in segregating populations. Further mapping studies are likely to reveal whether the same QTL reported in one population also account for the majority of variation in other, independent, populations. Ultimately, these results will be expanded to include the identification of orthologous QTL among different species. Conservation of QTL among species may provide opportunities for plant and animal breeders to use QTL mapping information from one species in the design and execution of breeding studies in another.

Finally, the ability to conduct high-resolution mapping experiments with QTL should help determine whether QTL are single genes or clusters of tightly linked genes and whether overdominance plays a significant role in conditioning heterosis. High-resolution mapping also sets the stage for map-based cloning of genes underlying quantitative variation—an event that will usher in the age of molecular quantitative genetics.

ACKNOWLEDGMENTS

I thank Drs. G Churchill, D Matthews, S McCouch, A Paterson, N Young, D Zamir for helpful comments, and S David for assistance in preparing the manuscript. Research in my laboratory was supported by grants the National Research Initiative Competitive Grants Program, USDA no. 91-37300-6418, the Binational Agricultural Research and Development Fund (no. IS-1822-90C) and The Rockefeller Foundation.

Literature Cited

1. Ahn S, Anderson JA, Sorrells ME, Tanksley SD. 1993. Homoeologous relationships of rice and wheat chromosomes. *Theor. Appl. Genet.* In press
2. Ahn S, Tanksley SD. 1993. Comparative linkage maps of the rice and maize genomes. *Proc. Natl. Acad. Sci. USA*. In press
3. Anderson JA, Sorrells ME, Tanksley SD. 1993. RFLP analysis of genomic regions associated with resistance to pre-harvest sprouting in wheat. *Genomics*. In press
4. Bateson W. 1907. The progress of genetics since the rediscovery of Mendel's paper. *Progr. Rei Bot.* 1:368
5. Beavis WD, Grant D, Albertsen M, Fincher R. 1991. Quantitative trait loci for plant height in four maize populations and their associations with qualitative genetic loci. *Theor. Appl. Genet.* 83:141–45
6. Bonierbale MW, Plaisted RL, Tanksley SD. 1988. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120:1095–103
7. Botstein D, White RL, Skolnick M, Davis RW. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32:314–31
8. Breese EL, Mather K. 1957. The organization of polygenic activity within a chromosome in Drosophila. I. Hair characters. *Heredity* 11:373–95
9. Burr B, Burr FA. 1991. Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. *Trends Genet.* 7:55–60
10. Churchill GA, Giovannoni JJ, Tanksley SD. 1993. Pooled-sampling makes high-resolution mapping practical with DNA markers. *Proc. Natl. Acad. Sci. USA* 90:16–20
11. Coe EH, Hoisington DS, Neuffer MG. 1990. Linkage map of corn (maize) (*Zea mays* L.) (2N=20). In *Genetic Maps*, ed. SJ O'Brien, pp. 6.39–6.67. New York: Cold Spring Harbor Lab. Press
12. Copeland NG, Jenkins NA. 1991. Development and applications of a molecular genetic linkage map of the mouse genome. *Trends Genet.* 7:113–18
13. Crow JF. 1952. Dominance and overdominance. In *Heterosis*, ed. JW Gowen, pp. 282–97. Iowa State Coll. Press
14. Darling SM, Abbott CM. 1992. Mouse models of human single gene disorders I: non-transgenic mice. *BioEssays* 14: 359–66
15. Davisson MT, Lalley PA, Peters J, Doolittle DP, Hillyard AL, Searle AG. 1991. Report of the comparative committee for human, mouse and other rodents (HGM11). *Cytogenet Cell Genet.* 58:1152–89
16. DeVicente MC, Tanksley SD. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics*. 134:585–96
17. Dietrich W, Katz H, Lincoln SE, Shin HS, Friedman J, et al. 1992. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* 131: 423–47
18. Doebley J, Stec A. 1991. Genetic analysis of the morphological differences between maize and teosinte. *Genetics* 129:285–95
19. Edwards MD, Helentjaris T, Wright S, Stuber CW. 1992. Molecular-marker-facilitated investigations of quantitative trait loci in maize. *Theor. Appl. Genet.* 83:765–74
20. Edwards MD, Stuber CW, Wendel JF. 1987. Molecular-marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics* 116:113–25

21. Falconer R. 1960. *Introduction to Quantitative Genetics*. New York: Ronald Press
22. Fatokun CA, Menancio-Hautea DI, Danesh D, Young ND. 1992. Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. *Genetics* 132:841–46
23. Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD. 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Res.* 23:6553–58
24. Haldane JBS. 1946. The interaction of nature and nurture. *Ann. Eugen.* 13: 197–205
25. Hulbert SH, Richer TE, Axtell JD, Bennetzen JL. 1990. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proc. Natl. Acad. Sci. USA* 87:4251–55
26. Hunter RL, Markert CL. 1957. Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science* 125:1294–95
27. Jacob HJ, Lindpaintner K, Lincoln SE, Kusumi K, Bunker RK, et al. 1991. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. *Cell* 67: 213–24
28. Lande R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* 99:541–53
29. Lander ES, Botstein D. 1986. Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map. *Cold Spring Harbor Symp. Quant. Biol.* 51:49–62
30. Lander ES, Botstein D. 1986. Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. *Proc. Natl. Acad. Sci. USA* 83:7353–57
31. Lander ES, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–99
32. Law CN. 1967. The location of factors controlling a number of quantitative characters in wheat. *Genetics* 56:445–61
33. Lewontin RC, Hubby JL. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* 54:595–609
34. Martin B, Nienhuis J, King G, Schaefer A. 1989. Restriction fragment length polymorphisms associated with water use efficiency in tomato. *Science* 243: 1725–28
35. Martin GB, Williams JGK, Tanksley SD. 1991. Rapid identification of markers linked to a Pseudomonas resistance gene in tomato by using random primers and near-isogenic lines. *Proc. Natl. Acad Sci. USA* 88:2336–40
36. Mather K, Jinks JL. 1971. *Biometrical Genetics*. Ithaca, NY: Cornell Univ. Press
37. Michelmore RW, Paran I, Kesselli RV. 1991. Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proc. Natl. Acad. Sci. USA* 88:9828–32
38. Moll RH, Cockerham CC, Stuber CW, Williams WP. 1978. Selection responses, genetic-environmental interactions, and heterosis with recurrent parent selection for yield in maize. *Crop Sci.* 18:641–45
39. Nodari RO, Tsai SM, Guzman P, Gilbertson RL, Gepts P. 1993. Toward and integrated linkage map of common bean. III. mapping genetic factors controlling host-bacteria interactions. *Genetics* 134:341–50
40. Park YC. 1997. Theory for the number of genes affecting quantitative characters 2. Biases from drift, inequality of gene effects, linkage disequilibrium and epistasis. *Theor. Appl. Genet.* 50:163–72
41. Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, et al. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. *Genetics* 127:181–97
42. Paterson AH, Deverna JW, Lanini B, Tanksley SD. 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecies cross of tomato. *Genetics* 124:735–42
43. Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD. 1988. Resolution of quantitative traits into Mendelian factors using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335: 721–26
44. Reiter RS, Williams JG, Feldman KA, Rafalsk JA, Tingey SV, Scolnik PA. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. USA* 89:1477–81
45. Rick CM, Smith PG. 1953. Novel variation in tomato species hybrids. *Am. Nat.* 88:359–73
46. Rise ML, Frankel WN, Coffin JM, Seyfried TN. 1991. Genes for epilepsy mapped in the mouse. *Science* 253: 669–73
47. Robertson DS. 1985. A possible technique for isolating genomic DNA for quantitative traits in plants. *J. Theor. Biol.* 117:1–10
48. Sax K. 1923. The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8:552–60
49. Serikawa T, Kuramoto T, Hilbert P, Mori M, Yamada J, et al. 1992. Rat gene mapping using PCR-analyzed microsatellites. *Genetics* 131:701–21
50. Smith HH. 1937. The relation between genes affecting size and color in certain species of Nicotiana. *Genetics* 22:361
51. Smithies O. 1955. Zone electrophoresis in starch gels. *Biochem. J.* 61:629
52. Stuber CW, Edwards MD, Wendel JF. 1987. Molecular-marker facilitated investigations of quantitative trait loci in maize II. Factors influencing yield and its component traits. *Crop Sci.* 27:639–48
53. Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132:823–39
54. Tanksley SD, Bernatzky R, Lapitan NL, Prince JP. 1988. Conservation of gene repertoire but not gene order in papper and tomato. *Proc. Natl. Acad. Sci. USA* 85:6419–23
55. Tanksley SD, Medina-Filho H, Rick CM. 1982. Use of naturally-occurring enzyme variation to detect and map gene controlling quantitative traits in an interspecific backcross of tomato. *Heredity* 49:11–25
56. Tanksley SD, Ganal MW, Prince JP, DVicente MC, Bonierbale MW, et al. 1992. High density molecular linkage maps of the tomato and potato genomes *Genetics* 132:1141–60
57. Taylor BA. 1976. Genetic analysis of susceptibility to isoniazid-induced seizures in mice. *Genetics* 83:373–77
58. Thoday JM. 1961. Location of polygenes. *Nature* 191:368–70
59. Thompson JN Jr. 1975. Quantitative variation and gene number. *Nature* 258:665–68
60. Thompson JN Jr, Thoday JM, eds. 1979. *Quantitative Genetic Variation*. New York: Academic
61. Vallejos CE, Sakiyama NS, Chase CD. 1992. A molecular marker-based linkage map of *Phaseolus vulgaris* L. *Genetics* 131:733–40
62. Vallejos CE, Tanksley SD. 1983. Segregation of isozyme markers and cold tolerance in an interspecific backcross of tomato. *Theor. Appl. Genet.* 66:241–47
63. Weber JL, May PE. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* 44:388–96
64. Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, et al. 1992. A second-generation linkage map of the human genome. *Nature* 359:794–801
65. Weller JI, Soller M, Brody T. 1988. Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* x *Lycopersicon pimpinellifolium*) by means of genetic markers. *Genetics* 118:329–39
66. Whitkus R, Doebley J, Lee M. 1992. Comparative genome mapping of sorghum and maize. *Genetics* 132:1119–30
67. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. Oligonucleotide primers of arbitrary sequence amplify DNA polymorphisms which are useful as genetic markers. *Nucleic Acids Res.* 18:6531–35
68. Winkelman DC, Hodgetts RB. 1992. RFLPs for somatotropic genes identify quantitative trait loci for growth in mice. *Genetics* 131:929
69. Womack JE. 1990. In *Mapping the Genomes of Agriculturally Important Animals*, ed. JE Womack, pp. 19–24. New York: Cold Spring Harbor Lab. Press

Plant Breeding 104, 1—19 (1990)

ISSN 0179-9541

Review

Institut für Pflanzenzüchtung, Saatgutforschung und Populationsgenetik, Universität Hohenheim (F. R. Germany)

Use of Molecular Markers in Breeding for Oligogenic Disease Resistance

A. E. Melchinger

With 4 figures and 2 tables

Received August 8, 1989 / Accepted November 15, 1989

Abstract

In several plant species, genetic maps have recently been developed for restriction fragment length polymorphisms (RFLPs). Together with isozymes, they offer alternative solutions to many breeding problems. This review deals with the current and potential uses of molecular markers in breeding for oligogenic resistance traits.

In the first part, segregation analyses and analyses of near-isogenic lines are collated with respect to the mapping of resistance genes. Also, various types of populations are compared for the amount of information obtained in segregation analyses. In the second part, theoretical and numerical results are presented dealing with the number of individuals required for marker-facilitated selection of a resistance gene in a backcross program. Both the use of a single marker and a marker bracket are considered and the influence of the following parameters is investigated: (a) recombination frequency between the resistance gene and marker(s), (b) size of backcross families, (c) number of backcross generations, and (d) number of carriers of the resistance gene to be recovered in the final backcross generation. The results provide information with regard to the optimum design of marker-facilitated selection programs and the required expenditures compared to direct selection of the resistance trait. In the third part, applications of molecular markers are discussed with respect to (1) advantages of marker-facilitated selection for resistance genes under a variety of special circumstances, (2) pyramiding of resistance genes, (3) selection against the genetic background of a donor parent, and (4) their use as a starting point for chromosome walking.

Key words: Restriction fragment length polymorphisms (RFLPs) — isozymes — genetic mapping — marker-facilitated selection — resistance breeding

1. Introduction

Advances in molecular biology during the last decade have provided a new class of genetic markers at the level of DNA, termed restriction fragment length polymorphisms (RFLPs). Investigations in maize (Helentjaris et al. 1986 b, Burr et al. 1988), rice (McCouch et al. 1988), soybeans (Apuya et al. 1988), tomato (Bernatzky and Tanksley 1986), potato (Bonierbale et al. 1988), and brassicas (Figdore et al. 1988) have demonstrated that a potentially unlimited number of RFLPs exist, which should enable plant geneticists to establish well saturated genetic maps for any species. These developments have stimulated new interest in exploring the applications of genetic markers in plant breeding.

The potential usefulness of genetic markers as an instrument for the plant breeder was recognized more than 60 years ago (Sax 1923). Until the past 20 years, however, its applica-

tion was largely hindered by the lack of suitable markers. Detailed genetic linkage maps were available only in a few crop species (KING 1975), and the majority of markers were recessive morphological mutations that often had detrimental side effects. The development of molecular markers (isozymes and, more recently, RFLPs) promises to overcome most of the previous limitations (TANKSLEY 1983) and thus could open a wide spectrum of alternative approaches to many breeding problems.

Molecular markers have several advantages over morphological markers. (a) Numerous markers can be identified in breeding materials. (b) A relatively large number of alleles can be found. (c) Most molecular markers show a codominant mode of inheritance. (d) Molecular markers are generally silent in their effect on the phenotype. (e) Genotypes of most molecular markers can be determined at a very early developmental stage, allowing early screening methods to be applied.

Several reviews are available that deal in a general manner with the potential utilization of molecular markers in basic plant genetic studies and applied plant breeding programs (SOLLER and BECKMANN 1983, TANKSLEY 1983, BECKMANN and SOLLER 1986a, b, LANDRY and MICHELMORE 1987). This review summarizes the various approaches and types of populations proposed for developing genetic linkage maps and describes the actual and potential uses of molecular markers in breeding for oligogenic resistance traits. In addition, theoretical and numerical results are presented dealing with the optimum design and the number of individuals to be used for marker-facilitated selection of monogenic resistances in a backcross program.

2. Molecular markers

2.1 Isozymes

During the past two decades, techniques have been developed for the analysis of allelic variants of specific enzymes (isozymes) in plant tissues (TANKSLEY and ORTON 1983). Proteins in crude extracts are separated by electrophoresis on starch or polyacrylamide gels. By staining the gel for a specific enzyme activity, it is possible to visualize the corresponding isozymes. Changes in the amino acid composition may alter the electrophoretic mobility and thus result in a different banding pattern. Different enzyme systems often require specific laboratory procedures (STUBER et al. 1988). Furthermore, some enzymes can be assayed only at a certain developmental stage or in specific tissues.

Once established, isozyme methods are fairly simple and can be employed on a large scale. Isozyme systems have been described for more than 30 crop species (TANKSLEY and ORTON 1983). Although in some species (e.g., *Zea mays* L., *Secale cereale* L.) about 40 or more isozyme systems are known (STUBER et al. 1988, WEHLING et al. 1985), only between 10 and 20 polymorphic isozyme loci are commonly found in most breeding populations (STUBER and GOODMANN 1983).

2.2 Restriction fragment length polymorphisms (RFLPs)

RFLPs reflect differences in homologous DNA sequences that alter the length of restriction fragments obtained by digestion with type II restriction enzymes. These differences result from base pair changes or other rearrangements (e.g., translocations and inversions) at the recognition site of the restriction enzyme or from internal deletion/insertion events. The restriction fragments are separated according to their size by agarose gel electrophoresis. Subsequently, a Southern blot analysis is performed: the DNA is transferred to and immobilized on a nylon membrane or a nitrocellulose filter, hybridized to a labelled DNA probe, and visualized by autoradiography or a color reaction (SOUTHERN 1975). In general, unique or low copy number DNA sequences are used as probes and labelled with either ^{32}P or biotin (WINNACKER 1987). Such probes can be cloned DNA of specific genes, cDNA, random genomic DNA, or specifically synthesized oligonucleotides (BECKMANN and SOLLER 1986a). A survey of the different types and sources of probes as well as the different techniques to detect RFLPs was presented by LANDRY and MICHELMORE (1987).

Since RFLPs directly reflect differences in the DNA sequence, they are independent of gene expression and can be determined at any stage of development and in all tissues, with few exceptions. Moreover, the laboratory methods for detecting RFLPs are generally the same, irrespective of the species or marker investigated. One of the foremost attributes of RFLPs compared to isozymes is the substantially greater number of polymorphic markers found within breeding materials (BECKMANN and SOLLER 1986a). The reasons are: (i) minor differences in the DNA can yield RFLPs; and (ii) not only translated DNA regions (exons) but also nontranslated DNA regions contribute to the variation detected. However, RFLP assays require expensive laboratory supplies and are rather time consuming compared to isozymes analyses.

3. Mapping of resistance genes

Most marker applications in breeding for disease resistance require a precise linkage map for both the markers and the target loci. For oligogenic traits, whose mode of inheritance can be determined by classical Mendelian factor analysis, genetic maps can be established by segregation analysis and/or comparison of near-isogenic lines. For a few plant species, mapping can also be facilitated by the use of aneuploids (HELENTJARIS et al. 1986a), chromosome substitution lines, or translocation stocks (WEBER and HELENTJARIS 1989). However, the latter methods are primarily applied to confirm results obtained by segregation analysis, to assign linkage groups to chromosomes and chromosome arms, and to correlate RFLP maps more closely with conventional genetic maps derived from cytological, morphological, and isozyme data.

3.1 Segregation analysis

Genetic maps of molecular markers in plants have conventionally been constructed by segregation analysis of generations derived from the F_1 cross between homozygous parents differing at appropriate loci. In most instances, F_2 or first backcross generations have been used, with individual plants being the unit of segregation (BERNATZKY and TANKSLEY 1986, HELENTJARIS et al. 1986b, LANDRY et al. 1987, MURRAY et al. 1988). Recently, random inbred lines developed by single seed descent (herein designated as F_∞ lines(S) but in literature often referred to as recombinant inbred lines) have been employed for estab-

Table 1. Summary of maximum likelihood equations for estimating the recombination frequency p between two loci from various types of populations derived from the cross of two homozygous lines and formulas for calculating the average amount of information, $I(p)$, contributed per segregation unit

Type of population[1]	Estimation equation[2]	Information $I(p)$[3]	Reference[4]
F_2	$\hat{p}_1$: Solve eqn. $\frac{2x}{p} + \frac{2y}{p-1} + (n - x - y - z)\left[\frac{1-2p}{p(1-p)}\right] + z\left[\frac{2(2p-1)}{1-2p+2p^2}\right] = 0$	$I_1(p) = \frac{2(1-3p+3p^2)}{p(1-p)(1-2p+2p^2)}$	1
F_2 Syn1	$\hat{p}_2 = (3/4) - [(9/16) - \hat{p}_1]^{1/2}$	$I_2(p) = I_1(p^*) \times (9 - 16p^*)/4$	4
DH lines; BC1	$\hat{p}_3 = y/n$; $\hat{p}_4 = w/n$	$I_3(p) = I_4(p) = 1/p(1-p)$	1
F_∞ lines (S)	$\hat{p}_5 = y/2x$	$I_5(p) = 2/p(1+2p)^2$	2, 4
F_∞ lines (FS)	$\hat{p}_6 = y/(4x - 2y)$	$I_6(p) = 4/p(1+2p)(1+6p)^2$	3

[1] For detailed description of terminology, see text.
[2] x: = [AB/AB, ab/ab], y: = [Ab/Ab, aB/aB], z: = [AB/ab, Ab/aB] and w: = [Ab/ab, aB/ab] denote the observed frequencies of genotypes AB/AB, ab/ab etc., assuming that the parental genotypes were AB/AB and ab/ab and the latter served as backcross parent; n refers to the total number of segregation units assayed.
[3] $p^* = p(1.5 - p)$.
[4] 1 = ALLARD (1956); 2 = SNAPE (1988); 3 = SILVER (1985); 4 = own derivations.

lishing a RFLP linkage map of maize (BURR et al. 1988). Moreover, random inbred lines derived by doubled haploid (DH) methods have been proposed as a further alternative (SNAPE 1988).

Published results for estimating recombination frequency p by the method of maximum likelihood (ML) with these types of populations are summarized in Table 1. In addition to the above-mentioned types of populations, random inbred lines developed by full-sib mating (designated as F_∞ lines(FS)) and the Syn1 generation produced by random mating an F_2 population (designated as F_2 Syn1) are considered. The underlying assumptions are: (a) segregation ratios at each locus are consistent with those expected for codominant genes under Mendelian inheritance, and (b) recombination frequencies are the same in both sexes. Implicit solutions of the ML equations are available except for the F_2 and F_2 Sny1 populations. Deviations from assumption (a) can be caused by irregularities during meiosis or gametic and zygotic selection, as reported for isozyme markers linked to incompatibility loci in rye (WRICKE and WEHLING 1985, GERTZ and WRICKE 1989). ML estimates of recombination frequencies in F_2 populations with distorted single gene segregation were derived for various modes of inheritance by HEUN and GREGORIUS (1987) and GERTZ (1989).

Table 1 also provides formulas for the average amount of information, $I(p)$, for estimating p, contributed by a single segregation unit. $I(p)$ depends on the underlying value of p and the type of population employed. The variance of the ML estimate of $\hat{p}$ is obtained as

$$\sigma^2_{\hat{p}} = \frac{1}{n\, I(p)},$$

where n is the number of segregation units assayed. For all types of populations, $\sigma^2_{\hat{p}}$ is largest for independently segregating loci ($p = 0.5$) and converges toward zero for $p \to 0$.

Following ALLARD (1956), the F_2 completely classified into the 10 possible genotypes is used as a standard for comparison of the different types of populations. Its average amount of information is $I^*(p) = 1/2p(1-p)$. Fig. 1 shows graphs of the relative information, i.e., $I(p)$ expressed in percent of $I^*(p)$, for all six types of progeny. DH lines and BC populations provide exactly half the information as the F_2 completely classified, because they are informative only for the recombination events in a single gamete. The relative information of the F_2 itself (i.e., without separate classification of the two types of double heterozygotes) declines almost linearly from 100 % for $p = 0$ to 50 % for $p = 0.5$. F_∞ lines(S) provide nearly the same amount of information as F_2 popula-

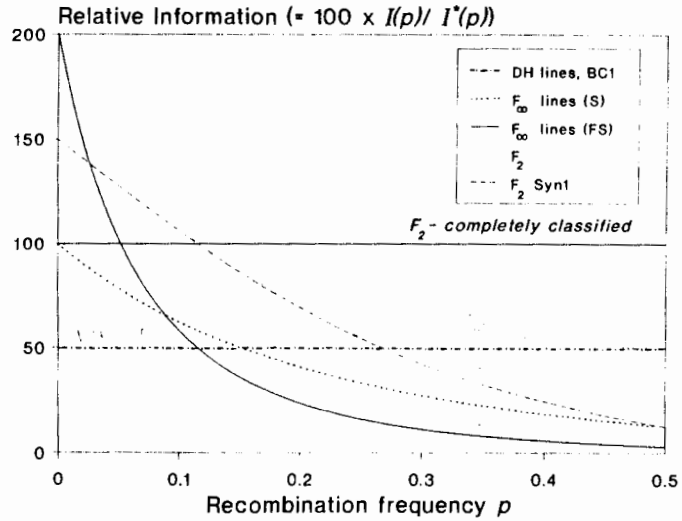


Fig. 1. Average amount of information $I(p)$ contributed per segregation unit (expressed in per cent of the F_2 completely classified) as a function of the recombination frequency p for various types of populations

tions for tightly linked loci. However, for $p > 0.15$ they provide even less information than DH lines. F_∞ lines(FS) supply about twice as much information as F_2 populations for extremely tightly linked loci ($p < 0.05$). Yet for larger values of p, they become rapidly less efficient than all other types of populations. Likewise, the F_2 Syn1 provides approximately 50 % more information than the F_2 for very tightly linked loci, but has lower efficiency than the F_2 for $p > 0.17$. In conclusion, additional opportunities for recombination due to intermating, sib mating or selfing result in a gain of information only for small values of p, but reduce the efficiency of estimating p for intermediate or large recombination values.

Besides constraints imposed by the reproduction biology of the species considered, choice of the optimum type of population employed in mapping studies depends primarily on the objectives pursued. In a species with a well saturated linkage map, mapping of a resistance gene relative to closely linked markers can be accomplished most efficiently by analyzing F_2 or F_2 Syn1 populations. Although F_∞ lines(FS) could increase the efficiency even further for extremely tightly linked loci, their use is usually prohibited by the large number of generations required for obtaining a sufficient level of inbreeding.

For the initial construction of a marker linkage map, F_2 populations are likewise superior to F_∞ lines(S) or DH lines from a biometrical point of view; i.e., for given sample size n, the ML estimate of p for the F_2 has uniformly the smallest variance in the entire interval (0, 0.5). In practical use, however, random inbred lines have two advantages compared to segregating populations (BURR et al. 1988): (a) They represent a permanent population because all alleles are fixed and, thus, can be used indefinitely for mapping of new DNA probes and passed on to other research groups. (b) They can be tested in replicated trials as might be required for determining resistance traits with incomplete penetrance. On the other hand, it should be pointed out that instead of sampling tissue from individual F_2 or F_2 Syn1 plants, their RFLP genotype can be determined with sufficient accuracy by analyzing tissue mixtures sampled from 10 to 20 plants of the respective F_2-derived lines obtained by selfing.

Segregation analysis and construction of linkage map are stepwise processes, which can be facilitated with the aid of computer programs (SUITER et al. 1983, HOISINGTON 1987). Goodness-of-fit to expected segregation ratios at each locus are checked by χ^2 analyses. Subsequently, two-way χ^2 contingency tests are employed to test for independent segregation of loci pairs. If significance is obtained, the recombination frequency p and its standard error is calculated by using the ML formulas. Finally, the linear order of loci in linkage groups is inferred from three-point linkage analyses (STRICKBERGER 1985).

Recently, considerable progress has been made in the computer-facilitated construction of genetic linkage maps based on multipoint linkage analyses. LANDER and GREEN (1987) described a new algorithm for computing the likelihood of a given genetic map, which is defined by the order of loci and the recombination values between them. Thus, it is possible to search for the "best" map, i.e., the map with the highest probability for the observed data. Employing this algorithm, LANDER et al. (1987) developed a computer package, MAPMAKER, that allows rapid construction of linkage maps from RFLP data of F_2 or backcross populations by simultaneous multipoint analysis of any number of loci. In line with the traditional procedure employed in human genetics, the program uses the maximum LOD scores (MORTON 1955) for a test of two-point linkage. (The maximum LOD score for a pair of loci is defined as the $\log_{10}$ of the ratio of the likelihoods obtained for the ML estimate p and for $p = 0.5$).

There are only a few reports in the literature in which molecular markers have actually been used for mapping resistance genes. LANDRY et al. (1987) have established a linkage map of lettuce (*Lactuca sativa*) that includes five downy mildew resistance genes. Furthermore, an isozyme linkage map of pea (*Pisum sativum*) was used to map resistance genes against bean yellow mosaic virus and pea enation mosaic virus (WEEDEN et al. 1984, WEEDEN and PROVVIDENTI 1988).

3.2 Analysis of near-isogenic lines

Although mapping of a single resistance gene on the newly emerging marker linkage maps by traditional segregation analysis is straightforward, it becomes a formidable task for a larger number of resistance genes. Suitable F_2

or backcross populations would have to be developed that segregate for both the resistance gene(s) and a large number of molecular markers covering the whole genome. These populations must then be assayed for the resistance trait(s) and, in the absence of prior information, for the entire set of markers. Consequently, integration of resistance genes and molecular markers into a single genetic map by means of conventional segregation analyses would be very time consuming and tedious.

MUEHLBAUER et al. (1988) and YOUNG et al. (1988) have recently described an alternative approach for rapidly identifying molecular markers that are tightly linked to one of a large number of important plant genes. This procedure involves the use of genetic stocks, commonly referred to as near-isogenic lines (NILs). Sets of NILs are available from conventional breeding programs for numerous resistance genes in the major crop species. NILs endowed with different resistance genes have been bred as components of multiline varieties in oats and wheat (FREY 1982).

The approach takes advantage of the fact that most NILs, in particular those for resistance genes, have been developed by means of the backcross method (cf. FEHR 1987). Accordingly, a donor parent (DP) carrying the gene of interest is repeatedly backcrossed to a recurrent parent (RP), usually a cultivated line chosen because of its otherwise favorable properties. Backcrossing accompanied by selection for the resistance gene and recovery of the recurrent parent phenotype is continued (generally for five to seven generations) until the newly developed resistant line is theoretically nearly isogenic with the RP, except for the chromosome segment containing the target gene. Theoretical calculations show that even after 10 backcross generations, this segment is expected to have a length of about 18 centimorgans (HANSON 1959, STAM and ZEVEN 1981).

The basic idea of the NIL gene mapping technique is to identify molecular markers located in the linkage block surrounding the introgressed gene. Linkage between a molecular marker and the target gene can be assessed *a posteriori* by determining the marker genotype of the RP, its various NIL derivatives, and their corresponding DPs. Putative evidence of linkage is obtained whenever the NIL has the same marker genotype as its DP, but a different marker genotype than its RP.

Successful application of the NIL mapping technique presupposes that the DP-derived chromosome segment flanking the target gene contains molecular markers with different allelic states in the RP and DP. Obviously, the chances of detecting such markers depend on the length of the introgressed chromosome segment and thus decrease with a higher number of backcross generations employed in developing the NIL(s). Moreover, the chances depend on the extent of molecular diversity between the RP and the DP genomes. Polymorphic molecular markers are more likely to be found when the RP and DP belong to cultivated and wild taxa, respectively. On the other hand, fewer polymorphic markers are anticipated if the RP and DP are more closely related.

With DNA markers, the chances of detecting RFLPs within the chromosome segment attached to the gene of interest can be enhanced by screening a large number of probes and employing several restriction enzymes. Probes hybridizing to DNA regions where the NIL is homozygous for the RP genome will display identical restriction fragment patterns in the RP and NIL. By contrast, probes hybridizing to genomic regions inside the introgressed segment may in combination with certain restriction enzymes uncover RFLPs between the RP and the NIL. Screening a large number of probes can be facilitated by simultaneous probing with sets of five to ten clones (YOUNG et al. 1987). Thus, the number of Southern hybridizations can be reduced considerably. The probes can be random genomic clones, preferentially single-copy sequences, but do not have to have been mapped previously.

A possible source of error in the NIL gene mapping technique is that the RP may not only differ for the chromosome segment carrying the target gene but also for other DP-derived DNA sequences scattered throughout the genome. This is due to the fact that with a finite number of backcross generations t, elimination of the DP-derived alleles at loci unlinked to the introgressed gene follows a geometric sequence $(1/2)^t$ and, consequently, is never complete. Obviously, these other DP-derived DNA sequences can harbor "false positive" markers, i.e., RFLPs or isozyme variants that are actually not linked to the target gene. Theoretical calculations by MUEHLBAUER et al. (1988) show that in a hypothetical species with 20 chromosomes of 50 centimorgan map length, about four out of 100 randomly chosen DP-derived molecular markers are expected to be retained in a NIL developed by five backcross generations. Of these four markers, one or two are expected to be not linked to the target gene.

The probability of errors P, i.e., occurrence of "false positive" markers, decreases with a higher number of backcross generations because the length of the introgressed chromosome segment in the NIL is reduced at a lower rate than the proportion of the DP-derived genome that is not linked to the target gene. Another means to increase the power of the approach is to include in the comparison not only a single NIL but rather two (or more) independently derived NILs from the same pair of RP and DP. Since the chances of obtaining the same "false positive" marker in each NIL are independent from each other, the probability of presumptive errors becomes P^2 for two NILs instead of P for a single NIL.

Since only few genotypes (RP and NIL(s)) have to be assayed, the NIL gene mapping technique provides a proper tool for narrowing down a large set of randomly chosen markers to a subset of a few markers, some of which are very tightly linked to the resistance gene introgressed in the NIL. With this subset of molecular markers, however, a traditional segregation analysis would have to be performed in order to confirm or refute the putative linkages and obtain information about the gene order and intervening map distances. NIL gene mapping represents, therefore, a complementary rather than an alternative approach for integrating monogenic resistances and molecular markers into a single genetic map.

Successful application of NILs for mapping purposes has been demonstrated in several recent studies (YOUNG et al. 1988, CHYI et al. 1989, GUPTA et al. 1989). YOUNG et al. were able to identify DNA markers in tomato (*Lycopersicon esculentum*) that were less than 0.5 centimorgans distant from a gene controlling resistance to tobacco mosaic virus. The results from this experiment substantiate that the NIL gene mapping technique should be a very powerful tool for finding RFLP markers extremely tightly linked to a gene of interest. Such markers are desired for marker-facilitated selection of resistance traits and would be an absolute necessity for attempts to clone resistance genes by chromosome walking.

4. Marker-facilitated selection

Tight linkage of a marker to a resistance gene can be exploited for indirect selection of monogenic resistances in a breeding program. Application of this method presupposes that the initial population is polymorphic for the marker and the resistance gene and both are in extreme linkage disequilibrium. Instead of testing for the resistance trait itself, selection in segregating generations is based on determination of the marker genotype. Only those individuals that carry the desired marker allele(s) are selected as parents for the next generation. As proposed by TANKSLEY (1983), use of two linked markers bracketing the target gene should improve the reliability of marker-facilitated selection.

So far, little is known about the optimum design of breeding programs utilizing marker-facilitated selection. Of special interest is the minimum number of individuals to be monitored and retained in each generation, which depends on the degree of linkage and whether a single marker or a marker bracket is employed. Moreover, information about the total number of individuals to be screened is needed as a basis for evaluating under which circumstances marker-facilitated selection is economically superior to direct selection. In the following, these questions shall be investigated for the breeding method most frequently used for introgressing monogenic resistances: recurrent backcrossing.

Fig. 2 shows a flow diagram of the various breeding steps and decisions involved in marker-facilitated selection for a major resistance gene with recurrent backcrossing. The first step after producing the BC1 generation involves marker assays of individual plants. All genotypes heterozygous for the desired marker allele(s) (i.e., single heterozygotes and double heterozygotes in the case of a single marker and a marker bracket, respectively) are retained and backcrossed to the recurrent parent (step 2). For each BC2 family obtained in this

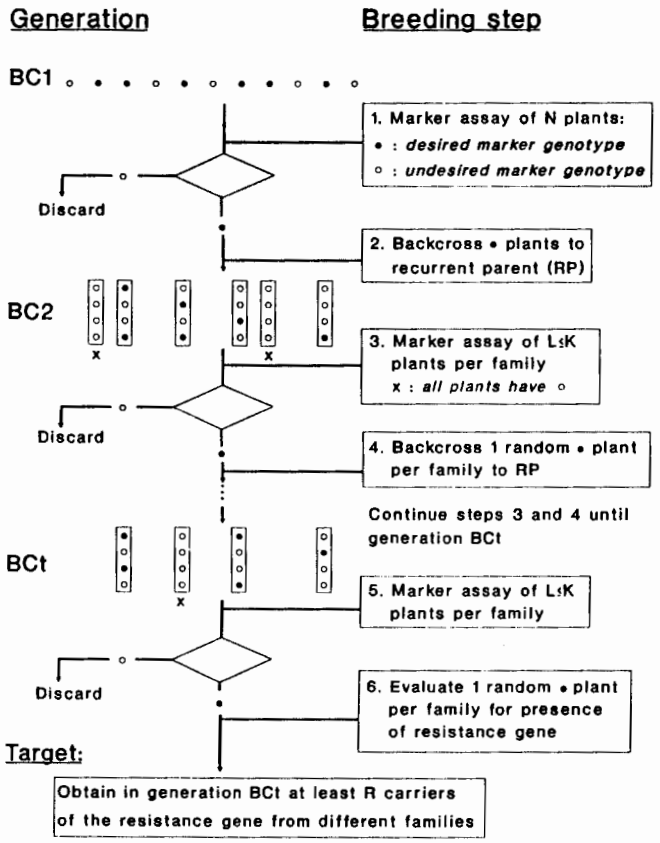


Fig. 2. Diagram of marker-facilitated selection for a monogenic resistance trait in a backcross program. For definitions of N, L, and K, see text

fashion, a variable number of plants, L, is screened for their marker genotype (step 3) according to one of the following two options. Option I involves sequential monitoring of a single plant per family at a time, until one plant of the desired marker genotype is obtained or K plants have been tested without success. Option II involves screening of L = K plants per family simultaneously. If none of these plants has the desired marker genotype, the family is discarded. Otherwise, one plant with the desired marker genotype is used from each BC2 family for further backcrossing (step 4). Steps 3 and 4 can be continued for several generations. In an intermediate of final generation following the marker assays (step 5), one plant of the desired marker genotype per family is evaluated for presence of the resistance gene by direct or progeny testing (step 6). The ultimate goal of the breeding program is to recover in generation BCt at least R carriers of the resistance gene originating from different BC1 individuals. These genotypes might be used as parents for further backcrossing or selfing and, after evaluating their agronomic performance, released as new cultivars or components of cultivars in hybrids, synthetics, or multiline varieties.

4.1 Theoretical considerations

Calculation of the number of plants required in a classical backcross program is generally based on the criterion that the breeders' goal of obtaining R carriers of the resistance gene should be met with a given probability of success α,

e.g., 95 % (SEDCOLE 1977). This criterion will also be adopted subsequently for marker-facilitated selection. Further assumptions are: (a) regular diploid Mendelian inheritance, (b) the recombination frequency between loci pairs is independent of the genotype and constant for all generations, and (c) the parents chosen for the backcross program are homozygous lines.

Consider a random individual in generation BC1. Let

δ denote the probability that it has the desired (heterozygous or double-heterozygous) marker genotype, and let

γ denote the probability that it has the resistance gene as well.

For a single marker locus i,

$$\delta = 1/2 \text{ and } \gamma = (1-p_{ir})/2. \quad (1)$$

For two markers i and j bracketing the resistance locus r,

$$\delta = (1-p_{ij})/2 \text{ and } \gamma = (2-p_{ir}-p_{jr}-p_{ij})/4, \quad (2)$$

where p_{ir}, p_{jr}, and p_{ij} are the recombination frequencies between the loci pair indicated by the indices.

Suppose the breeding plan outlined in Fig. 2 is followed until generation BCt ($t \geq 1$). Let

δ_t denote the probability

(a) to recover from a random BC1 individual a BCt family and

(b) at least one out of K randomly sampled plants from this family has the desired marker genotype. Furthermore, let

γ_t denote the probability that in addition to (a) and (b),

(c) a random plant from a BCt family has not only the desired marker genotype but the resistance gene as well.

The transition probabilities for these events from one backcross generation to the next can be described by Markov chains. Using elementary results about stochastic processes (cf. KARLIN 1966, Chap. 2), it can be shown that for $t \geq 1$

$$\delta_t = \delta\,\xi^{t-1}, \quad (3)$$

and

$$\gamma_t = \gamma^{\,t}\,(\xi/\delta)^{t-1}, \quad (4)$$

where $\xi = 1 - (1-\delta)^K$.

According to the above definition, γ_t describes the probability of recovering from a BC1 individual a carrier of the resistance gene in generation BCt under the breeding plan in Fig. 2. Consequently, the minimum number N of individuals to be marker assayed in generation BC1 so that in generation BCt at least R carriers of the resistance gene are recovered with probability α is obtained by solving the following inequality for N:

$$\sum_{i=R}^{N} \binom{N}{i} (\gamma_t)^i (1-\gamma_t)^{N-i} \geq \alpha. \quad (5)$$

Several methods leading to approximate solutions for this inequality were presented by SEDCOLE (1977). Direct numerical solutions can be obtained by use of statistical computing packages, e.g., the SAS function PROBBNML (SAS 1988).

After determination of the population size N in BC1, the expected number of backcross families, F_l, recovered in a later backcross generation BCl ($1 \leq l \leq t$) is obtained from equation (3) and the mean for a binomial distribution as

$$F_l = N\,\delta_{l-1}. \quad (6)$$

Hence, the expected total number of individuals, M, that must be marker-assayed in generation BCl under option II (simultaneous screening of K individuals per family) is

$$M^{II}_l = K\,N\,\delta_{l-1}. \quad (7)$$

Under option I (sequential screening in each family), the expected number of marker assayed individuals per family is ξ/δ. Thus, the expected total number of individuals to be marker-assayed in generation BCl becomes

$$M^{I}_l = N\,\xi^{l-1}. \quad (8)$$

Summation over l yields the expected total number of individuals, T, to be marker assayed during the entire backcross program from generation BC1 to BCt. Under option I, we obtain

$$T^{I} = N\,(1-\xi^t)/(1-\xi), \quad (9)$$

and under option II,

$$T^{II} = \begin{cases} N & \text{for } t = 1, \\ \text{and} & \\ N\,[1 + K \times & \\ \delta\,(1-\xi^{t-1})/(1-\xi)] & \text{for } t > 1. \end{cases} \quad (10)$$

Besides expenditures for the marker assays, the efficiency of a marker-facilitated selection program depends on the number of individuals in generation BCt that must be screened for presence of the resistance gene either directly or by using their testcross or selfed progenies. From equation (3), we obtain for the expectation of this random variable

$$V = N\,\delta\,\xi^{t-1}. \quad (11)$$

In order to keep the number of parameters as small as possible, two simplifying restric-

tions were imposed in the derivation of equations (5) to (11) that can easily be removed. First, for calculating N according to formula (5), it was assumed that the resistance genotype of BCt individuals (step 6 in Fig. 2) is determined without error. This holds true only if the resistance trait has full penetrance and/or a fairly large family size is employed for monitoring the resistance trait in testcross or selfed progenies. Otherwise, the expression for γ_t in equation (4) should be multiplied with the probability of correct classification of the resistance genotype. Second, instead of testing the resistance genotype of only a single plant from each BCt family, all plants with the desired marker genotype could be tested. In this case, the value of γ_t to be inserted in equation (5) becomes

$$\gamma_t^* = \gamma_{t-1}[1 - (1-\gamma)^K]. \quad (12)$$

Equations (8), (9), and (11) would have to be modified accordingly. For less stringent linkage between the resistance gene and the marker(s), this procedure allows a slight reduction in the total number of marker assays at the expense of testing more individuals for the resistance trait.

Furthermore, it should be pointed out that the above calculations are based exclusively on *a priori* probabilities. For $t > 1$, a reduction in T^I, T^{II}, and especially in V could be attained by choosing the number of individuals retained for the next generation on the basis of the actual outcome in each generation, i.e., by employing *a posteriori* probabilities. Since the mathematical treatment of this approach is considerably more complicated, it is not further pursued in this treatise.

4.2 Numerical results

Equations (4) and (5) were used to compute N, the minimum number of individuals required in generation BC1, as a function of parameters p, K, and R, if marker-facilitated selection is practised according to the scheme in Fig. 2 for $t = 1$, 2 or 3 backcross generations with a single marker (Table 2). The values for $p = 0$ (complete linkage between marker and resistance gene) correspond to the number of individuals required with direct selection. N increases only little from $p = 0$ to $p = 0.2$ for $t = 1$, but about 50 and 100 % for $t = 2$ and $t = 3$, respectively. Increasing family size K from one to four results in a substantial reduction of N, whereas a further increase to $K = 8$ leads only to a minor additional reduction of N. Choice of R determines the absolute size of N, but it hardly affects the behaviour of N as a function of the other parameters.

Table 2. Number of plants (N) in generation BC1 which are necessary to recover with probability $\alpha \geq 0.95$ at least R carriers of the resistance gene in generation BCt, when marker-assisted selection is practised according to the breeding plan in Fig. 2, using a single marker

	Generation		Recombination frequency p[2]					
R	BCt	K[1]	0.00	0.025	0.05	0.10	0.15	0.20	0.25
1	BC1	1	5	5	5	6	6	6	7
	BC2	1	11	12	12	14	16	18	20
		4	5	6	6	7	8	9	10
		8	5	5	6	7	7	8	10
	BC3	1	23	25	27	32	38	46	56
		4	6	6	7	8	10	12	15
		8	5	5	6	7	9	11	13
4	BC1	1	13	14	14	15	16	17	19
	BC2	1	29	31	32	36	41	47	53
		4	14	15	16	18	21	24	27
		8	13	14	15	17	19	22	26
	BC3	1	60	65	70	83	99	119	145
		4	16	17	18	22	27	32	40
		8	13	15	16	19	23	29	35
8	BC1	1	23	24	25	26	28	30	32
	BC2	1	50	53	56	62	70	79	91
		4	25	27	28	32	36	41	47
		8	23	25	26	30	34	38	44
	BC3	1	103	111	120	142	169	203	247
		4	27	29	32	38	46	56	68
		8	24	26	28	34	40	49	60

[1] K = family size in backcross generations BC2 and BC3.

[2] p = recombination frequency between resistance gene and marker.

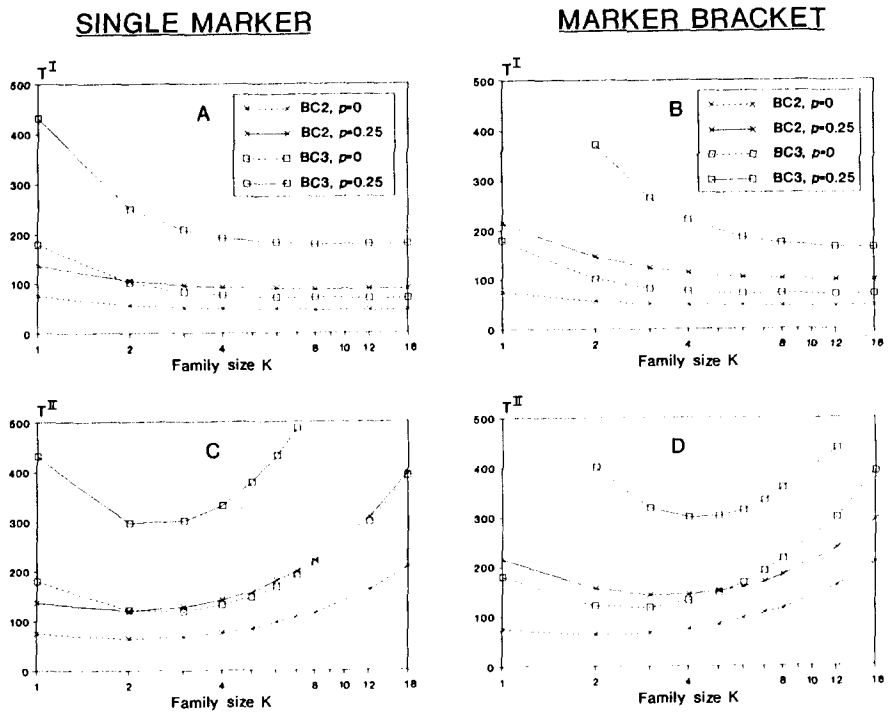


Fig. 3. Size of T^I and T^{II} (for definitions, see text) as a function of backcross family size K in order to recover with probability $\alpha \geq 0.95$ at least $R = 8$ carriers of the resistance gene in generations BC2 and BC3, when the breeding plan in Fig. 2 is employed. Marker-facilitated selection is based either on a single marker (A, C) or a marker bracket (B, D) with no interference of recombination, and p designates the recombination frequency between resistance gene and marker(s)

Fig. 3 shows the expected total number of marker-assayed individuals as a function of the family size K in backcross generations BCl ($2 \leq l \leq t$) for $t = 2$, 3 and $p = 0.0$ and 0.25. Calculations were made for both a single marker and a marker bracket. In the latter case, it is assumed that the resistance gene has the same recombination frequency p with both markers and interference of recombination is absent, resulting in $\gamma = (1-p)^2/2$. Under option I (sequential screening), T^I decreases with increasing K and approaches rapidly an asymptotic value (Figs. 3 A, 3 B). The reduction in T^I with increasing K becomes larger for greater values of t and p and is more pronounced with a marker bracket than a single marker. The asymptotic value of T^I is fairly closely attained at $K = 4$ and $K = 6$ for a single marker and a marker bracket, respectively. Taking into account practical aspects, there is probably no advantage employing a family size K greater

than three or four for t = 2 and t = 3, respectively.

Under option II (simultaneous screening), T^{II} is a convex function of K that rapidly approaches its minimum and increases thereafter (Figs. 3C, 3D). For the parameter values considered, the optimum family size is K = 2 or 3 and K = 3 or 4 for a single marker and a marker bracket, respectively. However, because the curves for T^{II} are rather flat in the vicinity of the minimum, K can be chosen within certain limits without significant increase in T^{II}.

Comparison of the curves for T^{I} and T^{II} reveals that sequential screening allows considerable savings in the number of marker assays compared to simultaneous screening of all K members of a family. Since sequential testing might be less convenient from a practical point of view, the results suggest that a reasonable compromise between options I and II would be sequential testing of two plants at a time and testing not more than four plants per family.

Fig. 4 presents graphs of N, T^{I} and V as a function of the recombination value p between the resistance gene and the marker(s) for t = 1, 2, 3 and K = 4, R = 8. The curves for N and T^{I} display essentially the same pattern. They increase almost linearly with p for t = 1 and show approximately an exponential increase for t = 3. This illustrates that tight linkage between the resistance gene and marker(s) becomes more and more important as the number of backcross generations with marker-facilitated selection increases.

The curves for N and T^{I} associated with a marker bracket have a steeper slope than those associated with a single marker. Hence, for a given value of p, the total number of marker-assayed individuals is greater for a marker bracket than a single marker. The reason being that the chances of obtaining individuals with the desired marker genotype are better for a single marker than for a pair of markers. Consequently, a large number of families must be screened in order to compensate for the higher risk of losing families because none of the K individuals has the desired marker genotype.

The curves for V (Figs. 4E, 4F) are very similar to those for N for a single marker, i.e., their slope increases considerably from t = 1 to t = 3. In contrast, with a marker bracket, V remains nearly constant from $p = 0$ to $p = 0.2$ and is practically identical for all values of t. Thus, using a marker bracket instead of a single marker has the advantage of reducing the number of individuals to be screened for the resistance gene in generation BCt. However, as pointed out above, this can be achieved only at the expense of a larger number of marker-assayed individuals.

For assessment of the efficiency of marker-facilitated vs. direct selection, the size of T^{I} (or T^{II}) and V for the respective recombination frequency p must be compared with the size of T^{I} (or T^{II}) for $p = 0$. As an example, let us consider marker-facilitated selection for t = 3 backcross generations using a marker bracket with $p = 0.1$. In this case, the expected total number of marker-assayed individuals T^{I} amounts to 106. The expected number of BC3 individuals (or their progenies) to be tested for presence of the resistance gene is V = 13. In comparison, the total number T^{I} of individuals (or their progenies) to be tested from BC1 to BC3 for presence of the resistance gene under direct selection ($p = 0$) is 76. The question of which selection method is superior when economical constraints are considered depends, therefore, primarily on the costs of marker assays relative to the costs of direct testing for the resistance trait. However, other merits of marker-facilitated selection such as acceleration of the breeding process might be equally important under certain conditions, outlined in the following section.

4.3 Practical applications

There are as yet few reports in the literature dealing with the application of marker-facilitated selection for the transfer of resistance genes. This section describes a variety of circumstances in which marker-facilitated selection might be advantageous and reviews some examples of its application.

4.3.1 Unreliable natural inoculum

With numerous plant diseases, selection progress is hampered by the fact that it is difficult to ensure uniform exposure to inoculum in the field. This is especially true for diseases caused by soil-borne pathogens and nematodes. Moreover, production and application of artificial inoculum can be difficult or expensive. Under these circumstances, marker assays are more reliable or simpler than tests for resistance and offer the possibility of screening breeding materials at a large scale with acceptable costs.

For example, RICK and FOBES (1974) proposed that the *Mi* gene conferring nematode

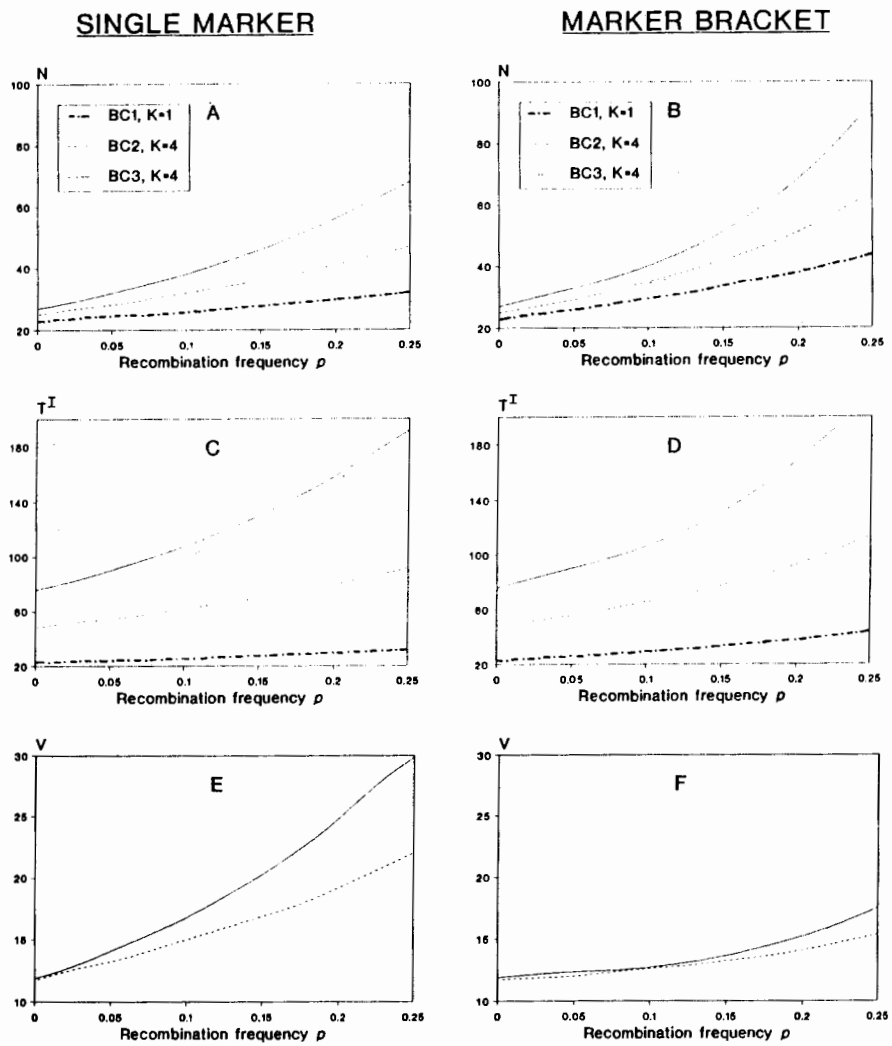


Fig. 4. Size of N, T^{I} and V (for definitions, see text) as a function of the recombination frequency p between resistance gene and marker(s) in order to recover with probability $\alpha \geq 0.95$ at least R = 8 carriers of the resistance gene in backcross generation BCt, when the breeding plan in Fig. 2 is employed with family size K = 4. Marker-facilitated selection is based either on a single marker (A, C, E) or a marker bracket (B, D, F) with no interference of recombination

resistance in tomato could be selected for by virtue of its tight linkage ($p \leq 0.01$, MEDINA-FILHO 1980) to the isozyme locus, *Aps-1*. Many seed companies have utilized this marker for the transfer of the *Mi* nematode resistance between breeding lines (TANKSLEY 1983).

4.3.2 Exotic or quarantined pathogens

If a virulent pathogen does not naturally occur in the test environment, artificial inoculation is usually prohibited for safety reasons. Nevertheless, in many instances the breeder would like to select for resistance. This is especially a problem for the off-season nurseries employed in breeding of major crop species like wheat and maize, which are usually located in other geographical areas than the target region of the newly developed varieties.

Under these conditions, marker-facilitated selection may be the only possible way to screen for resistances. With this in mind, the International Maize and Wheat Improvement Center (CIMMYT) is currently attempting to identify molecular markers linked to resistances against downy mildews in maize and corn streak virus (C. JAMES, pers. communication). The latter disease occurs under natural conditions erratically in some African countries and marker-facilitated selection could ensure consistent progress for materials grown in African nurseries or off-season nurseries elsewhere.

4.4.3 Strong environmental effects on resistance

Expression of many resistance genes can be strongly influenced by environmental conditions. Well-documented instances in which expression of infection type is affected by temperature occur among the cereal rusts, including stripe rust (*Puccinia striiformis* W.) and wheat stem rust (*Puccinia graminis* f. sp. *tritici*) (DYCK and KERBER 1985, ROELFS 1988). Marker-facilitated selection allows overcoming a similar problem in breeding for resistance to pea enation mosaic virus (PEMV). For this pathogen, disease symptoms are suppressed on susceptible plants at high temperatures, whereas low temperatures may cause disease-like symptoms on resistant plants (WEEDEN and PROVVIDENTI 1988). The authors found a resistance gene for PEMV to be closely linked (about 4 centimorgans) to an isozyme locus. They concluded that, in this case, marker-facilitated selection would be more reliable than direct screening using inoculum.

4.3.4 Developmentally regulated resistance

Some resistance phenotypes are not expressed until a late developmental stage. In contrast, molecular markers can be screened generally at a very early growth stage. Consequently, there would be no need to maintain a large number of plants until the age when direct resistance tests can be carried out. Early selection based on the marker genotype of young seedlings would be particularly beneficial for late expressed resistance traits in species with a long generation interval and high cultivation costs per plant (e.g. perennials, tree and timber species). For example, the resistance of western white pine (*Pinus monticola*) to blister rust (*Cronartium ribicola*) can be screened directly only with three- to five-year-old plants (GERHOLD 1973), whereas selection based on RFLP markers could be carried out much earlier.

4.3.5 Transfer of recessive resistance genes

The classical procedure of transferring a recessive resistance gene includes a progeny test after each or every second backcross generation to determine the presence of the desired allele (cf. FEHR 1987, pp. 363—365). In contrast, with marker-facilitated selection the transfer can be accomplished by an uninterrupted series of backcrosses because most markers are codominantly inherited. Hence, marker-facilitated selection could be used as a tool to accelerate resistance breeding programs especially in species with a long generation interval. WEEDEN et al. (1984) described the case of a tight linkage (approx. 2 centimorgans) between an isozyme marker and a recessive gene bestowing resistance against bean yellow mosaic virus in pea, where marker-facilitated selection could be superior to direct selection due to the savings in time.

4.4 Pyramiding of resistance genes

Pyramiding of resistance genes has been suggested as a strategy to provide durable resistance to both virulent and avirulent races of a pathogen (NELSON 1978). Pyramiding involves the accumulation of several resistance genes against the same pathogen into a single line or cultivar. Each of the resistance genes is either effective (pathogen avirulent) or defeated (pathogen virulent). Essentially, two arguments have been given for the assumed greater stability of resistance (durability) of a gene pyramid over time, depending on whether the genes are effective or defeated. With respect to effective resistance genes (i.e., genes conditioning resistance to all existing races), SCHAFER and ROELFS (1985) calculated that the probability of a pathogen overcoming the resistance of a pyramid with four to six is very low. Either virulant mutants that arise independently must be combined, or they must arise simultaneously or sequentially in the same isolate. For defeated resistance genes conferring some "residual" resistance, NELSON (1979) argued that a genotype whose resistance relies on the accumulated partial effects of numerous resistance genes should put little selection pressure on the pathogen and hence its resistance should be more durable. Although this concept has not been universally accepted, there is experimental evidence supporting the existence of residual resistance effects in some host/parasite systems (BRODNY et al. 1986, PEDERSEN and LEATH 1988).

The actual construction of a pyramid of resistance genes can be a long and costly process. With effective resistance genes, breeding can be complicated by the fact that, is difficult or often impossible to distinguish the various resistance genotypes. Pyramiding resistance genes with residual effects could also be very laborious, because the number of isolates needed for detecting a specific resistance gene can increase rapidly with the number of genes to be integrated (PEDERSEN and LEATH 1988).

By the same token as described previously for monogenic resistances, molecular markers could be exploited to facilitate pyramiding of resistance genes. Once the genes conferring resistance to the same pathogen are tagged by tightly linked markers, they could relatively easily be accumulated into a single genotype via marker-facilitated selection.

4.5 Selection against the "genetic background" of a donor parent

New, useful resistance genes are frequently found in wild relatives of a species or unadapted germplasm. Incorporation of genes from these sources (= DP) into breeding materials (= RP) may be accomplished by either simple crossing or using special techniques for interspecific hybridization. In both cases it is a difficult breeding task to eliminate undesirable genes from the donor parent, which by virtue of the breeding technique are also transferred into the breeding materials along with the desired trait.

Conventionally, undesirable genes from the donor parent are gradually eliminated by repeated backcrossing to the recurrent parent. ALLARD (1960) suggested six backcross generations as adequate in most instances, but more may be required for wide crosses. Although repeated backcrossing is generally regarded as a very efficient method for transfer of monogenic traits, its major shortcoming is the time requirement, particularly for species with a long generation interval.

TANKSLEY and RICK (1980) have proposed a method that could accelerate return to the RP genotype by exploiting markers as tags for chromosome segments. It presupposes a large number of markers covering the entire genome, preferably at least one marker per chromosome of chromosome arm. Plants from the first backcross generation or some other early segregating generation, which carry the desired resistance gene, are screened for their marker genotype at the seedling stage. Subsequently, multilocus marker selection is practised. Individuals with the maximum number of homozygous markers of the RP genotype are kept as parents of the next generation. A theoretical example by TANKSLEY and RICK (1980) demonstrates that, depending on the number of markers screened, the proportion of the RP genome in the selected individuals of the first backcross generation can correspond to that expected after three generations of backcrossing without selection. However, it should be stressed that effective selection for return to the RP genotype requires screening of several hundreds to thousands of plants. Considering economic aspects, this method seems appropriate only when marker assays are inexpensive and saving time is very important.

A special problem in introgression breeding is the recovery of the recurrent parent genome in the vicinity of the desirable resistance gene.

As expected from theoretical calculations mentioned earlier (STAM and ZEVEN 1981) and recently corroborated by experimental investigations with RFLPs (CHYI et al. 1989, GUPTA et al. 1989), even after several generations of backcrossing the target gene is flanked by a fragment from the DP of considerable length. By using markers closely adjacent to the target gene, one could screen for recombinants in this region. Thus it should be possible to shorten the length of this segment during an acceptable number of backcross generations.

An interesting modification of this approach was recently proposed by JUNG et al. (1986) to integrate an isozyme screening test into a resistance testing program aimed at selecting diploid sugar beet (*Beta vulgaris* L.) lines resistant to the beet cyst nematode (*Heterodera Schachtii* Schm.). They described isozyme markers that could be used to screen nematode resistant progenies obtained from crosses of diploid lines with monosomic addition lines, which carry the resistance gene(s) on an alien chromosome from a wild species, for presence of the entire alien chromosome. This procedure should allow to detect the rarely occurring resistant diploid genotypes obtained by translocation. Thus, resistant genotypes with a wild beet chromosome fragment as short as possible could be selected to avoid expression of wild beet characters except nematode resistance.

5. Cloning of resistance genes by chromosome walking

In recent years, there has been tremendous progress in the development of methods for gene transfer in plants (WEISING et al. 1988). However, application of these techniques for genetic engineering of plants resistant to pathogens is hampered by the lack of cloned genes for disease resistances (WENZEL 1985).

One of the strategies proposed for identifying and cloning genes is "chromosome walking." It uses DNA markers as the starting point. The first step in this approach is to identify DNA markers that are extremely tightly linked to the gene of interest by one of the mapping procedures described earlier. The corresponding DNA probe serves to initiate iterative screening of a genomic library that allows chromosome walking in the direction of the target gene (WINNACKER 1987).

Until recently, chromosome walking was not practicable in plants for several reasons. First, genetic linkage data in terms of recombination frequencies imply little about the actual physical distance between loci because of variation in the ratio of these two measures within the genome (MEAGHER et al. 1988). Second, even if one assumes a constant ratio over the whole genome, the highest resolution of a genetic linkage map obtainable with reasonable expenditure corresponds to a physical distance of several megabases in most plant species — a distance too long for walking with the available techniques. Third, the large proportion of repetitive DNA sequences in higher plants (FLAVELL 1982) could present another obstacle.

Newly developed methods such as preparative pulsed field gel electrophoresis (PFGE) for separation of very large DNA fragments, cloning of large DNA fragments as artificial chromosomes in yeast, or use of "jumping libraries" promise to overcome the present limitations for chromosome walking (JORDAN 1988). Furthermore, MEAGHER et al. (1988) recently described an assay for determination of the physical distance between a DNA probe and a target gene that uses a special subset of RFLPs (denoted as class II RFLPs) in combination with classical genetics and PFGE.

6. Conclusions and prospects

The introduction of molecular markers, particularly RFLPs, has opened new avenues in breeding for disease resistance. Whether these techniques will actually be employed in practical breeding programs depends on several factors such as the degree of linkage between the marker and the target gene, the saving in time, and the relative costs of direct vs. marker-facilitated selection. Therefore, it is not possible to make general inferences about the superiority of marker-based techniques, but rather it is necessary to decide case by case.

As mentioned earlier, efficient utilization of marker-facilitated techniques presupposes the existence of a detailed linkage map that is highly saturated with molecular markers. Such maps are either already available or presently developed for the major crop species. However, it appears questionable whether the considerable investments of time, labor, and funds required for developing a map will be available and justified for less important crop species.

At present, routine use of RFLP markers is severely limited by the significant costs and time required for RFLP assays. Since RFLP technology is now used in many areas of biology, it is anticipated that more efficient and simpler laboratory methods will be developed (LANDRY and MICHELMORE 1987). Moreover, many steps of the laboratory procedure could be automated by the use of robots, resulting in a substantial reduction of the costs. In addition, new classes of molecular markers might become available, such as the oligonucleotide polymorphisms recently proposed by BECKMANN (1988), that are less expensive and more suitable than RFLPs for large-scale assays. For these reasons, the prospects for a cost-effective implementation of marker-based techniques in breeding for disease resistances are very promising in the near future.

The author is indebted to Drs. C. R. BRONSON and M. LEE for critical reading of the manuscript and valuable suggestions; to Mrs. B. MCCLURKIN for her correction of the English text; and to Mrs. I. E. MELCHINGER for typing the manuscript and preparing the figures.

7. References

ALLARD, R. W., 1956: Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24**, 235—278.

——, 1960: Principles of Plant Breeding. John Wiley & Sons, New York.

APUYA, N. R., B. L. FRAZIER, P. KEIM, E. J. ROTH, and K. G. LARK, 1988: Restriction fragment length polymorphisms as genetic markers in soybean, *Glycine max* (L.) merrill. Theor. Appl. Genet. **75**, 889—901.

BECKMANN, J. S., 1988: Oligonucleotide polymorphisms: A new tool for genomic genetics. Bio/Technology **6**, 1061—1064.

——, and M. SOLLER, 1986a: Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica **35**, 111—124.

——, and ——, 1986b: Restriction fragment length polymorphisms in plant genetic improvement. In: Oxford Surveys of Plant Molecular and Cell Biology, vol.3, pp. 196—250.

BERNATZKY, R., and S. D. TANKSLEY, 1986: Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics **112**, 887—898.

BONIERBALE, M. W., R. L. PLAISTED, and S. D. TANKSLEY, 1988: RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120**, 1095—1103.

BRODNY, U., R. R. NELSON, and L. V. GREGORY, 1986: The residual and interactive expression of "defeated" wheat stem rust resistance genes. Phytopathology **76**, 546—549.

BURR, D., F. A. BURR, K. H. THOMPSON, M. C. ALBERTSON, and C. W. STUBER, 1988: Gene mapping with recombinant inbreds in maize. Genetics **118**, 519—526.

CHYI, Y., M. GUPTA, M. J. FROME, F. D. PARK, B. HOO, D. ZAITLIN, W. D. BANKS, J. PERKINS, and P. FILNER, 1989: Identification of RFLP markers for the *rhm* gene. Maize Genet. Coop. Newsl. **63**, 112.

DYCK, P. L., and E. R. KERBER, 1985: Resistance of the race-specific type. In: BUSHNELL, W., and A. ROELFS (eds.), The Cereal Rusts, vol. II, Academic Press, New York, p. 469—500.

FEHR, W., 1987: Principles of Cultivar Development, vol. 1. Macmillan, New York.

FIGDORE, S. S., W. C. KENNARD, K. M. SONG, M. K. SLOCUM, and T. C. OSBORN, 1988: Assessment of the degree of restriction fragment length polymorphism in *Brassica*. Theor. Appl. Genet. **75**, 833—840.

FLAVELL, R., 1982: Sequence amplification, deletion and rearrangement: Major sources of variation during species divergence. In: DOVER, G. A., and R. B. FLAVELL (eds.), Genome Evolution, Academic Press, London, pp. 301—323.

FREY, K. J., 1982: Multiline breeding. In: VASIL, I. K., W. R. SCOWCROFT, and K. J. FREY (eds.), Plant Improvement and Somatic Cell Genetics, Academic Press, New York, pp. 43—71.

GERHOLD, H. D., 1973: Forest trees. In: NELSON, R. R. (ed.), Breeding Plants for Disease Resistance — Concepts and Applications, Pennsylvania State University Press, p. 377—386.

GERTZ, A., 1989: Die Nutzung der Selbstinkompatibilität in der Roggenzüchtung unter besonderer Berücksichtigung der Entwicklung von Hybridsorten. Dissertation, Universität Hannover.

——, and G. WRICKE, 1989: Linkage between the incompatibility locus Z and a β-glucosidase locus in rye. Plant Breeding **102**, 255—259.

GUPTA, M., F. D. PARK, B. HOO, M. FROME, D. ZAITLIN, Y. CHYI, W. D. BANKS, and P. FILNER, 1989: Identification of RFLP markers for the *Ht1* gene by comparison of inbreds and their *Ht1*-conversions. Maize Genet. Coop. Newsl. **63**, 112.

HANSON, W. D., 1959: Early generation analysis of lengths of heterozygous chromosome segments

around a locus held heterozygous with backcrossing or selfing. Genetics **44**, 833—837.

HELENTJARIS, T., D. F. WEBER, and S. WRIGHT, 1986a: Use of monosomics to map cloned DNA fragments in maize. Proc. Natl. Acad. Sci. USA **83**, 6035—6039.

— —, M. SLOCUM, S. WRIGHT, A. SCHAEFER, and J. NIENHUIS, 1986b: Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. **72**, 761—769.

HEUN, M., and H. R. GREGORIUS, 1987: A theoretical model for estimating linkage in F_2 populations with distorted single gene segregation. Biom. J. **29**, 397—406.

HOISINGTON, D., 1987: Computer programs for use in linkage analysis. Maize Genet. Coop. Newsl. **61**, 48.

JORDAN, B. R., 1988: Megabase methods: A quantum jump in recombinant DNA techniques. BioEssays **8**, 140—145.

JUNG, C., P. WEHLING, and H. LÖPTIEN, 1986: Electrophoretic investigations of nematode resistant sugar beets. Plant Breeding **97**, 39—45.

KARLIN, S., 1966: A First Course in Stochastic Processes. Academic Press, New York.

KING, R. C., 1975: Handbook of Genetics, vol. 2. Plenum Press, New York.

LANDER, E. S., and P. GREEN, 1987: Construction of multilocus genetic linkage maps in humans. Proc. Natl. Acad. Sci. USA **84**, 2363—2367.

— —, — —, J. ABRAHAMSON, A. BARLOW, M. J. DALY, S. E. LINCOLN, and L. NEWBURG, 1987: MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1**, 174—181.

LANDRY, B. R., and R. W. MICHELMORE, 1987: Methods and applications of restriction fragment length polymorphism analysis to plants. In: BRUENING, G., J. HARADA, and A. HOLLAENDER (eds.), Tailoring Genes for Crop Improvement: An Agricultural Perspective, Plenum Press, New York.

LANDRY, B. S., R. V. KESSELI, B. FARRARA, and R. W. MICHELMORE, 1987: A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics **116**, 331—337.

MCCOUCH, S. R., G. KOCHERT, Z. H. YU, Z. Y. WANG, G. S. KUSH, W. R. COFFMAN, and S. D. TANKSLEY, 1988: Molecular mapping of rice chromosomes. Theor. Appl. Genet. **76**, 815—824.

MEAGHER, R. B., M. D. MCLEAN, and J. ARNOLD, 1988: Recombination within a subclass of restriction fragment length polymorphisms may help link classical and molecular genetics. Genetics **120**, 809—818.

MEDINA-FILHO, H. P., 1980: Linkage of Aps-1, Mi and other markers on chromosome 6. Rep. Tomato Genet. Coop. No. **30**, 26—28.

MORTON, N., 1955: Sequential tests for the detection of linkage. Am. J. Hum. Genet. **7**, 277—318.

MUEHLBAUER, G. J., J. E. SPECHT, M. A. THOMAS-COMPTON, P. E. STASWICK, and R. L. BERNARD, 1988: Near-isogenic lines — a potential resource in the integration of conventional and molecular marker linkage maps. Crop Sci. **28**, 729—735.

MURRAY, M., J. CRAMER, Y. MA, D. WEST, J. ROMERO-SEVERSON, J. PITAS, S. DEMARS, L. VILBRANDT, J. KIRSHMAN, R. MCLEESTER, J. SCHILZ, and J. LOTZER, 1988: Agrigenetics maize RFLP linkage map. Maize Genet. Coop. Newsl. **62**, 89—91.

NELSON, R. R., 1978: Genetics of horizontal resistance to plant diseases. Ann. Rev. Phytopathol. **16**, 359—378.

— —, 1979: The evolution of parasitic fitness. In: HORSFALL, J. G., and E. B. COWLING (eds.), Plant Disease, An Advanced Treatise, vol. 4, Academic Press, New York, pp. 23—46.

PEDERSEN, W. L., and S. LEATH, 1988: Pyramiding major genes for resistance to maintain residual effects. Ann. Rev. Phytopathol. **26**, 369—378.

RICK, C. M., and J. F. FOBES, 1974: Association of an allozyme with nematode resistance. Rep. Tomato Genet. Coop. No. **24**, 25.

ROELFS, A. P., 1988: Genetic control of phenotypes in wheat stem rust. Ann. Rev. Phytopathol. **26**, 351—367.

SAS, 1988: SAS Language Guide for Personal Computer, Release 6.03 Edition. SAS Institute Inc., Cary, NC.

SAX, K., 1923: The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgarus*. Genetics **8**, 552—560.

SCHAFER, J. F., and A. P. ROELFS, 1985: Estimated relation between numbers of urediniospores of *Puccinia graminis* f. sp. *tritici* and rates of occurrence of virulence. Phytopathology **75**, 749—750.

SEDCOLE, J. R., 1977: Number of plants necessary to recover a trait. Crop Sci. **17**, 667—668.

SILVER, J., 1985: Confidence limits for estimates of gene linkage based on analysis of recombinant inbred strains. J. Heredity **76**, 436—440.

SNAPE, J. W., 1988: The detection and estimation of linkage using double haploid or single seed descent populations. Theor. Appl. Genet. **76**, 125—128.

SOLLER, M., and J. S. BECKMANN, 1983: Genetic polymorphisms in varietal identification and genetic improvement. Theor. Appl. Genet. **67**, 25—33.

SOUTHERN, E., 1975: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**, 503—517.

STAM, P., and C. ZEVEN, 1981: The theoretical proportion of the donor genome in near-isogenic

lines of self-fertilizers bred by backcrossing. Euphytica **30**, 227—238.

STRICKBERGER, M. W., 1985: Genetics (3rd ed.). Macmillan, New York.

STUBER, C. W., and M. M. GOODMAN, 1983: Allozyme genotypes for popular and historically important inbred lines of corn. U. S. Dep. Agric., Agric. Res. Serv., Southern Ser., No. 16.

— —, J. F. WENDEL, M. M. GOODMAN, and J. S. C. SMITH, 1988: Techniques and scoring procedures for starch gel electrophoresis for maize (*Zea mays* L.). Tech. Bull. 286, Raleigh, NC.

SUITER, K. A., J. F. WENDEL, and J. S. CASE, 1983: LINKAGE-1: A PASCAL computer program for the detection and analysis of genetic linkage. J. Heredity **74**, 203—204.

TANKSLEY, S. D., 1983: Molecular markers in plant breeding. Plant Mol. Biol. Rep. **1**, 3—8.

— —, and T. J. ORTON, 1983: Isozymes in Plant Genetics and Breeding. Elsevier, New York.

— —, and C. M. RICK, 1980: Isozymic gene linkage map of the tomato: applications in genetics and breeding. Theor. Appl. Genet. **57**, 161—170.

WEBER, D., and T. HELENTJARIS, 1989: Mapping RFLP loci in maize using B—A translocations. Genetics **121**, 583—590.

WEEDEN, N. F., and R. PROVVIDENTI, 1988: A marker locus, Adh-1, for resistance to pea enation mosaic virus in *Pisum sativum*. J. Heredity **79**, 128—131.

— —, — —, and G. A. MARX, 1984: An isozyme marker for resistance to bean yellow mosaic virus in *Pisum sativum*. J. Heredity **75**, 411—412.

WEHLING, P., G. SCHMIDT-STOHN, and G. WRICKE, 1985: Chromosomal location of esterase, peroxidase and phosphoglucomutase isozyme structural genes in cultivated rye (*Secale cereale* L.). Theor. Appl. Genet. **70**, 377—382.

WEISING, K., J. SCHELL, and G. KAHL, 1988: Foreign genes in plants: Transfer, structure, expression, and applications. Ann. Rev. Genet. **22**, 421—477.

WENZEL, G., 1985: Strategies in unconventional breeding for disease resistance. Ann. Rev. Phytopathol. **23**, 149—172.

WINNACKER, E. L., 1987: From Genes to Clones: Introduction to Gene Technology. VCH Verlagsgesellschaft, Weinheim.

WRICKE, G., and P. WEHLING, 1985: Linkage between an incompatibility locus and a peroxidase isozyme locus (Prx7) in rye. Theor. Appl. Genet. **71**, 289—291.

YOUNG, N. D., J. C. MILLER, and S. D. TANKSLEY, 1987: Rapid chromosomal assignment of multiple genomic clones using primary trisomics. Nucleic Acids Res. **15**, 9339—9348.

— —, D. ZAMIR, M. W. GANAL, and S. D. TANKSLEY, 1988: Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics **120**, 579—585.

Author's address: Dr. A. E. MELCHINGER, Agronomy Department, Iowa State University, Ames/IA, 50011 (U.S.A.)

Plant Molecular Biology Reporter
15 (2) 1997 pages 154–162

Protocols

Use of STSs and SSRs as Rapid and Reliable Preselection Tools in a Marker-Assisted Selection-Backcross Scheme

Jean-Marcel Ribaut,[1] *Xueyi Hu, David Hoisington, and Diego González-de-León*

E-mail: jribaut@cimmyt.mx
Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), Int., Lisboa 27, A.P. 6-641, México 06600 D.F., México

Key Words: DNA isolation, maize, MAS, microsatellites, multiplexing, PCR screening, sequence-tagged sites

Abstract: We describe a new approach for using suitable STS and SSR markers as a powerful molecular tool for screening segregating populations involved in backcross schemes for marker-assisted selection, as a preselection step. Since it can be applied to very large populations, this preselection strategy allows one to increase substantially the pressure of selection at each backcross generation. The technique is fast and reproducible, and can be made even more efficient and cost-effective by simultaneous DNA amplification from different primer pairs. In the example illustrated here, three suitable PCR-based markers were used to complete the selection of 300 individuals out of 2300 in less than one month with two people working on the project.

During the last ten years, the development of new molecular marker systems has been one of the most dynamic areas in applied molecular genetics. The recent increase in the number of available PCR primers for STSs, developed principally from RFLP probes (MaizeDB, at http://www.agron.missouri.edu/top.html), and SSRs or microsatellites for species, such as maize (e.g., Taramino and Tingey, 1996), has opened new doors for genome mapping and especially for developing new PCR selection strategies based on markers. Given the typical length of STS and SSR primers (18–30 bp), the products of PCR amplification are highly reproducible and can be used reliably as molecular markers of plant genomes whenever they reveal polymorphisms. Such markers should map to specific loci, irrespective of which segregating population is used for mapping. If such loci are linked to genes involved in the expression of a trait of interest, it is possible, therefore, to use the corresponding markers for selection of that trait. Before the availability of suitable PCR-based markers, marker-assisted selection (MAS) schemes had to be designed within the constraints imposed by the most reliable marker system, RFLPs. This technique imposed a practical (both financial and logistical) limit on the number of individuals that could be genotyped efficiently. Here, we present a simple strategy that can be used for marker-based selection on populations that are at least ten times larger than is currently practical with RFLPs. This approach will be illustrated with a practical example of preselection, using three PCR markers at the BC_2 level, a part of the MAS project conducted at CIMMYT to improve drought tolerance in tropical maize.

[1]Author for correspondence

Abbreviations: BC, backcross; BME, β–mercaptoethanol; CTAB, alkyltrimethylammonium bromide; MAS, marker-assisted selection; PCR, polymerase chain reaction; QTL, quantitative trait locus; RFLP, restriction-fragment-length polymorphism; SSR, simple sequence repeat; STS, sequence-tagged site; Taq, *Thermus aquaticus*.

Procedures

The preselection strategy

The basic goal of a BC MAS scheme consists of the transfer of one or more pertinent genomic segments from a donor to a target variety. By screening a large BC population with a reliable PCR-based marker for a particular locus, all the individuals not having the desired allele (statistically half of the population) are discarded. Those retained are then assayed at a second locus to repeat the preselection cycle, and so on, for as many loci as is practical. The population can thus be quickly reduced in size to one that can then be handled more efficiently using RFLPs for genotyping at further loci of interest, and to evaluate the remaining percentage of the donor parent genome in the different preselected individuals. Here we describe the implementation and application of this powerful preselection tool, using a specific example of preselection that employs three PCR-based markers on 2300 BC_2F_1 individuals.

Sample harvest

Approximately 15 cm (about 350 to 500 mg) of a leaf tip of each segregat-

ing BC maize plant were harvested, placed in a 15-mL plastic tube, kept on ice until arrival at the laboratory, and then stored directly at -80°C. Harvesting early in the morning is recommended to avoid the dehydration of leaf tissue that might be encountered later in the day.

Reagents

extraction buffer: 0.1 M tris pH 8.0, 0.15 M EDTA pH 8.0, 2.1 M NaCl, 2% PVP, 2% CTAB, 1% 14 M BME

DNA extraction for PCR amplification

- Extract leaf juice from frozen tissue using a sap extractor (Clarke et al., 1989). As the tissue is pressed, add 1.2 mL of extraction buffer, and collect the extract, about 1 mL, at the tip of the rollers in 2-mL tubes.
- Incubate the extracts in a water bath or an oven at 65°C for 20–40 min; mix twice by inversion during this incubation. Remove the tubes from the heat, and let cool down for 5–10 min.
- Extract the samples with 1 mL of octanol-chloroform (1:24). Mix by inversion for 5 min and spin in a table-top centrifuge at 3200 rpm for 10 min.
- Transfer 600 µL of the aqueous supernatant containing the DNA to 1.5-mL Eppendorf tubes.
- Add 75 µL of 5 M NaCl and precipitate DNA with 700 µL of cold absolute ethanol.
- Spin DNA down, decant ethanol, and dry under a weak vacuum for 30 min.
- Resuspend in 500 µL TE, pH 8.0.
- Quantify DNA of a small number of randomly chosen samples, using a fluorometer, as a rough guide for adjusting sample DNA concentration.
- One person can easily handle 100 to 150 samples per day.

Selection of PCR primers

Primers for both STS and SSR markers were tested. A marker is considered of interest when its theoretical map position fits within the two flanking markers of a genomic segment (typically resulting from the detection of a QTL for a given trait) that is to be transferred. In the case of STSs, many terminal sequences of RFLP probes have been published, and thus appropriate primers can be designed. For maize, many sequences are available from the Maize Genome Database (MaizeDB), and primers can be designed using various computer programs available commercially or in the public domain. In the case of maize SSRs, primer

sequences are directly available on the Internet (MaizeDB) or from published papers (e.g., Taramino, and Tingey, 1996; Maize Genetics Cooperation Newsletter No. 70, 1996). After amplification of parental DNA, polymorphic products can be mapped using a small number of segregating individuals to confirm the locus position in the map of the segregating population under study.

PCR protocol and identification of polymorphisms

In order to increase the screening efficiency of markers, more than one pair of primers can be used for multiplex amplification, using the same PCR conditions (Table I). This implies that care must be taken in designing or choosing primer pairs for which the annealing temperatures are reasonably compatible, and for which there is no overlap in the size of the amplification products. PCR reactions were run in 96-well thermal cyclers (ERICOMP, Inc.) under the following conditions: 2 min at 93°C; 30 cycles of 1 min at 93°C, 2 min at 56°C, and 2 min at 72°C; and finally 5 min at 72°C. Different reaction volumes can be used according to the final number of gel wells that might need to be loaded for separation of the amplification products (Table I). Different gel systems can be used according to the sizes of the amplification products and clarity of the polymorphisms: normal agarose gels (1.5%, e.g., SeaKem LE, FMC BioProducts, Rockland, USA) for larger fragments, high-resolution aga-

Table I. Composition of PCR reactions as a function of the number of primer pairs that can be used for simple or multiplex amplification.

Component	Stock	Volume in µL for 1 primer pair	2 primer pairs	3 primer pairs
ddH_2O		9.4	4.4	8.2
Primer 1.1	2 µM	2.0	2.0	2.0
Primer 1.2	2 µM	2.0	2.0	2.0
Primer 2.1	2 µM	-	2.0	2.0
Primer 2.2	2 µM	-	2.0	2.0
Primer 3.1	2 µM	-	-	2.0
Primer 3.2	2 µM	-	-	2.0
Taq buffer	10	2.0	2.0	3.0
dNTP mix	2.5 mM each	1.6	1.6	2.4
$MgCl_2$	50 mM	0.8	0.8	1.2
Template DNA	10 ng/µL	2.0	3.0	3.0
Taq polymerase	5 U/µL	0.2	0.2	0.2
Total reaction vol.		20	20	30

rose gels (3.0%, e.g., 2% MetaPhor agarose, FMC BioProducts, Rockland, USA, and 1% SeaKem LE agarose) for a wide range of fragment sizes, or polyacrylamide gels (6%) for small fragments and tight polymorphisms. For both normal and high-resolution agarose gels, we routinely run 20x25-cm gels with eight 30-well combs for a total of 240 samples. Expensive high-resolution agarose can be reused at least three times after running off the previous samples from the gel and remelting it. An increase in the use-efficiency of acrylamide gels can also be achieved by loading additional sets of samples after the first set has been electrophoresed for a period of time.

Results and Discussion

Three key issues were considered during the development of each step of the preselection protocol: high reproducibility and reliability of the amplification products of large numbers of samples, relative rapidity of the operation, and cost effectiveness.

Efficiency of DNA extraction

The harvest of predetermined lengths of leaf-tip tissue into pre-labeled plastic tubes can be very fast, and samples do not need to be weighed. Since little material is needed for DNA extraction, harvesting can be performed at an early stage of plant development. Direct extraction of frozen samples using a sap extractor obviated the need for lyophilization and/or laborious grinding procedures, which would be too time consuming for a large number of samples. The DNA extraction protocol is not the fastest one (see Rogers et al., 1996), but the chloroform-partition step assures good DNA quality and more uniform amounts. High molecular DNA is obtained with relatively low partial degradation (about 30%), and could be amplified by PCR. No effect on PCR amplification was detected in the presence of RNA, and therefore an RNAse treatment was not necessary. Even with the variation in the initial fresh weight among leaf samples, the range of DNA concentrations at the end of the extraction never exceeded a factor of four. This variation in the final DNA concentration was found not to affect the proper amplification of the DNA samples. Consequently, with a DNA quantification of only a few randomly chosen samples (e.g., 30) the final resuspension volume of TE could be established. In our example, the mean DNA concentration chosen was 10 ng/µL, obtained with a final resuspension volume of 500 µL of TE. This amount of DNA can be used for more than a hundred PCR amplifications.

PCR primer selection criteria

The two kinds of PCR-based markers used in our strategy, STSs and SSRs, have different advantages and weaknesses. In both cases, their number in publications and databases is increasing rapidly, thereby facilitating the choice of those most appropriate for a given purpose. In the case of STS markers, their genomic location can be directly associated with the position of the corresponding RFLP locus detected by the probe from which the STS primers were derived. The chance of detecting directly a polymorphism with a pair of STS primers is, however, very low compared to the corresponding RFLP probe. In effect, the latter will usually reveal polymorphisms based on changes in restriction sites within or flanking the sequence homologous to the probe, and therefore over a relatively long DNA sequence (typically up to 20 kb). Polymorphisms revealed by an STS will be strictly limited to sequence variation within the limits of the amplified sequence, which is much shorter (generally 0.6 to 3 kb). Often, a restriction digest of the amplified product may be needed to increase the chance of revealing a polymorphism between two individuals. By contrast, since SSRs are hypervariable regions, these detect much higher levels of polymorphism. Moreover, the initial use of SSR primers is faster, since the primer sequences are publicly available, while for STSs, primers may have to be designed from published sequence information. SSRs generally yield much smaller amplification products and thus polymorphisms are more difficult to resolve (e.g., higher resolution gel systems may be needed), whereas STS products are longer and polymorphisms can be resolved using cheaper conventional gel systems, as suggested in the **Procedures** section. The quality of selected PCR-based markers is crucial for successfully achieving the preselection protocol, since the same primers will be reused at the successive levels of BC selection and over a very large number of samples. So far, we have found that all the maize SSR and STS primers tested in our laboratory reveal loci that map at the expected positions.

Increased efficiency of polymorphism screening through multiplex PCR amplifications and gel loading

For both STSs and SSRs, primer length is at least 18 bases. PCR amplification using such primer lengths is highly sequence specific and reproducible. Although each primer pair may have an optimum temperature for annealing to the template DNA, the sequence specificity of the amplification reaction due to high primer length allows some degree of flexibility in selecting a single annealing temperature for two or more

pairs of primers. This means that more than one target sequence could be amplified, under the same PCR conditions, in the same reaction tube (Table I). Since STSs and SSRs usually produce very different amplification product lengths, they often can be combined. It is often possible to resolve the two polymorphisms simultaneaouly on the same gel (Figs. 1 and 2). This strategy, which in some cases might be extended to more than two markers, significantly reduced selection costs and screening time. Note in Table I that the same amount of Taq polymerase (by far the most expensive item in the reaction) could be used irrespective of the number of primer pairs included in the reaction.

A practical example

The use of PCR-based markers as a preselection step was successfully tested for the transfer of quantitative QTLs involved in the expression of a favorable trait for drought tolerance in tropical maize. This MAS experiment is in progress at CIMMYT. Based on selected QTL intervals, several STS and SSR markers were tested. After confirmation of the map position of loci identified by several markers, three different primer pairs were selected for screening the BC_2F_1 individuals, and thus reduce the large initial population size by approximately a factor of eight. Selection was based on the quality of the polymorphism identified and their

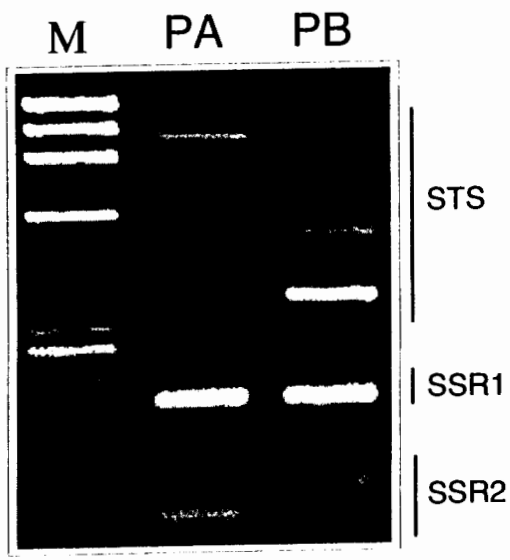


Fig. 1. Multiplex PCR amplification of DNA of two parental maize lines that gave rise to the BC population. (PA = donor line Ac7643, PB = recurrent line CML247). The three PCR primer pairs used were for an STS (umc67) and two SSRs (1: phi071; 2: MAG.IA01). The molecular weight marker (M) was ϕx174-*Hae* III.

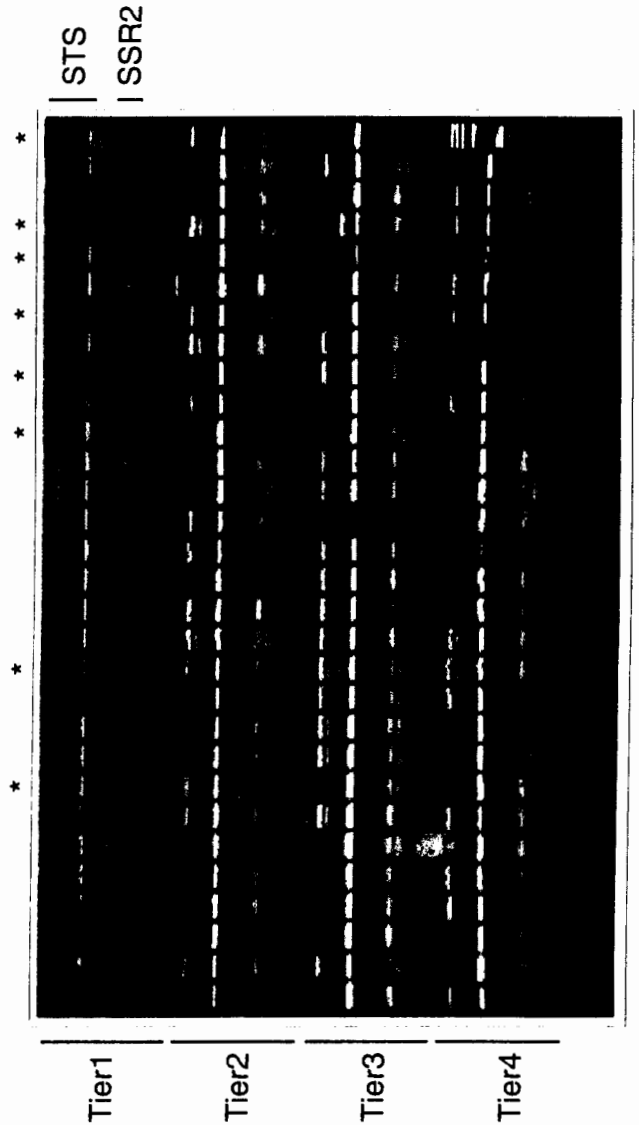


Fig. 2. Screening of BC_2F_1 plants. Two PCR-based markers pre-select individuals heterozygous at the two loci, STS (umc67) and SSR2 (MAG.IA01), with the gel lanes marked by an asterisk (*) for the first tier. Half the Metaphor agarose gel is shown, with 119 individuals arranged in four tiers, the last lane being ϕx174-*Hae III*.

combinability for multiplex amplification. These primers amplified an STS (umc67) on chromosome 1, an SSR (MAG.IA01, DuPont) on chromosome 2, and an SSR (phi071, Pioneer) on chromosome 10. Multiplex amplification for the three different primers was possible, and the amplification products with DNA from the donor (Ac7643) and the recurrent elite line (CML247) (Ribaut et al., 1996) are presented in Fig. 1. Fragments of DNA amplified with phi071 had to be separated on polyacrylamide gels, while for the other two, polymorphisms were clear enough to resolve them simultaneously on the same MetaPhor agarose gel. Results on Fig. 2 show the screening of a sample of BC_2F_1 plants with STS umc67 and MAG.IA01, after multiplex amplification. The percentage of PCR reaction failures was low, less than 3%, demonstrating the high efficiency of the method. The early harvest of plant material for the preselection step allowed one to achieve the total selection process, using PCR-based and RFLP markers, within a growing plant season. No more than a month was needed for the preselection PCR-based step (300 plants selected out of 2300, two technicians), and about three months were needed for the RFLP screening of the whole genome for final selection (300 to 10 plants, one technician).

Acknowledgments. We wish to thank E. Perotti for advice and helpful discussion, as well as E. Huerta and E. Magaña for their excellent technical work. This work was supported by the Swiss Development Corporation (SDC) and the United Nations Development Program (UNDP)

References

Clarke, B.C., L.B. Moran, and R. Appels. 1989. DNA analyses in wheat breeding. Genome 32:334-339.

Ribaut, J.-M., D.A. Hoisington, J.A. Deutsch, C. Jiang, and D. González-de-León. 1996. Identification of quantitative traits loci under drought conditions in tropical maize. 1. Flowering parameters and the anthesis-silking interval. Theor. Appl. Genet. 92:905-914.

Rogers, H.J., N.A.. Burns, and H.C. Parkes. 1996. Comparison of small-scale methods for the rapid extraction of plant DNA suitable for PCR analysis. Plant Mol. Biol. Reptr. 14:170-183.

Taramino, G., and S. Tingey. 1996. Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39:277-287

Plant Molecular Biology Reporter
15 (2) 1997 pages 163–169

Protocols

A Plant-Based Expression System for Matching cDNA Clones and Isozymes

P. R. Matthews, F. Gubler, and J. V. Jacobsen

E-mail: petermat@pican.pi.csiro.au
CSIRO Division of Plant Industry, P.O. Box 1600, Canberra, ACT 2601, Australia

Key Words: *Avena sativa*, α-amylase cDNAs, cDNA isozyme matching, *Hordeum vulgare*, isoelectric focusing, microparticle bombardment, oat aleurone, transient expression

Abstract: Isozymes of barley α-amylase were matched to cDNAs that encode them using transient expression in oat aleurone layers. Four cDNAs, including two that are previously unpublished, were inserted into oat aleurone cells by microparticle bombardment. The cDNAs were under the control of the *Act1* promoter of rice. Expression levels were sufficient for in-gel detection of enzyme activity following isoelectric focusing of aleurone homogenates. The system has also proved useful in characterizing a hybrid β-glucanase gene.

The use of biotechnology to improve plant growth and performance often entails the characterization of cDNA clones encoding enzymes influencing desirable plant traits. The traits can then be manipulated using molecular genetic techniques. The enzymes are often part of a family of enzymes that all perform the same or similar function, but it is usually desirable to select one particular isozyme for manipulation.

The matching of cloned cDNA to the isozyme encoded is thus a necessary step in many projects for the genetic engineering of plants. Although genes encoding cereal a-amylases have been expressed in heterologous eukaroytic systems, such as yeast (Rothstein et al., 1984; Sogaard et al., 1991) and *Xenopus laevis* oocytes (Aoyagi et al., 1990), it is

Abbreviations: ABA, abscisic acid; *Act1*, rice actin gene; GA_3, gibberellic acid; IEF, isoelectric focusing.

PH

Techniques et utilisations des marqueurs moléculaires
Montpellier (France), 29-31 mars 1994
Ed. INRA, Paris 1995 (Les Colloques, n°72)

Marker-assisted backcrossing: a practical example

M. RAGOT[1], M. BIASIOLLI[1], M.F. DELBUT[2], A. DELL'ORCO[1], L. MALGARINI[1], P. THEVENIN[1], J. VERNOY[2], J. VIVANT[2], R. ZIMMERMANN[1] and G. GAY[1].

1 Ciba Seeds, CH-4002 Basle, Switzerland.
2 Département d'Amélioration des Plantes, INRA-Domaine d'Epoisses, F-21110 Genlis, France

Summary

That molecular markers allow fast recovery of recurrent parent genotype in backcross programs is undisputed. Restriction Fragment Length Polymorphisms (RFLP's) were used in maize to introgress by backcross a transgene construct, containing phosphinothricin resistance and insecticidal protein genes, from a transformed parent into an elite inbred line. At each generation plants carrying the transgene construct were selected based on their phosphinothricin resistance, and further characterized with RFLP's. Both maximum recovery of recurrent parent genotype and minimum linkage drag were taken into account for marker-based selection. Embryo rescue was used to shorten generation time. Progress towards recurrent parent genotype was spectacular. Levels of recurrent parent genotype recovery which would normally be observed, in the absence of selection, in the BC_6 generation were obtained at the BC_3 generation, about one year after BC_1 seeds had been planted. Besides the evidence already provided by RFLP's, phenotypic evaluation of the backcross-derived near-isogenic lines will constitute an additional check of the completeness of the conversion.

Introduction

Backcrossing has been a common breeding practice for as long as elite germplasm has been available. It has mainly been used to introgress single Mendelian traits, such as disease resistances or quality factors, into elite germplasm (Allard 1960; Hallauer and Miranda 1981). One of the most attractive attributes of backcrossing is that it allows to perform targeted modifications without disrupting the existing overall genetic balance of the recurrent parent.

However, production of fully converted near isogenic lines through classical backcrossing procedures is a lengthy procedure, if at all possible. Theoretically, a minimum

of seven classical backcross generations are required to recover more than 99% of recurrent parent genotype, assuming no linkage drag. The attractiveness of classical backcross procedures is therefore substantially diminished for crops, such as maize (*Zea mays* L.), where the turn-over of elite cultivars is very fast. In addition, full recovery of recurrent parent genotype is usually not achieved through classical backcrossing, which may result in deleterious agronomic effects. Murray *et al.* (1988) reported about 90% recurrent parent genotype recovery in two BC_{10}-equivalent conversions (A632Ht and A632Rp) of the maize line A632. The conversions had retained respectively 4 and 7 donor fragments in addition to the one carrying the gene of interest.

Reduction in the number of backcross generations needed to obtain fully converted individuals has been shown theoretically, or from simulations, to be achievable through the use of molecular markers (Tanksley *et al.* 1989; Hospital *et al.* 1992; Jarboe *et al.* 1994). Because they provide thorough characterization of the genetic variability at each backcross generation, markers allow to take full advantage of this variability by applying the highest possible selection intensity.

Efficiency of marker-assisted backcrossing was investigated through an experiment aimed at introgressing a single genetic factor (a transgene construct) from a donor into a recipient maize line.

Materials and methods

Plant Material

A hemizygous transgenic maize line of Lancaster origin was used as donor parent to introgress its transgene construct, through repeated backcrossing, into a recipient parent from the Stiff Stalk germplasm group. Both parents are proprietary elite lines. The transgene construct carries both a phosphinothricin resistance gene and synthetic genes encoding the entomotoxic fragment of the CryIA(b) *Bacillus thuringiensis* protein (Koziel *et al.* 1993). Transformation was achieved through microprojectile bombardment (Koziel *et al.* 1993) and resulted in a single insertion (*Bt* locus), on chromosome 1 (Figure 1).

Backcross protocol

The F1 progeny of the cross between the donor and the recipient was screened for the presence of the transgene construct by applying Basta, a phosphinothricin-based herbicide, onto each plant. Resistant individuals were then used to generate BC_1 progeny.

For each backcross generation, except the BC_4, individuals were planted in multipots and sprayed with Basta to eliminate those which did not carry the transgene construct. To avoid the stress resulting from treatment with Basta, BC_4 plants carrying the transgene construct were identified using Southern blots probed with the *pat* and *Bt* genes. Resistant plants were transplanted in an open-soil greenhouse and leaf-sampled for molecular marker

analyses. Results of marker analyses were made available at the latest two weeks after flowering. A single plant was selected, of which all backcross-derived embryos were rescued and transferred onto tissue culture medium. Plantlets that developed from these embryos first underwent a greenhouse acclimation phase, while still growing on tissue culture medium, before being transplanted into multipots. Backcross cycles lasted, on average, four months.

Molecular marker analyses

Restriction Fragment Length Polymorphisms (RFLP's) were used to establish genotypes in all four generations. RFLP detection involved either radioactive or chemiluminescent techniques. For the BC_1 generation, 61 marker-enzyme combinations were chosen from among those revealing polymorphism between donor and recipient. They provided coverage of the entire genome, defining intervals of about 25 cM in size, and contained two loci tightly linked to the *Bt* locus, CG320 and CG415, respectively 5 and 16 recombination units away (Figure 1). For subsequent generations, markers analyzed in the BC_{n+1} generation comprised both those for which the selected BC_n plant was heterozygous, or tightly linked ones, and additional ones located in chromosomal segments for which the selected BC_n plant was heterozygous (Table 1). Marker map positions were obtained from independent reference populations and confirmed by analysis of segregation in the BC_1 generation.

Selection procedure

At each generation plants were ranked based both on the percentage of homozygous recurrent-parent-genotype and on the extent of linkage drag around the *Bt* locus, in an attempt to integrate both criteria. Plants for which two or more adjacent markers had missing values were not included in the analyses. Success or failure of the pollinations also contributed to the selection procedure. One single plant was selected at each generation: the best ranking one of those for which a backcross progeny of size 100 or more (50 or more for the BC_3 selection) was available.

Results and discussion

Selection for the gene of interest

The observed segregation ratios for phosphinothricin resistance (Table 1) were not significantly different ($P < 0.05$) from the expected 1:1, as shown by Chi-square tests.

Recurrent parent genotype recovery

Statistics for the genotyped plants are summarized in Table 1. Calculations were performed taking the whole genome into account, including the *Bt* locus. The "perfect" backcross-derived plant therefore counts one heterozygous chromosome segment, that

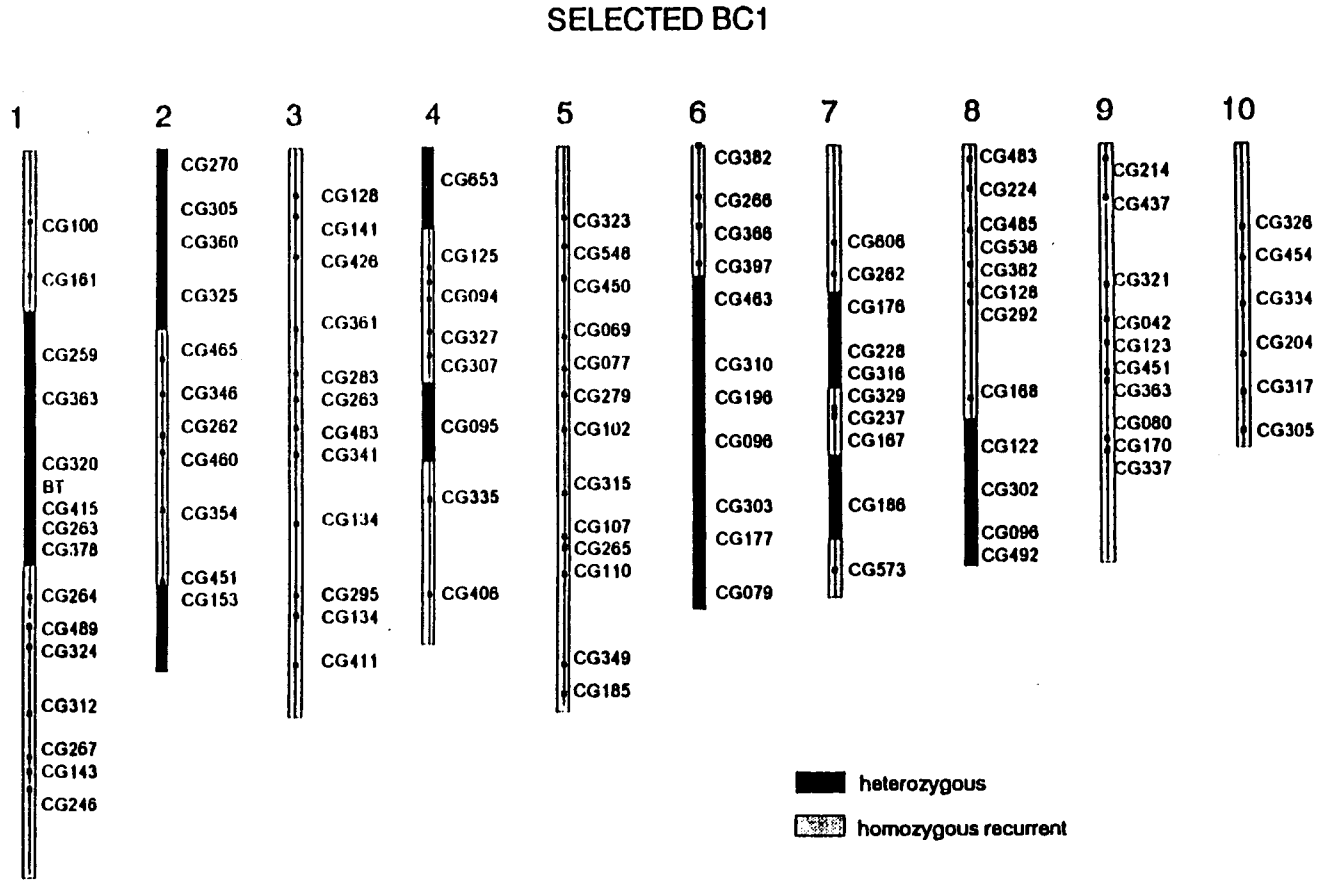


Figure 1-a: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

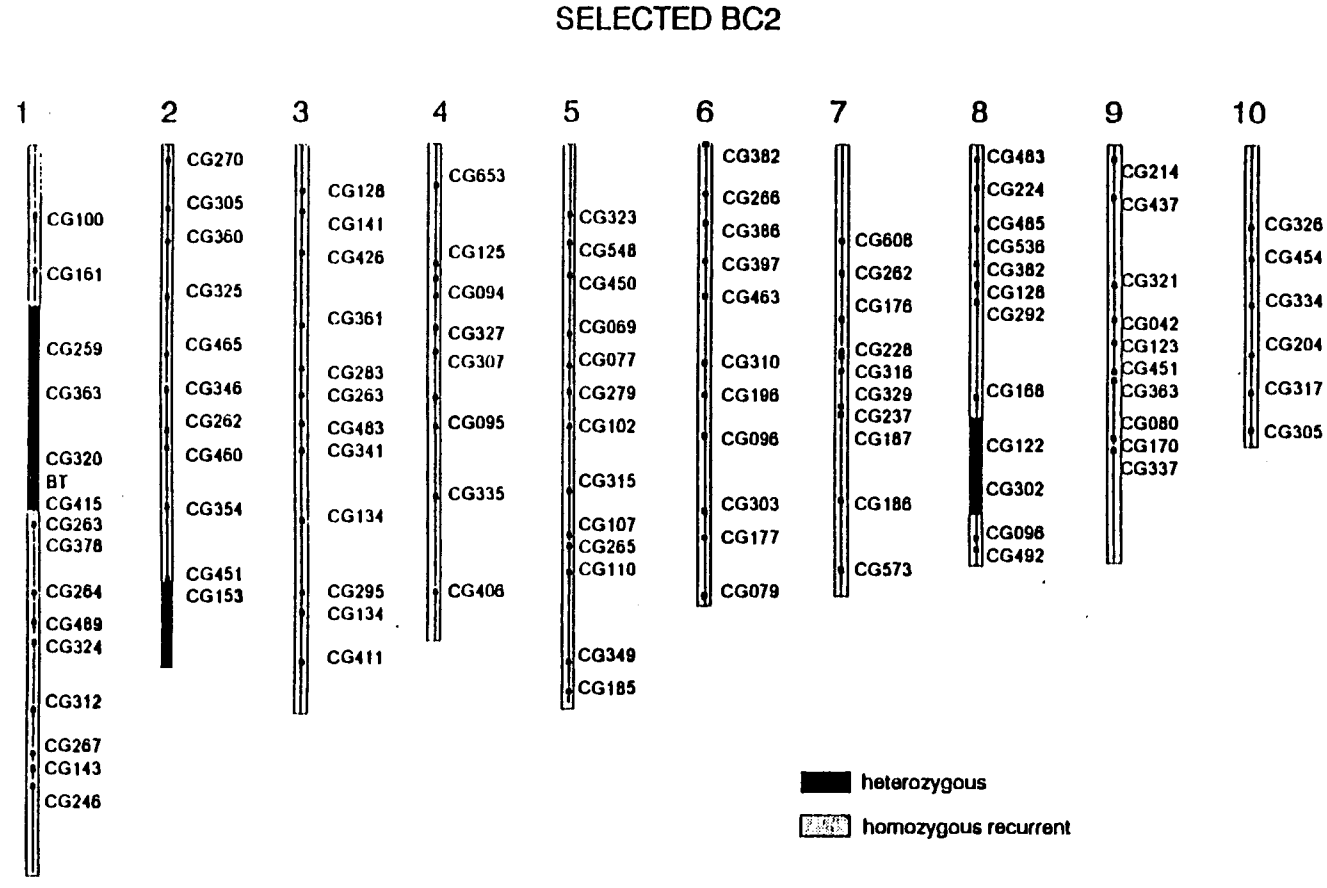


Figure 1-b: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

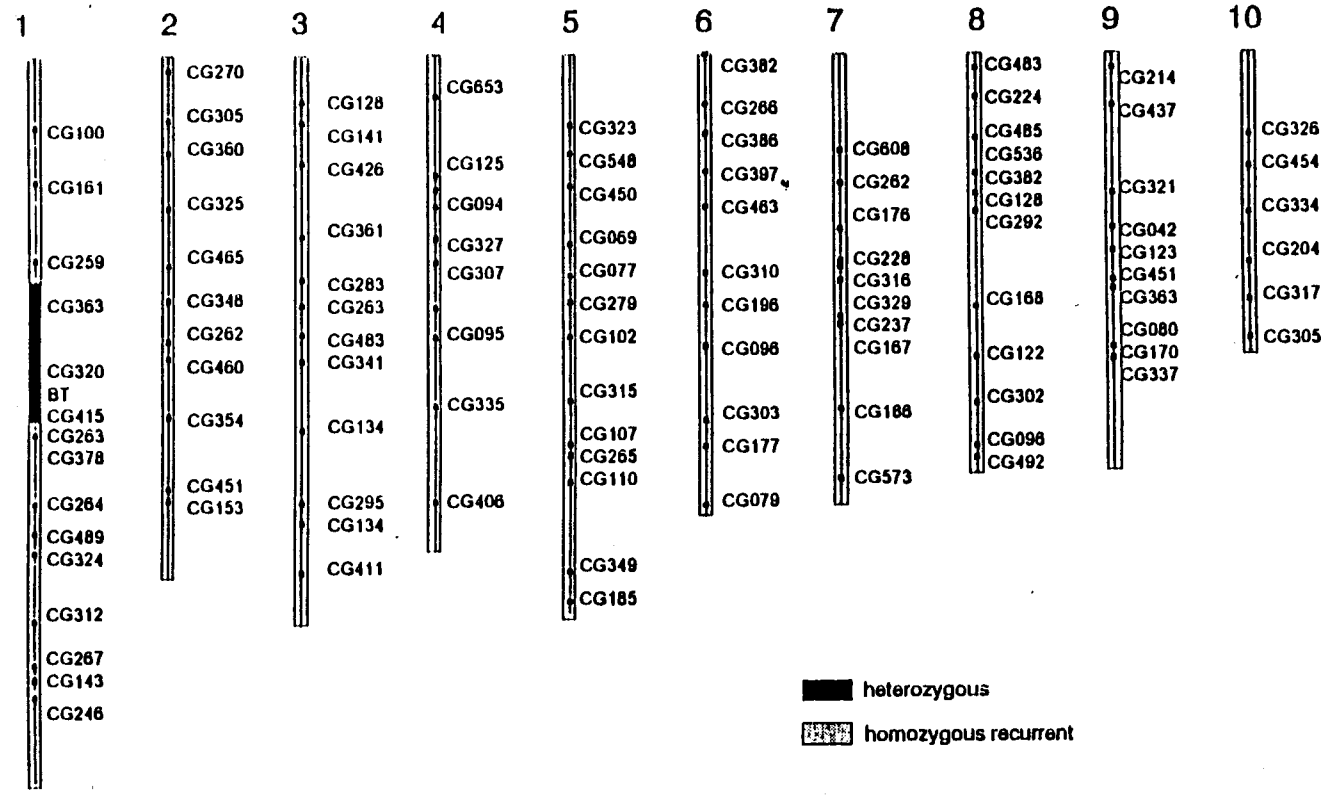


Figure 1-c: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

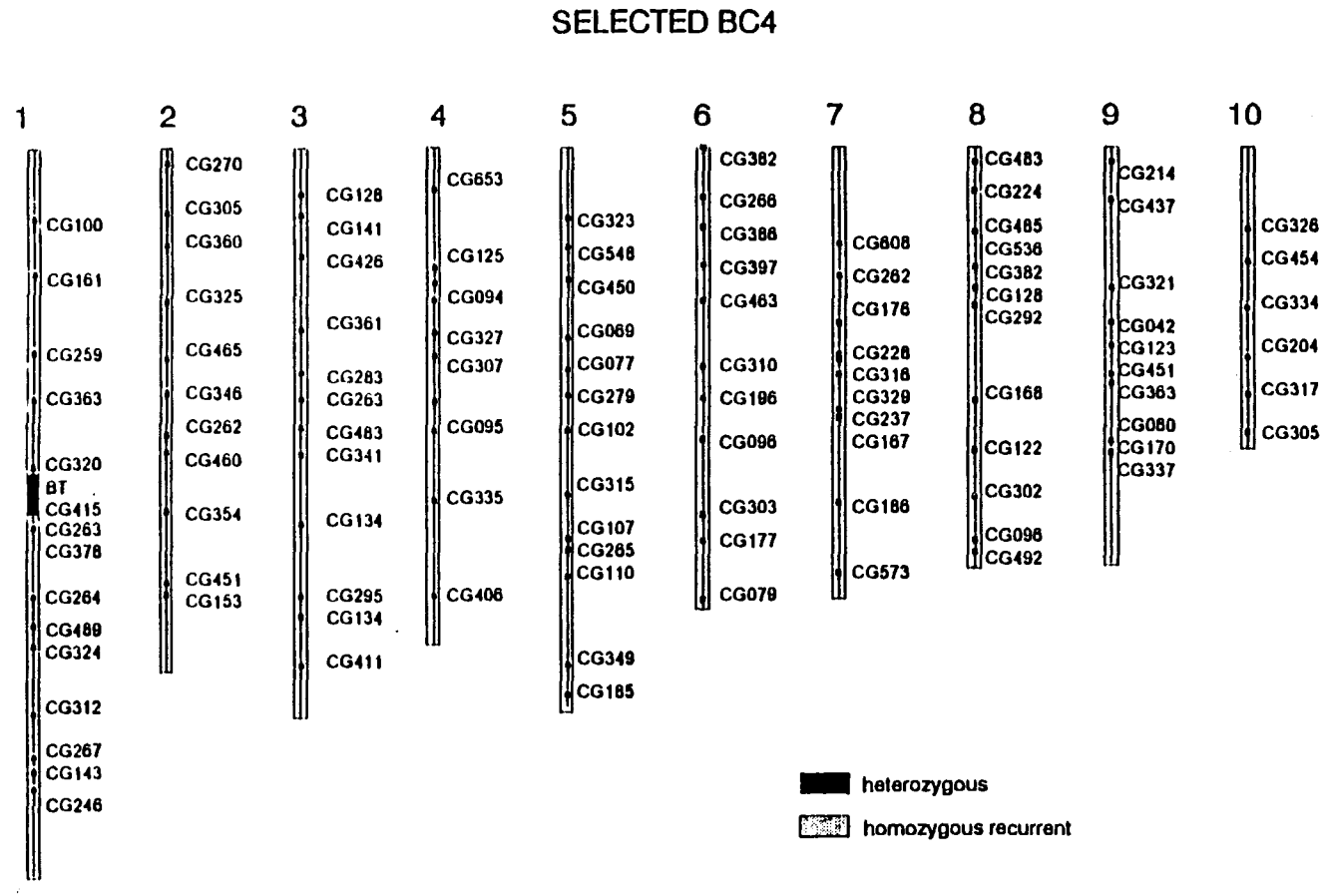


Figure 1-d: Genetic maps of the backcross-derived individuals selected in the first four generations of a marker-assisted backcross program. The locus to be introgressed (*Bt*) is located on chromosome 1.

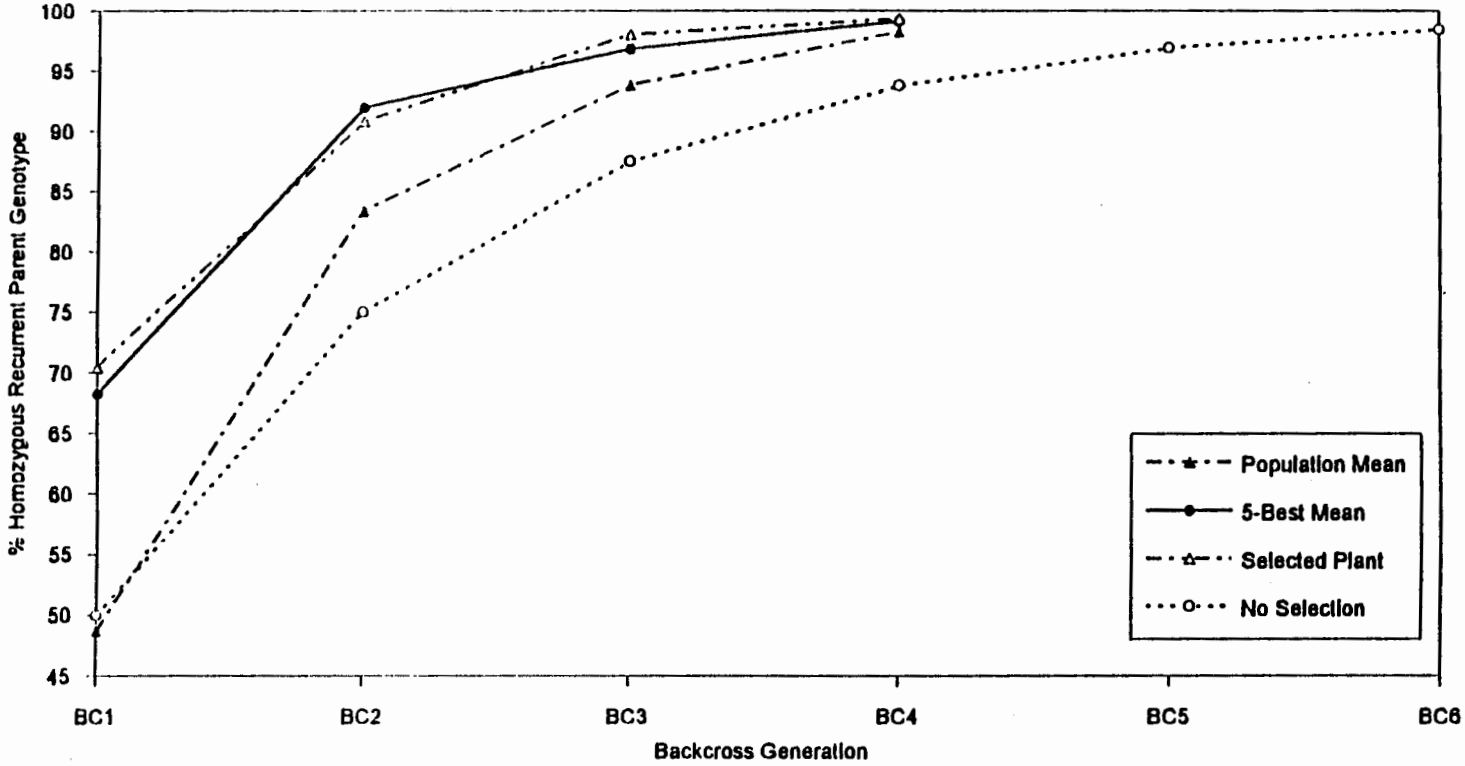


Figure 2: Recovery of recurrent parent genotype through backcrossing, with or without marker-assisted selection

Table 1: Proportion and characteristics of plants carrying the genes of interest, in the first four generations of a marker-assisted backcross program.

generation	% phosphinothricin resistant plants	RFLP genotyping			nb plants analyzed *	% homozygous recurrent parent genotype				nb heterozygous chromosome segments ***			
		nb plants	nb loci	nb datapoints		mean	std dev	5-best mean **	selected plant	mean	std dev	5-best mean **	selected plant
BC1	49.05	96	61	5856	87	48.72	10.35	68.31	70.45	11.01	2.17	7.75	6
BC2	44.65	61	22	1342	30	83.42	5.64	91.98	90.84	5.03	1.54	3.20	3
BC3	46.32	72	10	720	71	93.83	1.85	96.82	98.03	2.20	0.71	1.60	1
BC4	-	26	3	78	26	98.23	0.49	99.09	99.36	1.00	0.00	1.00	1

* Plants for which two or more adjacent markers had missing values were not included in the analyses

** Mean value of the five individuals having the five highest percentages of homozygous recurrent parent genotype.

*** Including the segment carrying the transgene construct.

comprising the *Bt* locus. It also displays 99.36% of homozygous recurrent-parent-genotype. The remaining 0.64% corresponds to the average relative length of the chromosome segment containing the *Bt* locus, which depends on the two flanking markers chosen.

The mean percentage of homozygous recurrent-parent-genotype of the BC_1 generation was slightly lower than the expected 50%. This can be explained by linkage drag around the *Bt* locus, given that this percentage was computed based only on plants selected for heterozygosity at the *Bt* locus. For all other backcross generations the mean percentage of homozygous recurrent-parent-genotype was much higher than what would have been observed, should no selection have been done (Figure 2).

The percentage of homozygous recurrent-parent-genotype of the selected plant (Table 1) and the average of the five largest values (Table 1) were always very similar to one another, and much superior to the population mean value (Figure 2). The percentage of homozygous recurrent-parent-genotype of the selected plant was found only once, in the BC_2 generation, to be smaller than the average of the five largest values. This corresponded to the only time when the selected plant was not the one with the maximum percentage of homozygous recurrent-parent-genotype. The plant had been selected because it displayed a favorable recombination on one side of the *Bt* locus (Figure 1).

The percentage of homozygous recurrent-parent-genotype of the selected BC_1 plant was almost equal to that of an unselected BC_2, that of the selected BC_2 was larger than that of an unselected BC_3, that of the selected BC_3 was barely smaller than that of an unselected BC_6, and that of the selected BC_4 was equal to that of the "perfect" backcross-derived plant, given the set of markers that was used. Such rates of recurrent parent genotype recovery are consistent with results of simulation analyses. Jarboe *et al.* (1994) who used the maize genome as a model reported that three backcross generations and 80 markers were needed to recover 99% of recurrent parent genotype.

Number of donor chromosome segments

The number of heterozygous chromosomal segments decreased from one backcross generation to the next. Plants selected at each generation were not necessarily those which had the lowest number of heterozygous chromosomal segments (Table 1). However, with the set of markers used, BC_3 and BC_4 plants were recovered which contained only one heterozygous chromosomal segment: that comprising the *Bt* locus.

Linkage drag

Linkage drag around the *Bt* locus was estimated, relative to the length of chromosome 1. Its value was found to lie between 24.0 and 48.4% for the selected BC_1 individual, between 17.6 and 34.8% for the selected BC_2, between 2.0 and 24.0% for the selected BC_3, and between 0.0 and 8.4% (respectively 0.0 and 14.5 cM) for the selected BC_4.

The two values given for each generation are extreme values of linkage drag, which correspond to extreme positions of the crossing-overs in the marker-defined intervals flanking the transgene construct locus. Therefore the true linkage drag value of the selected BC_4 is likely to be less than 1.3% of the genome. Although this maximum value may appear to be somewhat high, reflecting the limited selection pressure put here on linkage drag, it is much lower than what would be expected from classical backcross programs (Stam and Zeven 1981; Tanksley *et al.* 1989). Practically, in a study of *Tm*-2 conversions of tomato cultivars obtained by a large number of classical backcross cycles, Young and Tanksley (1989) found that the sizes of the introgressed fragments ranged between 4 and 51 cM.

Conclusion

These results clearly demonstrate that molecular markers provide important time and quality advantages over classical procedures for the production of near-isogenic lines through backcrossing. Only four backcross generations were necessary to recover, in less than a year and a half from planting of the BC_1's, individuals which appeared to be genotypically fully converted. Nevertheless, it is likely that recovery of recurrent parent genotype could proceed even faster than in the experiment described herein, should the appropriate protocol and resources (population size, number and position of markers) be allocated.

Comparison of BC_4-derived lines with the recurrent parent for both morphological markers and agronomic performance (including hybrid performance) will be performed in order to confirm the completeness of the conversion.

References

ALLARD, R.W. (1960) Principles of plant breeding. Wiley, New York, NY.

HALLAUER, A.R., and J.B.MIRANDA, Fo. (1981) Quantitative genetics in maize breeding. Iowa State University Press, Ames, IA.

HOSPITAL, F., C.CHEVALET, and P.MULSANT (1992) Using markers in gene introgression breeding programs. Genetics 132:1199-1210.

JARBOE, S.G., W.D.BEAVIS, and S.J.OPENSHAW (1994) Prediction of responses to selection in marker-assisted backcross programs by computer simulation. In: Abstracts of the second international conference on the plant genome. Scherago International Inc. 38.

KOZIEL, M.G., G.L .BELAND, C. BOWMAN, N.B. CAROZZI, R. CRENSHAW, L. CROSSLAND, J .DAWSON, N. DESAI, M. HILL, S. KADWELL, K. LAUNIS, K. LEWIS, D. MADDOX, K. McPHERSON, M.R. MEGHJI, E. MERLIN, R. RHODES, G.W. WARREN, M. WRIGHT, and S.V. EVOLA (1993) Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Bio/Technology 11:194-200.

MURRAY, M.G., Y.MA, J.ROMERO-SEVERSON, D.P.WEST, and J.H.CRAMER (1988) Restriction fragment length polymorphisms: what are they and how can breeders use them ? In: D.Wilkinson ed.,

Proceedings of the 43rd annual corn and sorghum industry research conference. American Seed Trade Association 43:72-87.

STAM, P., and C. ZEVEN (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. Euphytica 30:227-238.

TANKSLEY, S.D., N.D. YOUNG, A.H.PATERSON, and M.W.BONIERBALE (1989) RFLP mapping in plant breeding: new tools for an old science. Bio/Technology 7:257-264.

YOUNG, N.D., and S.D.TANKSLEY (1989) RFLP analysis of the size of chromosomal segments retained around the *Tm*-2 locus of tomato during backcross breeding. Theor. Appl. Genet. 77:353-359.

Les Colloques, n° 72

Techniques et utilisations des marqueurs moléculaires

Montpellier (France), 29-31 mars 1994

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
147, rue de l'Université - 75338 Paris Cedex 07

Ce colloque a été organisé par l'INRA, l'Université de Perpignan, le CNRS, le CIR
l'ORSTOM, l'ENSAM, AGROPOLIS et l'Université de Montpellier II

avec la participation financière

du Conseil Régional Languedoc-Roussillon,
du District de Montpellier
et des sociétés
RHONE POULENC AGROCHIMIE,
AMERSHAM LIFE SCIENCE,
AMILABO,
APPLIED BIOSYSTEMS
BALCO,
BECKMAN,
GENSET
LABIMAP,
MOLECULAR DYNAMICS,
PERKIN ELMER,
PHARMACIA BIOTECH,
PROLABO,
SCHLEICHER ET SCHUELL,
SIGMA ALDRICH

Editeurs / *Editors*

A. BERVILLE - M. TERSAC
INRA - Station d'Amélioration des Plantes
2 Place Viala, 34060 Montpellier Cedex 1, France

En vente / *For sale*

INRA Editions
Route de St Cyr, 78026 Versailles Cedex, France

ISBN : 2-7380-0614-0

RFLPs for Rapid Recurrent Selection

Marlin Edwards[1] and Lane Johnson[2]

The Pillsbury Company/Green Giant, LeSueur, MN 56058

Molecular markers have been used widely for genetic studies in plants during the past 10 years. Many researchers have evaluated genetic factors underlying the inheritance of quantitative traits. However, there are few published accounts of efforts to improve expression of quantitative traits with marker-based breeding strategies. Stuber and Edwards (1986) demonstrated gains in yield and other quantitative traits in two maize populations from a single generation of selection based on genotypes at many isozyme loci that were known to be linked to quantitative trait loci (QTL). Marker loci in that study were poorly distributed, representing no more than 30% to 40% of the genome. Nevertheless, one cycle of marker-based selection was found to be as effective as one cycle of mass selection, which should have exploited QTL throughout the genome. Stuber and Sisco (1992) later showed improvements in the performance of an elite maize hybrid by backcrossing marker-linked QTL into the original parental inbred genotypes. Although limited numbers of QTL were successfully introgressed through three cycles of backcrossing into parental genotypes, "improved" hybrids often performed significantly better than the original "check" hybrid.

Although these earlier investigations are encouraging, they do not represent the comprehensive effort to improve genotypes that are characteristic of most commercial breeding efforts in crop plants. These former examples were limited in the number of genes or traits that were targeted by the marker-based improvement process. This contrasts with the commercial breeding procedures of pedigree breeding or recurrent selection, which often attempt simultaneous improvement in a several different quantitative attributes and which allow the manipulation of gene frequencies at numerous loci across the genome.

Theoretical studies have also provided encouragement that molecular markers might be used to enhance selection response under some circumstances (Edwards and Page, 1994; Lande and Thompson, 1990; Zhang and Smith, 1992). However, numerous aspects of quantitative trait inheritance, which may profoundly affect the potential of marker-based selection, still elude basic understanding. Examples include: the degree to which marker-based investigations are detecting the net effect of numerous, linked QTL rather than single QTL; the degree to which trait expression is controlled by epistasis; the importance of pleiotropy vs. linkage in producing trait correlations; and the degree to which the genomic regions important for a given trait are conserved across populations. Because many of these issues are not easily addressed using currently available technology, empirical studies remain necessary in determining the usefulness of markers as a selection tool.

Population development and QTL estimation

This investigation represents an effort to develop improved, new inbred lines using QTL-mapping information from elite sweet corn breeding populations. A brief diagram of the procedures that were used is presented in Fig. 1. Each of the two populations was derived by crossing two elite, proprietary inbred lines of sweet corn and developing a population of 160 single-seed-descent F_4 families without selection from the F_2. These families were testcrossed to two other elite inbreds to develop F_1 hybrid seed for performance testing. The testers were chosen based on known combining ability with the parental inbreds used to develop the improvement cross population. Crosses between parental inbreds and tester inbreds produce either production hybrids or elite experimental hybrids. Thus, the germplasm used represents promising breeding populations, and improvements from such a scheme would have commercial utility.

The F_4 families were grown in the greenhouse in the winter of 1989–1990. Leaf samples from about 10 seedlings per F_4 family were bulked for genetic analysis. Genomic DNA was isolated from each sample (Dellaporta et al., 1983) using 0.5 to 1.0 g of fresh leaf tissue. DNA was digested with *Eco*RI, loaded in 0.7% agarose gels (0.5 µg per lane), electrophoresed, and transferred to nylon membranes. The probe (100 ng) was radiolabeled by using the random-primer method (Feinberg and Vogelstein, 1983). Hybridizations were performed as described by Church and Gilbert (1984) in a Robbins Hybridization Oven (Robbins Scientific). Genotypes of each F_4 family were determined at each of 61 marker loci distributed throughout the genome in population A and 52 loci in population B. Approximate locations of the RFLP loci in each of the two populations, based on mapping data from the F_4 families, are indicated in Fig. 2.

The testcross F_1 progenies derived by crossing each F_4 family by the chosen inbred testers were evaluated in the field in the summer of 1990. Experimental design for each population was a randomized complete-block design with two replicates; each was planted at a different location near LeSueur, Minn. The planting dates for population A were June 20 and June 21. Population B was planted on June 1 and June 9. Plant spacing was 34 kernels per 7.6-m plot with 1.5-m alleys between ranges. Row spacing was 0.76 m, for an effective within-plot plant population of 56,800 plants/ha. Days to 50% silk exposure were recorded for each plot and used as a basis for harvesting each plot at an equivalent stage of ear development. Plots were hand-harvested and ear samples were weighed and taken to a pilot processing line, where they were subjected to the same handling processes that are used in commercial processing. These include: mechanical husking,

[1]To whom correspondence should be sent. Current address: Northrup King Co., 317 330th St., Stanton, MN 55018–4308.
[2]Agricultural Research, Green Giant Co., 1201 N. 4th St., LeSueur, MN 56058.

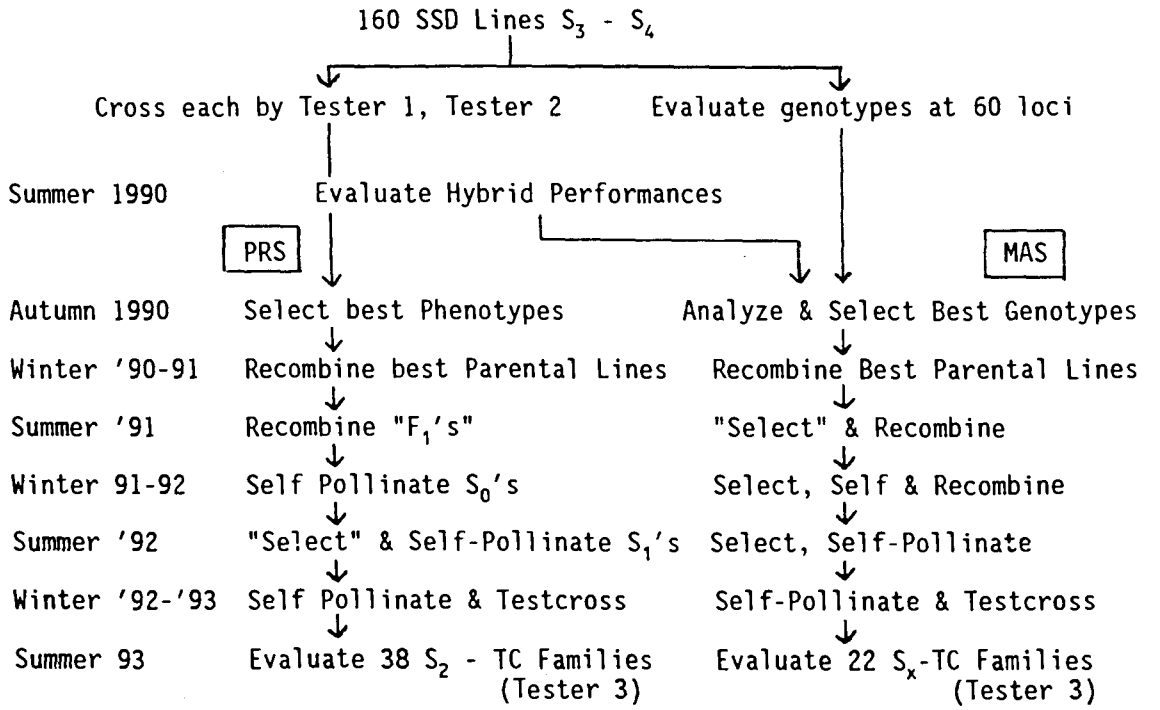


Fig. 1. A flow chart of the procedures used for identification and selection of QTLs in two, elite, proprietary sweet corn populations. The left-hand column indicates the season in which each activity was conducted. The center column, labeled "PRS," indicates the process used for phenotypic recurrent selection in population B. The right-hand column, labeled "MAS," indicates the process used for marker-based (or assisted) selection both in populations A and B (with the following exceptions: 50, rather than 22, lines were selected for evaluation in population A, and two testers were used to evaluate the responses in population A and one was used to evaluate population B).

mechanical cutting, sample washing, and vacuum-pack canning and thermal processing of a subsample of the produce of each plot. The gathered data included a number of attributes in the field and processing facility, as well as final consumer quality attributes. A total of more than 40 attributes were recorded on each of the 640 plots that were evaluated in each population.

The data from each population and for each tester was subjected to a single-factor analysis of variance for each marker locus with genotypic classes at the locus considered as the treatment classes. Variation attributable to each locus was subjected to a single-degree-of-freedom contrast to determine the average effect of gene substitution at the locus. A selection index was constructed using a linear function of a number of component traits, with each F_4 line, Y_i, receiving a value calculated as: $\Sigma_{j=1}{}^{n}\omega_j[(\gamma_{ij}-\mu_j)/\sigma_j]$ where ω_j is the weight assigned to the jth trait, γ_{ij} is the performance of the ith line (in hybrid combination) for the jth trait, μ_j is the mean performance of the jth trait across lines, and σ_j is the phenotypic standard deviation among lines for the jth trait. The rationale for this approach is as follows. First, the performance of each line (for each trait) was converted to normal, standard units so all traits exhibited an equivalent mean and variance, regardless of the original units. It was then possible to multiply trait performances by a weight for each trait in accordance with the perceived importance of changing the performance of the trait by 1 SD. Traits for which greater improvements were desired were assigned large weights; less-important traits were given smaller weights. The weighted performances of each F_4 testcross progeny, when summed across traits, determine a "net value" of the F_4 progeny for the selection index.

Thirty-four traits were used in the index. These traits can be categorized in three classes: 11 traits were measures of consumer quality acceptability, 13 were measures of processability in manufacturing (process suitability or performance), and 10 were measures of agronomic performance. The weights placed on each trait ranged from a minimum of <1% to a maximum of 20% of the total index weight. Index performances for the F_4 progeny were then subjected to a single-factor analysis, similar to that used for the original trait data to determine the marker loci that were associated with genetic factors having an apparently desirable or undesirable effect on index performance.

In this fashion, genetic factors were identified throughout the genome that appeared to affect index performance. Some of these apparently affected only one (or a few) of the component traits; others affected a larger number of traits. Some chromosome regions had no apparent effect on the selection index, despite their importance for some traits, due to antagonistic effects on other traits that resulted in a nonsignificant average affect on the selection index.

The chromosome regions that were chosen for subsequent use in marker-based selection are indicated in Fig. 3 A and B, for populations A and B, respectively. The small tickmarks along the chromosomes indicate the location of marker loci used in the investigations. The width of the ellipses indicate the relative importance of each region, which are based on the significance level of the contrasts that estimate additive effects in each chromosome region. The type of crosshatching in the ellipse indicates which of the two parental genotypic classes exhibited the desired direction of effect on the performance index. Regions without ellipses either were judged to be unimportant in the inheritance of the performance index or were regions where genetic markers were unavailable for interpretation of genetic effects. Figure 3 can be interpreted as a "genetic recipe" of the inbred line, which, if it could be developed, would be presumed to exhibit the optimum performance index—hereafter, the "ideal" genotype. The objective of deriving individuals like the ideal genotype was then undertaken using marker-based recurrent selection.

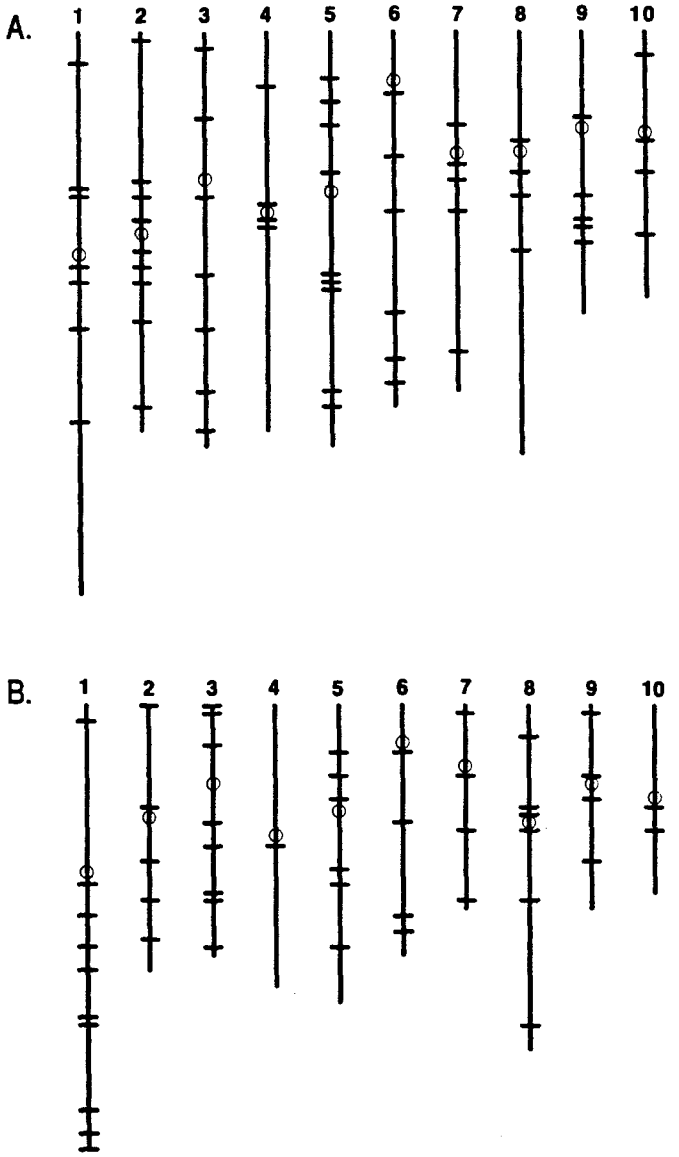


Fig. 2. The approximate positions of RFLP markers used for identification and manipulation of QTLs in two sweet corn populations, designated A and B, respectively.

Marker-based selection

The original 160 F_4 lines were surveyed to determine which of them possessed a genotype most closely resembling the ideal genotype. This was done by calculating a "breeding value" for each line by summing across loci the directional estimates of the effects of gene substitution at each locus. The top 10% of the lines were then examined to determine which pairwise matings would have the best probability of producing offspring like the ideal genotype. This was done by computer simulation where haploid genotypes (gametes) were drawn at random from the parental genotypes and united to generate potential offspring from each pairwise combination among the parental lines. One-hundred potential offspring genotypes were sampled from each pairwise combination. The average breeding value of the top 10% of these offspring was used as a basis for choosing six F_4 parental lines as well as the crosses to be conducted among them. A similar process was followed for subsequent selection cycles.

This method was easily implemented with simple, computerized algorithms and identified selections analogously to conventional methods in plant breeding, wherein the optimum improvement population is characterized by the highest frequency of favorable alleles and the largest genetic variance (Sprague and Eberhart, 1977).

The above methodology was applied for each of four cycles of recurrent selection that were conducted between Oct. 1990 and July 1992. For each cycle, seedlings were grown to about the five- to six-leaf stage before leaf tissue was sampled and the genotype of each seedling determined. Typically, 100 to 300 individuals were evaluated at each cycle of selection. Genotypic information was always determined before flowering, such that selection could proceed at the rate of two cycles per year.

Fifty lines were selected for evaluation from the last cycle of selection in population A and 22 from population B. Hybrids were also reconstituted from cycle 1 of population B by testcrossing the six originally selected F_4 families to provide a measure of whether additional gains were achieved from marker-based selection between cycle 1 and cycle 4. These hybrids were evaluated along with the marker-based selections from cycle 4.

Methods for evaluation of selection response

The selected lines were crossed to testers to develop testcross F_1 progeny for evaluation of the response to selection. Two testers were used to evaluate the performance of breeding lines from population A. Only one of the original tester inbreds was used to evaluate the responses in population B, due to the larger number of breeding lines that required evaluation. Testcross progeny were also developed using the original, elite inbred parental lines to provide a baseline for comparison to the "improved" lines. These hybrids, hereafter denoted "parental hybrid checks," provide a valid response "check" due to the fact that all possible additive and dominance effects (excluding epistasis) are represented among the four, parent–tester hybrid combinations in each population.

Testcross progeny of the selected line-hybrids and parental hybrid checks for each population were planted in a randomized complete-block at each of three planting dates in the summer of 1993. Parental hybrid checks were replicated for each planting date. Experimental protocol was the same as that previously described for mapping QTL effects. The flooding conditions in the Midwest during 1993 led to the loss of all but the first plantings of each population (May 15 planting date for both populations). The data presented was collected from the earliest planting. Due to the preliminary nature of conclusions drawn from this limited data, evaluations are being conducted again during the summer of 1994.

Analysis for each trait was conducted using analysis of variance with the model $Y_{ij} = \mu + \gamma_i + \lambda_{j(i)}$, where μ is the grand mean, γ_i is the mean of the ith class of lines (marker selected or parental testcross for population A), and $\lambda_{j(i)}$ is the performance of the jth line in the ith population. The performance index (which was the basis of the genotypic selection) was recalculated by summing normal standard deviation units, which were weighted as before, except that the standard deviation unit in the denominator of each trait expression was that calculated from the original investigation in 1990. This was done such that the relative weights per trait would be equivalent to those used in selection.

Phenotypic (traditional) selection

In addition to developing and evaluating a series of lines from marker-based selection, another group of lines was developed from population B using traditional phenotypic selection. This process involved selecting the top four lines based on average testcross performance and recombining them. The progeny were exposed to visual selection during two cycles of

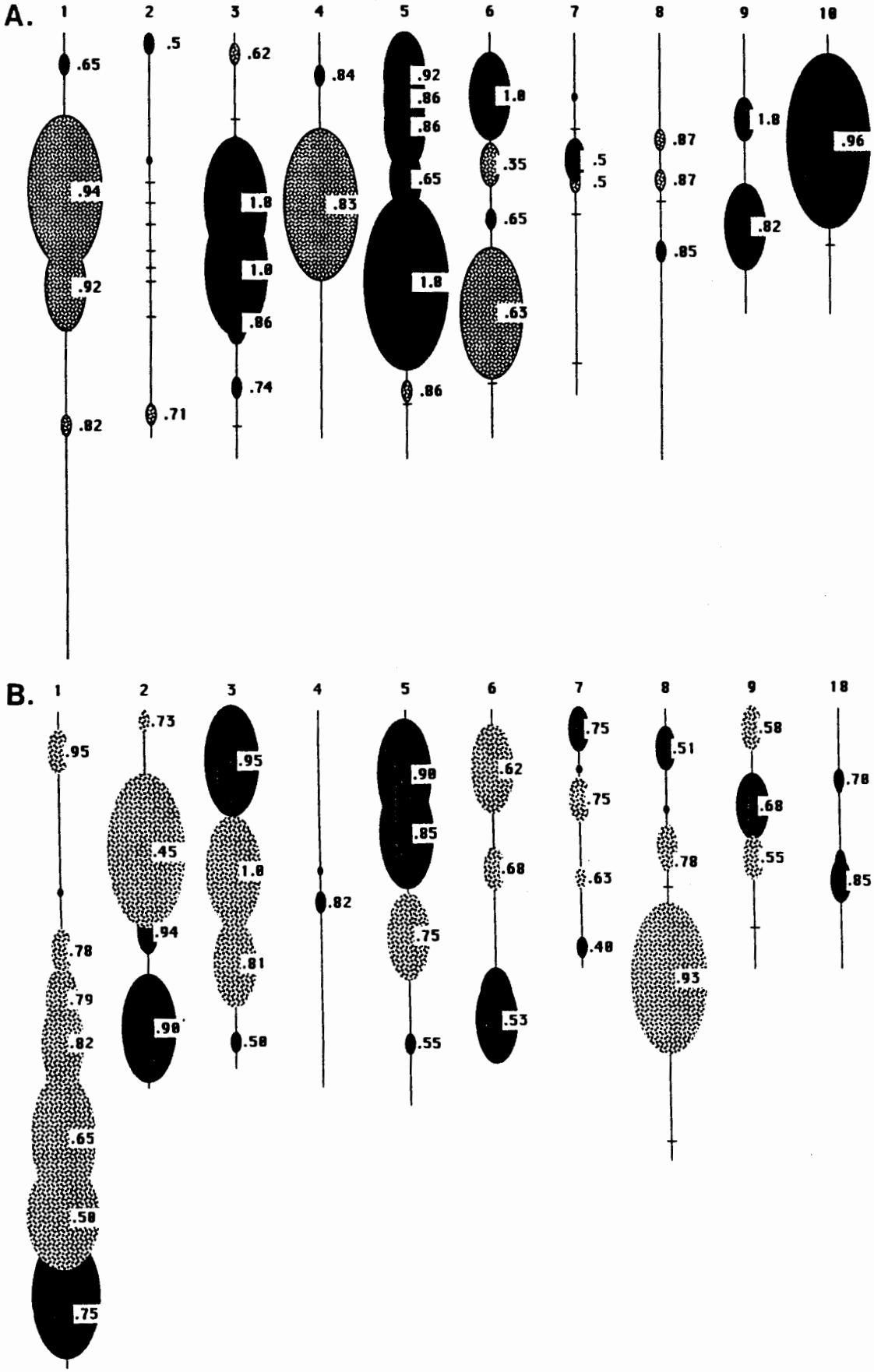


Fig. 3. The chromosome regions that were chosen for manipulation using RFLP marker loci in each of two sweet corn populations, designated A and B, respectively. The "–" marks along the 10 chromosomes in each population indicate the locations of marker loci that were not used in the selection procedure. Ellipses indicate locations of RFLP loci that were judged important in the inheritance of a multi-trait selection index. The width of the ellipse in each region corresponds to the magnitude of the additive effect upon the selection. The type of crosshatching in the ellipses indicates which parental "allele" has the desirable effect—a "dark" pattern indicates the superiority of parent 1 vs. parent 2. The numbers beside each ellipse indicate the final gene frequency of the favorable "allele" among the selected individuals at the RFLP locus at that site. A value of 1.0 indicates complete fixation of the favorable allele; 0.5 represents no change from the F_2 population.

recombination and two cycles of selfing. Thirty-eight selected progeny were testcrossed to the same tester that was used with the marker-selected lines. These provide a benchmark of the progress we might have achieved without using markers in the selection process. It was not possible to conduct another cycle of testcross performance-based selection with these lines within

the time required for three additional cycles of marker-based selection. This is true because the first cycle of random-mating among selected F_4 families produced individuals that were hybrid for the loci at which the F_4 lines differed. A second generation of random-mating was required to generate appreciable segregation. Seed from these progeny were selfed and testcrossed for evaluation in 1993. These progeny were evaluated in the same trial as were the lines developed using marker-based selection. Due to the greater expense of the genotypic characterization of the marker-based selection procedure than was required for the phenotypic selection procedure, different numbers of lines were developed and evaluated using the two methods to keep total costs more comparable for each breeding method.

Gene frequency changes from marker-based selection

The gene frequencies in Fig. 3 A and B providc a measure of the success of the selection procedures in capturing the ideal genotype. They represent the average gene frequency of the favorable allele at each marker locus among the selected lines. In population A (Fig. 3A), most chromosomal regions that appeared to contain QTL of large effect exhibited dramatic changes in gene frequency. Many regions were fixed, or nearly fixed, for the positive genotype. The desired genotype in population B was more difficult to achieve due to more regions for which selection pressure was applied and to more regions with repulsion-phase linkages of desired factors. As a result, smaller changes in gene frequencies were achieved for most chromosome regions. Specific chromosome regions are discussed later.

The observed gene frequency changes attest to the ability of marker-based breeding methods to achieve profound gene frequency changes in a short time. This does not necessarily mean that marker-based breeding will be effective in improving trait performance. These gains also depend on accurate identification of QTL and the ability of indirect selection to achieve gene frequency changes at QTL based on selection at linked, marker loci.

Results of selection in population A

The response to selection in population A is graphically depicted for 11 key traits in Fig. 4. These include some from each category (quality, processing, and agronomic traits) and collectively account for almost 70% of the total selection pressure in the index. They are ordered in Fig. 4 from those that received the greatest weights on the left to those with lesser weights on the right. The responses are expressed as a percentage deviation from the average performance of the parental hybrid checks. A positive response indicates that the average of

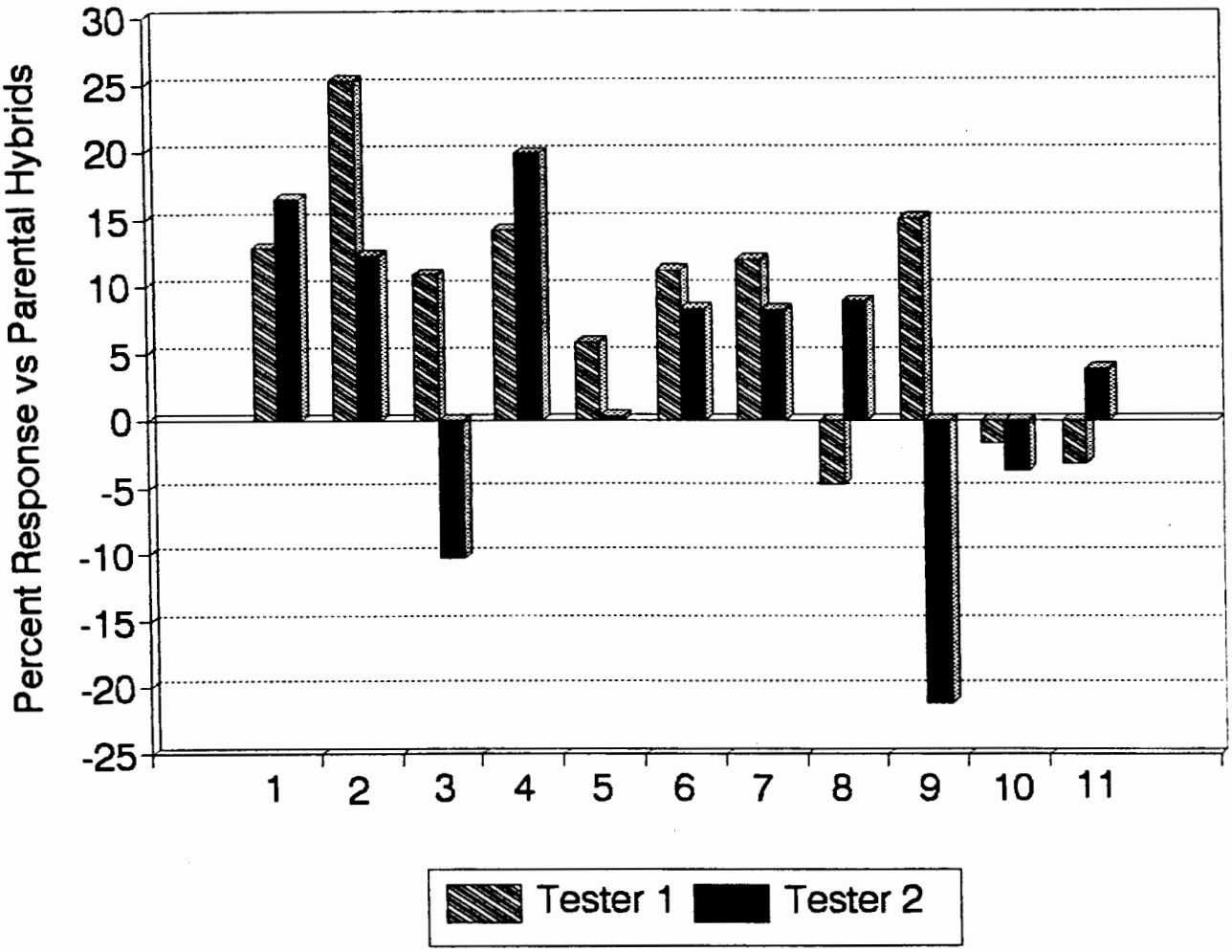


Fig. 4. The average performance of lines (represented as testcross F_1 performance with each of two testers) derived through four cycles of marker-based selection on a selection index. Depicted are performances key component traits that comprised 70% of the weight in the selection index. Traits are ordered from those receiving the greatest weights on the left to those receiving lesser weights on the right. Performances on the Y-axis are scaled to reflect the percent change of the selected group mean from the mean performance of the original parental inbreds when testcrossed by the respective testers. Traits are: 1) overall eating preference; 2) cases of finished product per acre; 3) percent process recovery; 4) tons of raw corn per acre; 5) processed corn color desirability; 6) processed corn sweetness; 7) suitability of ear appearance for corn-on-the-cob (COC); 8) processed corn tenderness; 9) percent of ears eligible for COC packaging; 10) heatunits to harvest; and 11) rating of ear damage during mechanical husking.

the marker-selected hybrid combinations exceeded the estimate of the unimproved population performance. Performances are indicated by pairs of histogram bars for each trait; members of each pair correspond to performances in hybrid combination with testers 1 and 2 on the left and right, respectively. In general, responses were similar in direction, regardless of which of the testers was used to evaluate the selected lines. Exceptions were noted for process recovery and eligibility for corn-on-the-cob, two process traits for which the responses were positive with tester 1 and negative with tester 2, and for tenderness and mechanical husking damage, for which responses were positive with tester 2 and negative with tester 1. Negative responses with both testers were observed with one trait, the number of heat units to flowering.

Consistently positive responses were observed for six of the 11 primary traits shown in Fig. 4, including the two with the greatest weight. The overwhelming direction of response was positive for most traits. The performance index scores of the marker-selected population were substantially improved over the parental hybrid checks with both testers. Net improvement for the weighted index was 0.28 unit with tester 1 and 0.34 unit with tester 2. Because of the way the index was calculated, these correspond to weighted averages of 0.28 to 0.34 phenotypic standard deviation units of response for each trait in the selection index.

The differences between the average performances of the marker-selected hybrids and the parental hybrid checks were not significant for most traits, based on the limited amount of data available from the one planting that could be harvested in 1993. However, the positive directional responses for so many component traits are very encouraging, even if preliminary.

Comparison of marker-based and phenotypic selection responses

This first experiment failed to address the obvious issue of evaluating marker-based selection gains in comparison to gains that may be derived from the same population using traditional breeding methods. Selections in population B were conducted to allow such comparisons. Unfortunately, the initial evaluation of QTL in this population produced a less-optimistic result than was obtained for population A. Many chromosome regions appeared to contain genes, or combinations of genes, that had antagonistic effects on different traits in the selection index. A common observation was for a region to affect yield and quality in an antagonistic manner; the parental genotype that appeared to increase yield also appeared to reduce quality, and vice versa.

There were also regions that were difficult to interpret due to apparent linkages of desirable and undesirable factors. The long arm of chromosome 1 was an example of such a region. Despite fitting numerous multiple-factor models using the GLM Procedure of SAS (SAS Institute, Cary, N.C.), we were unable to establish with confidence a genetic model for this region. This was due, in large part, to an insufficient number of recombination events in this region to represent many of the genotypic classes involved in the complex models. As a result, the weights that were placed on the region effectively held the region in a heterozygous state throughout the selection process (Fig. 3b). Other regions for which great changes in gene frequency did not occur include the long arms of chromosomes 6 and 7 and chromosome 9.

The responses from selection in population B, evaluated based on testcross performance with tester 3, are represented in Figs. 5 and 6. Figure 5 shows the responses for the subset of 12 traits that are performance-related. Positive responses from four cycles of marker-based selection were observed for all traits except heat units to harvest and ear diameter. For eight of the 12 traits, directional changes were positive from cycle 1 to cycle 4 of marker-based selection. Four of the 12 traits exhib-

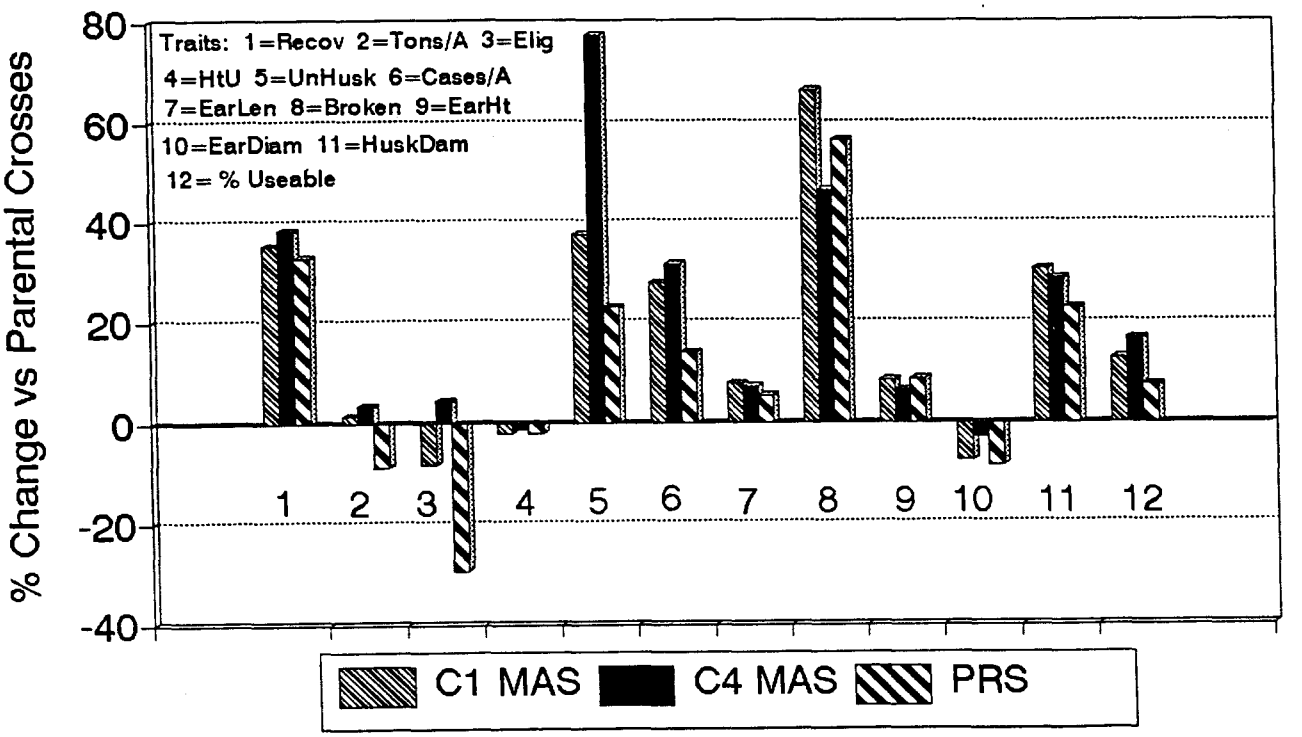


Fig. 5. Comparison of the average responses (in performance-related traits) to marker-based selection (after cycle 1 and after cycle 4) to responses from phenotypic selection during the same period. All selection responses are based on testcross performance using a single, elite inbred tester. Improvements or decreases in performance resulting from selection for an index score are given relative to the performance of testcrosses involving the unimproved parental inbreds. Traits are: 1) percent process recovery; 2) raw product tons/acre; 3) percent COC eligible ears; 4) heatunits to harvest; 5) percent of ears successful in mechanical husking; 6) cases of finished product per acre; 7) ear length; 8) percent unbroken ears during processing; 9) ear height; 10) ear diameter; 11) degree of ear damage during husking; and 12) percent usable ears.

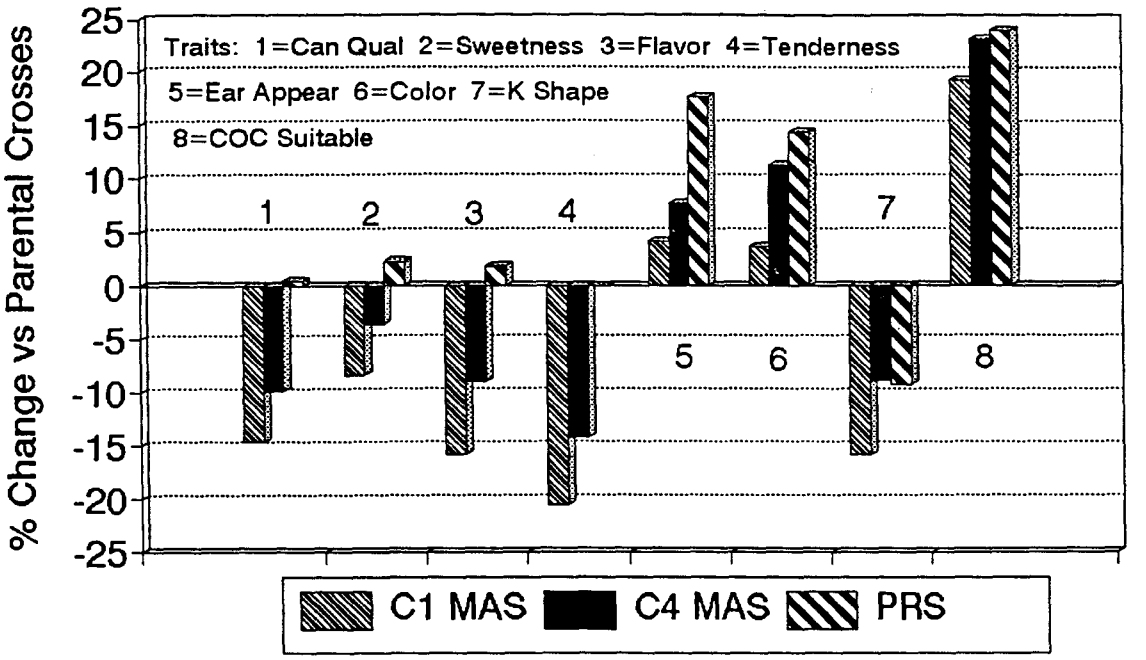


Fig. 6. Comparison of the average responses (in quality related traits) to marker-based selection (after cycle 1 and after cycle 4) to responses from phenotypic selection during the same period. All selection responses are based on testcross performance using a single, elite inbred tester. Improvements or decreases in performance resulting from selection for an index score are given relative to the performance of testcrosses involving the unimproved parental inbreds. Traits are: 1) overall eating preference; 2) sweetness; 3) flavor desirability; 4) tenderness; 5) ear appearance; 6) kernel color desirability; 7) kernel shape; and 8) COC suitability.

ited negative directional responses to phenotypic selection. For only two of the 12 traits was the response from phenotypic selection directionally superior to that of the final, marker-based selection. Both phenotypic and marker-based selection achieved generally positive responses for most traits when compared to the average of the parental hybrid checks.

Responses for the quality related traits were unlike those obtained for performance-related traits used in the selection index, as indicated in Fig. 6. In every case, the response to phenotypic selection exceeded the response to marker-based selection. Gains were achieved from cycle 1 to cycle 4 in every case; however, these were never sufficient to equal the performance of the phenotypically selected population. In five of the eight quality related traits, marker-based selection achieved a negative selection response. This compares to a negative response from phenotypic selection for only one of the eight traits.

The weighted and summed response to the selection index is decomposed into two portions: performance (yield-related) and quality traits, and depicted along with the overall response for the selection index in Fig. 7. The final response from marker-based selection was nearly two-and-a-half times as great as that of phenotypic selection for the performance-related portion of

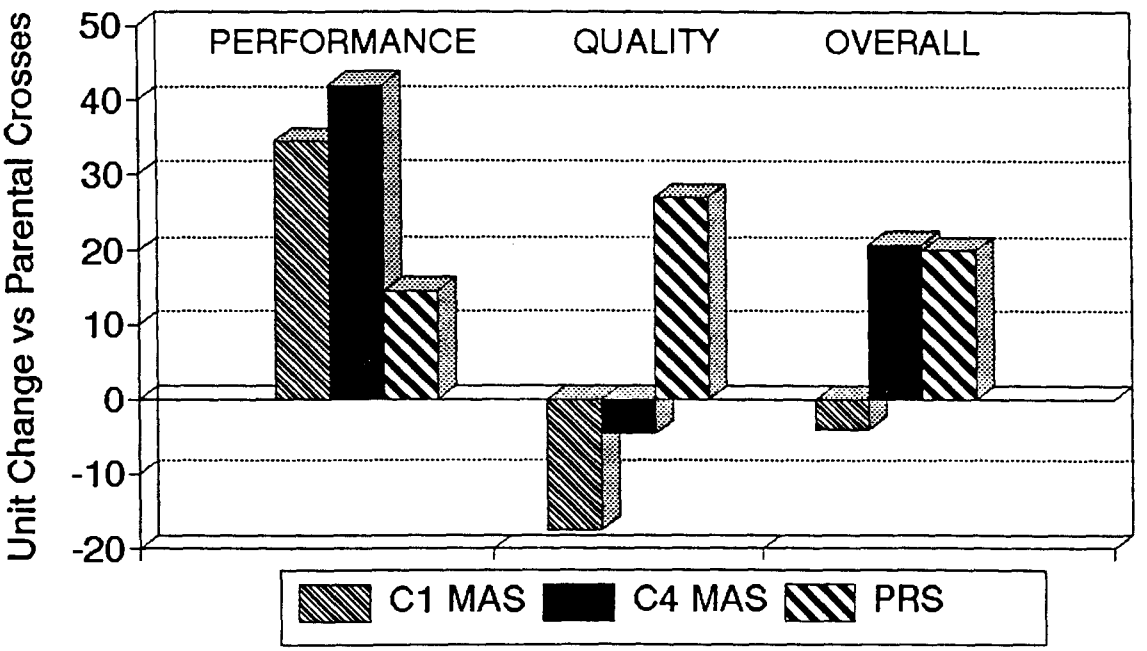


Fig. 7. The average weighted responses in two components of the final selection index and response of the overall selection index performance. The Y-axis is in weighted average phenotypic standard deviation units x 100 (thus, e.g., the average performance-related trait responded to four cycles of MAS, marker-based selection by increasing in value 0.42 standard deviation unit).

the traits. However, the weighted average response from marker-based selection for the quality traits was profoundly poorer than the response from phenotypic selection; it was negative in direction, whereas phenotypic selection produced a positive directional response. The overall response to the selection index was similar for both selection methods due to offsetting effects of performance and quality components.

The explanation for the selection responses in population B is not clear. It is known that yield and quality traits were negatively correlated in the base population that underwent selection. If the negative correlations observed between yield- and quality related traits in this population were due to genetic causes, we would expect gains for yield to tend to be accompanied by decreases in quality, and vice versa. However, the same index was applied to phenotypic measurements as the basis for phenotypic selection as was used in estimates of the genetic effects that were employed in marker-based selection. Therefore, the existence of genetic correlations should have affected the results of both selection methods.

In addition to ignoring genetic covariances among traits, the crude selection index that was applied also failed to consider differential heritabilities of the component traits. Previous studies in our research program have indicated that quality traits tend to have a higher heritability than yield-related traits (data not presented). If this was the case in this population, we might expect phenotypic selection, as exercised here, to have been more effective in changing gene frequencies at loci affecting quality than at those loci affecting yield. This interpretation is consistent with the observation that gains (in normal, standard units) from phenotypic selection for quality related traits were, on average, twice as great as were those for yield-related traits. Selection based on markers may be less sensitive to differences in heritability of component traits because each estimate is based on information from all individuals in the population. This observation is consistent with simulation data from Lande and Thompson (1991), which shows an increasing relative advantage of marker-assisted selection when compared to phenotypic selection as heritability decreased (at a given level of genetic informativeness of markers). Therefore, we might expect marker-based selection to achieve a greater relative response in yield-related traits than in quality related traits when compared to phenotypic selection. This appeared to be the case—response for the performance-related component of the index was about two-and-a-half times as great with marker-based selection as with phenotype-based selection. Nevertheless, we did not expect a negative response in quality from marker-based selection. There did appear to be an improvement in quality from cycle 1 to cycle 4 of marker-based selection, such that the final quality was near the original population performance.

As a practical matter, the results were encouraging because this population already possessed outstanding quality, but, being early in maturity, it lacked the yield performance of full-season sweet corn hybrids. Further data will be required to assign significance to these trends and for more conclusive interpretation of the results obtained in this first year of evaluation. Obviously, the final and most-important measure of success is the derivation of superior, finished hybrids that can replace progenitors in sweet corn production acreage. A determination of this result will require several more years of evaluation.

The data obtained to date from marker-based selection in sweet corn is encouraging. Significant gains in hybrid performance were obtained from selection based only on marker genotypes in two sweet corn populations. The selection appeared, furthermore, to allow simultaneous gains for a number of traits. Many of these traits require evaluation in a processing plant and are difficult and expensive to characterize. Marker-based selection allowed four selection cycles to be conducted within to 2 years rather than the 8 years that would have been required using traditional selection based on testcross performance. Although marker-based selection is relatively expensive, so is the alternative—a traditional breeding program that requires extensive processing of samples. In this respect, markers may be more competitive when used in a horticultural crop for the processing industry or in similar circumstances requiring expensive characterization than they would be for agronomic crops with well-developed and inexpensive evaluation protocols.

Previous simulation results (Edwards and Page, 1994) predict that marker-based selection would be more competitive with traditional breeding methods when used for traits of lower heritability. Markers are not required to make gains for traits of high heritability, and may, in fact, be inferior to phenotypic selection because selection is indirect (i.e., selection is applied on the genotypes of marker loci that are initially linked to QTL, but become dissociated from the QTL due to recombination that accompanies any selection procedure).

The mixed results obtained in population B cause us to question how marker and phenotypic data might be employed optimally, especially when the objective is to select for many traits simultaneously where the traits differ in heritability. If selection is based on genotypes and phenotypes at each selection cycle, one consequence is to slow the rate of genotypic selection to the rate that can be accommodated by phenotypic selection. In the case of selection based on testcross hybrid performance, this slows selection from a rate of two cycles per year to a rate of 2 years per cycle; i.e., by a factor of four. Alternatively, one might use phenotypic selection for improvement of higher-heritability traits at the first cycle (when phenotypes are available, by necessity, in order to support QTL mapping), then base subsequent cycles entirely upon genotypes. Further empirical data are required from this and other investigations to evaluate the merits of marker-based selection as an adjunct to traditional methods or as an exclusive selection procedure for rapid selection during certain phases of crop improvement processes.

REFERENCES

Church, G. and W. Gilbert. 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81:1991–1995.

Dellaporta, S.L., J. Wood, and J.B. Hicks. 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rptr. 1:19–21.

Edwards, M.D. and N.J. Page. 1994. Evaluation of marker-assisted selection through computer simulation. Theor. Applied Genet. (In press).

Feinberg, A.B. and B. Vogelstein. 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6–13.

Lande, R. and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756.

Sprague, G.F. and S.A. Eberhart. 1977. Corn breeding, p. 305–362. In: G.F. Sprague (ed.). Corn and corn improvement. Amer. Soc. Agron., Madison, Wis.

Stuber, C.W. and M.D. Edwards. 1986. Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. Annu. Corn Sorghum Res. Conf. Proc. 41:70–83.

Stuber, C.W. and P.H. Sisco. 1991. Marker-facilitated transfer of QTL alleles between elite inbred lines and responses in hybrids. Annu. Corn Sorghum Res. Conf. Proc. 46:104–113.

Zhang, W. and C. Smith. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. Theor. Applied Genet. 83:813–820.

MPMI Vol. 8, No. 6, 1995, pp. 811–815

Commentary

Genomic Organization of Disease and Insect Resistance Genes in Maize

Michael D. McMullen[1,2] and Kevin D. Simcox[2]

[1]Plant Genetics Research Unit, Agricultural Research Service-USDA, and [2]Plant Sciences Unit, Curtis Hall, University of Missouri, Columbia 65211 U.S.A.

Received 23 May 1995. Accepted 31 July 1995.

The development of molecular marker techniques has enhanced our ability to map loci involved in disease and insect resistance in maize (*Zea mays* L.) and other plant species (Prince and Tanksley 1992). Prior to the advent of molecular marker analysis, mapping resistance genes in maize typically required the use of endosperm marker-linked translocations and chromosome arm tester stocks (Burnham 1982; Louie et al. 1991). These specialized stocks provided only partial coverage of the genome and, in many cases, were present in inappropriate genetic backgrounds for mapping of disease and insect resistance traits. Restriction fragment length polymorphism (RFLP) analysis alleviates the need to use cytological and morphological markers to map genes, and, in maize, virtually any combination of parental lines can be used for RFLP analysis (Coe et al. 1988; Hoisington and Coe 1989). Over the past 5 years, there has been a large increase in the number of qualitative genes and quantitative trait loci (QTL) mapped in maize for plant response to pests and pathogens. It is becoming apparent, as more genes for disease response traits are placed on the maize linkage map, that loci for disease and insect resistance are not randomly distributed over the maize genome.

Our laboratory has been developing a high resolution genetic map in the region of the maize dwarf mosaic virus (MDMV) resistance gene, *mdm1*, on the short arm of chromosome 6 (Simcox et al. 1995). We discovered that there were two additional resistance genes that map near *mdm1*: (i) *wsm1*, which confers dominant resistance to a related potyvirus, wheat streak mosaic virus (WSMV) (McMullen and Louie 1991; McMullen et al. 1994); and (ii) a recessive gene, *rhm1*, which confers resistance to the fungal pathogen *Cochliobolus heterostrophus* (Drechs.) Drechs. race O (Zaitlin et al. 1993). A search of the literature was undertaken to determine to what extent genes for disease and insect response traits were linked or "clustered" in the maize genome.

Descriptions of various disease and insect traits are listed in Table 1. Table 2 summarizes the map locations of disease and insect resistance genes and QTL reported in the literature. Map positions are reported in the form of "chromosomal bin" locations taken from the 1995 UMC maize RFLP linkage map, generated by the USDA-ARS Plant Genetics Research Unit and the Plant Science Unit at the University of Missouri. For the bin map, each chromosome is divided into approximately equal segments of about 20 cM. The boundaries of the bins are fixed by "core" markers (Gardiner et al. 1993). Map positions given for QTL correspond to the most significant RFLP marker associated with that trait. All information presented in this article can be accessed through the USDA maize genome database, MaizeDB. Information for accessing MaizeDB can be obtained by contacting the MaizeDB curator, Dr. Mary Polacco, at the University of Missouri (maryp@teosinte.agron.missouri.edu).

Corresponding author: Michael D. McMullen;
E-mail: mcmullen@teosinte.agron.missouri.edu

Clustering of disease and insect resistance genes.

All 10 of the maize linkage groups contain either disease/insect resistance genes or QTL (Table 2). The majority of the resistance loci and QTL are located in chromosomal bins containing other disease or insect resistance factors. A chi-square test of independence of gene distribution within bins was performed after pooling adjacent bins to increase the expected class numbers to greater than 1.0 (Snedecor and Cochran 1967). The probability the observed distribution could be obtained by chance was $P < 0.001$. The distribution obtained did not fit a random model. These clusters of resistance genes occurred on each linkage group with the exception of chromosomes 7 and 9. The cM distance between loci within a particular cluster varied. In some instances, such as the gene clusters present in bins 3.04 and 6.01, loci were tightly linked, whereas resistance genes in other clusters were distributed over 20- to 40-cM regions. This variation in the size of the clusters may reflect the actual position of the loci, or be due to differences in experimental procedures used in the different mapping studies.

Do resistance genes cluster to a greater extent than other genes? The distribution of expressed sequences in the maize genome has not been adequately addressed. Certainly clusters of gene families such as the zeins, rDNA genes, and P-450 monoxgenases have been identified. Recently, Khavkin and Coe (1995a, 1995b) have reported that single genes and QTL for morphological traits and growth regulation are clustered in the maize genome.

Table 1. Description of pathogens and pests

Disease/insect	Causal organism	Symptom[a]	Vector	Reference
Maize dwarf mosaic virus	Potyvirus	S, Mo	Many aphid species	Simcox et al. 1995
Wheat streak mosaic virus	Potyvirus	S, Mo	*Eriophyes tulipae* (Keifer) mite	McMullen et al. 1994
Maize mosaic virus	Rhabdovirus	S, Mo	*Peregrinus maidis* (Ashmead) leafhopper	Ming et al. 1995b
Maize streak virus		S	*Cicadulina mbila* Naude leafhopper	Kyetere et al. 1995
Stewart's wilt	*Erwinia stewartii*	F, V, Ne	*Chaetocnema pulicaia* corn flea beetle	Ming et al. 1995a
Carbonum leaf spot	*Cochliobolus carbonum* Nelson	F, M, Ne		Coe et al. 1988
Gray leaf spot	*Cercospora zeae-maydis* Theon and Daniels	F, M, Ne		Bubeck et al. 1993
Southern corn leaf blight	*Cochliobolus heterostrophus* (Drechs.) Drechs.	F, M, Ne		Zaitlin et al. 1993
Northern corn leaf blight	*Setosphaeria turcica* (Luttrell) K. J. Leonard & E. G. Suggs	F, V, Ne		Simcox and Bennetzen 1993 Zaitlin et al. 1992 Bentolila et al. 1991 Hoisington and Coe 1989
Common rust	*Puccinia sorghi* (Schwein.)	F		Coe et al. 1988
Southern rust	*Puccinia polysora* Underw.	F		Coe et al. 1988
Anthracnose stalk rot	*Collectotrichum graminicola* (Ces.) Wils.	N		Jung et al. 1994
Fusarium stalk rot	*Gibberella zeae* (Schwein.) Petch	N		Pè et al. 1993
European corn borer	*Ostrinia nubilalis* Hübner	F, N		Schön et al. 1993
Corn earworm	*Helicoverpa zea* (Boddie)	E		Byrne et al. 1995

[a] E = mature ear, F = foliar, M = mesophyll cell layer, Mo = mosaic symptoms, N = stalk internode invasion, Ne = necrosis, R = root infection, S = systemic infection, V = vascular wilt

Table 2. Chromosomal bin locations of disease and insect resistance genes and quantitative trait loci (QTL)[a]

		Chromosome									
Disease/insect trait	**Locus**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Northern corn leaf blight	*ht1*		2.08[b]								
	ht2								8.06		
	htn1								8.06		
	QTL	1.01/2		3.07/8	4.02/3	5.01/2 5.06		7.03	8.03/4		
DIMBOA	*bx1*				4.02/3						
European corn borer	QTL	1.01 1.07	2.03 2.08/9	3.04/5	4.02/3			7.04			10.04/5
Corn earworm	QTL	1.03								9.01/2	10.06
Gray leaf spot	QTL	1.04	2.04/5		4.02 4.04 4.08				8.05		10.05
Anthracnose stalk rot	QTL				4.08						
Maize streak	*msv1*	1.04									
Stewart's wilt	QTL	1.05									
Carbonum leaf spot	*hm1*	1.07									
	hm2									9.04/5	
Fusarium stalk rot	QTL	1.07	2.04	3.04/5	4.04	5.02 5.04					10.06
Common rust	*rp3*			3.04							
	rp4				4.02/3						
	rp1										10.01
	rp5										10.01
	rp1-G										10.01
Southern rust	*rpp9*										10.01
Maize mosaic	*mv1*			3.04							
Wheat streak mosaic	*wsm1*						6.01				
	wsm2			3.04							
	wsm3										10.05
Maize dwarf mosaic	*mdm1*						6.01				
Southern corn leaf blight	*rhm1*						6.01				

[a] Map positions and scores of restriction fragment length polymorphism loci can be accessed through the maize genome database, MaizeDB. Information concerning access to MaizeDB can be obtained by contacting the MaizeDB curator, Dr. Mary Polacco, at the University of Missouri (maryp@teosinte.agron.missouri.edu).

[b] Bin locations are designated by an X.Y code, where X is the linkage group containing the bin and Y is the location of the bin within the linkage group.

An interesting aspect of these clusters is the spectrum of pathogens/pests and host-pathogen/pest interactions represented (Table 1). For example, near the centromere on chromosome 3 is a tight cluster of resistance genes and QTL, located within bins 3.04 and 3.05 (Table 2). This cluster contains three dominant disease resistance genes mapping within 5 cM of *umc102*: (i) the *rp3* locus, which confers resistance to *Puccina sorghi* (Schwein.) (Sanz-Alferez et al. 1995); (ii) the *wsm2* locus for resistance to WSMV (McMullen et al. 1994); and (iii) *mv1* for resistance to maize mosaic virus (MMV) (Ming et al. 1995b). Also mapping in bin 3.04 were QTL for resistance to *Fusarium* stalk rot (FSR), caused by *Gibberella zeae* (Schwein.) Petch, and the European corn borer (ECB), *Ostrinia nubilalis* Hübner (Pè et al. 1993; Schön et al. 1993). The *rp3* locus appears to be involved in the pathogen recognition pathway; infection by an avirulent race of *P. sorghi* will result in a localized hypersensitive response (HR) (Bennetzen et al. 1988). Both *wsm2* and *mv1* confer resistance to viral pathogens. While the mechanisms of resistance to these unrelated viruses are currently unknown, neither resistance is characterized by an HR similar to *rp3*. The QTL for ECB resistance in bin 3.04 is associated with degree of tunnel length caused by second generation larvae (Schön et al. 1993), and the FSR QTL is associated with percent stalk internodes infected (Pè et al. 1993). The association between resistance to ECB tunneling and FSR resistance is interesting since ECB tunneling provides entry points for stalk-rotting pathogens. However, the association between QTL for resistance to ECB and FSR is not universal, since only three of the seven ECB QTL identified by Schön and co-workers mapped to the same region as a QTL for FSR identified by Pè and co-workers. Although *O. nubilalis* and *G. zeae* share part of their life cycle within the stalk, the interactions between maize and these organisms are very different. Each interaction between maize and a pathogen or pest will have aspects unique to the invading organism.

Functional relationships.

The question of whether resistance genes within a cluster are functionally related is unclear, particularly because resistance to unrelated organisms is often found within a cluster. Except for a few cases, the biochemical and physiological bases of resistance to pathogens and pests have yet to be elucidated in maize. The only example of a disease resistance gene having been cloned in maize is the *hm1* locus, which confers resistance to carbonum leaf spot caused by *Cochliobolus carbonum* Nelson race 1 (Johal and Briggs 1992). The *Hm1* allele encodes a functional NADP reductase that detoxifies the HC-toxin, a fungal pathotoxin required by *C. carbonum* race 1 for disease development.

Another example of a known maize defense compound is the involvement of the flavone glycoside, maysin, in silk-feeding resistance to the corn earworm, *Helicoverpa zea* (Boddie) (Byrne et al. 1995). Resistance to corn earworm is a quantitative trait, with maysin synthesized via the flavonoid metabolic pathway (Coe et al. 1988). This pathway is well characterized in maize and efforts are currently underway in our laboratory to equate QTL for corn earworm resistance to specific steps in the pathway. The major QTL for maysin concentration corresponds to the transcription regulator gene *p1* located in bin 1.03 near the gene for resistance to maize streak virus, *msv1*, and an FSR QTL (Table 2).

An additional preformed antimicrobial compound—2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)—has been identified in maize and is associated with resistance to feeding by the first-generation ECB and correlated with northern corn leaf blight (NCLB) lesion expansion, caused by *Setosphaeria turcica* (Luttrell) K. J. Leonard & E. G. Suggs (Coulture et al. 1971; Guthrie et al. 1986). The accumulation of DIMBOA is regulated by the action of the *bx1* locus, which is located in bin 4.02 or 4.03 on the short arm of chromosome 4 (Table 2; Simcox and Weber 1985). This region also contains a dominant resistance gene to *P. sorghi*, *rp4*, and QTL for resistance to first brood larvae of ECB, gray leaf spot (GLS) caused by *Cercosprora zeae-maydis* Theon and Daniels, and NCLB (Schön et al. 1993; Bubeck et al. 1993; Freymark et al. 1993). It would seem unlikely that a preformed antimicrobial compound, such as DIMBOA, would be a primary factor in an *rp4*-mediated HR. However, the presence of a phytotoxic and antimicrobial compound might affect QTL for lesion size and sporulation, which are quantitative parameters for resistance to NCLB and GLS.

Genome duplications in the region of disease resistance loci.

The genome of maize is highly duplicated, with as many as 40% of the genomic RFLP and cDNA probes detecting more than one unlinked locus under high strigency hybridization conditions (Helentjaris 1995; Helentjaris et al. 1988). Furthermore, the existence of duplicate arrays of colinear loci, identified by the same RFLP probe, indicates whole segments of chromosomes are likely duplicated. Examination of the chromosomal position of disease resistance genes and QTL within duplication regions reveals that many resistance loci may be related through duplication. An example is the region around *hm1*, bins 1.06 to 1.07, which is duplicated on the long arm of chromosome 9, near 9.04 to 9.05. The chromosome 9 region contains the *hm2* locus that is involved in adult plant resistance to *C. carbonum* race 1 (Coe et al. 1988; Johal and Briggs 1992). The gene product of the *hm2* locus has yet to be determined, but the *Hm2* allele confers resistance against the same HC-toxin-producing *C. carbonum* isolates as *Hm1*.

On the long arm of chromosome 8, within bin 8.06, are two dominant NCLB resistance genes, *ht2* and *htn1* (Table 1; Simcox and Bennetzen 1993; Zaitlin et al. 1992). The *Ht2* allele confers a chlorotic-necrotic lesion response to infection by an avirulent race of *S. turcica*, analogous to an HR, which inhibits or limits sporulation. On the other hand, the *Htn1* allele results in delay of symptom expression when challenged with an avirulent race. Reduction of sporulation and inhibition of lesion development are two of the parameters that are used to measure quantitative resistance to NCLB (Freymark et al. 1993, 1994). It is interesting to note that when chromosomal regions containing homoeology to bin 8.05 (Helentjaris 1995) are examined, NCLB QTL are also present. RFLP probes mapping to bin 8.05 also detected loci mapping to chromosome 3, bin 3.07 to 3.08, and chromosome 5, bin 5.06 to 5.07. Both of these chromosome regions contain NCLB QTL (Freymark et al. 1993). It therefore appeared that these three genes may represent duplications of the same gene or represent genes derived from an ancient cluster (Helentjaris 1993).

Origin and implication of resistance gene clusters.

Recent reports describing cloning of disease resistance genes conferring HR resistance from *Arabidopsis thaliana*, tobacco, and tomato suggest that resistance genes involved in the recognition of an invading organism represent steps in the signal transduction pathway (Bent et al. 1994; Jones et al. 1994; Martin et al. 1993; Mindrinos et al. 1994; Whitman et al. 1994). There are often linked, duplicated sequences related to the resistance genes (Martin et al. 1994; Sanz-Alferez et al. 1995; Whitman et al. 1994). If these duplicated sequences also function in resistance to other pathotypes of the same pathogen, or other pathogens, a basis for resistance gene clustering for HR genes can be envisioned. The *rp1* complex on the short arm of chromosome 10 consists of duplicated genetic elements that encode alleles for resistance to *P. sorghi* (Bennetzen and Hulbert 1993; Sudapak et al. 1993). Unequal exchange and gene conversion events have been implicated in the instability of resistance at this complex, but also in the evolution of new alleles with altered race specificites (Hu and Hulbert 1994; Scot Hulbert, personal communication). In the case of resistance to rapidly evolving pathogens, the ability of a resistance gene cluster to evolve new specificities might be influenced by the chromosomal location. Distinct regions of the genome might evolve at a pace sufficient to coevolve with changing pathogen/pest challenges.

However, the diversity of organisms and resistance responses of loci mapping within a particular cluster probably rule out the possibility that all genes in clusters are involved in signal transduction. The *bx1* locus controlling the synthesis of DIMBOA and the *p1* locus controlling the synthesis of maysin may be clues to a second class of genes that may lead to gene clustering through duplication. Regulation of the anthocyanin pathway is controlled by a set of *myb*-like (*c1*, *p1*, and *p11*) and *myc*-like (*b1* and *r1*) transcription regulators (Grotewold et al. 1994). Duplicated elements are often associated with these loci. The involvment of the *p1* locus in the synthesis of maysin has already been discussed, but the action of the *bx1* gene is also consistent with this locus acting as a transcription regulator for genes in the DIMBOA biosynthetic pathway. If other QTL or single genes for resistance are transcription regulators then the resistance genes could be structurally related, such as *myb*-like or *myc*-like, but control biochemical pathways for compounds effective against very diverse pathogens or pests.

Applications in gene mapping and cloning.

Clustering of disease resistance genes suggests several applications to enhance the efficiency of cloning disease and pest-resistance genes in maize. Maize contains several active transposable element systems that have been used successfully to transposon tag and clone genes (Johal and Briggs 1992). One useful characteristic of these transposable element systems is propensity of active elements to transpose more frequently to linked sites than to more distant chromosome regions (Dellaporta and Moreno 1994). This property of maize transposable elements has recently been successfully used to clone disease resistance genes in both tobacco and tomato using the maize *Activator* transposable element (Jones et al. 1994; Whitman et al. 1994). Chromosomal regions containing clusters of resistance genes in maize could be more efficiently targeted for transposon tagging by preselecting active transposable element populations that contain transposable elements linked to the target region (Chang and Peterson 1994).

A second application of the phenomena of resistance gene clustering is the construction of high resolution genetic maps in the region of disease resistance gene clusters, which can then be used in positional cloning of a number of disease resistance traits (Martin et al. 1993; Simcox et al. 1995). Positional cloning in maize is at best a difficult undertaking, due mainly to high amounts of repetitive DNA (Springer et al. 1994). However, the realization that the genomes of maize and related grasses are colinear, despite the occurrence of translocations and inversions, has led to novel approaches for positional cloning (Ahn and Tanksley 1993; Hulbert et al. 1990). This colinearity of grass genomes has led to the premise of the cereals as a single experimental system allowing the exchange of tools and techniques of genome analysis between grass species (Bennetzen and Freeling 1993; Helentjaris 1993). Difficulties in positional cloning in maize could be overcome by utilizing high molecular weight libraries from smaller genomes that are somewhat colinear, such as rice and sorghum (Bennetzen and Freeling 1993). Establishment of colinear relationships of resistance genes across species will greatly aid gene isolation and understanding the actions of genes for resistance.

ACKNOWLEDGMENTS

We would like to thank Jeff Bennetzen, Tim Helentjaris, and Dave Zaitlin for permission to cite unpublished data, Jeff Bennetzen and Scot Hulbert for critical appraisal of the manuscript, and Larry Darrah and Gary Krause for assistance with statistical analysis of data. This was a cooperative investigation of the USDA-ARS and the Plant Sciences Unit, University of Missouri, and was submitted as Missouri Agricultural Experimental Station Journal Series #12,330.

LITERATURE CITED

Ahn, S., and Tanksley, S. D. 1993. Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. USA 90:7980-7984.

Bennetzen, J. L., Blevins, W. E., and Ellingboe, A. H. 1988. Cell-autonomous recognition of the rust pathogen determines *Rp1*-specified resistance in maize. Science 241:208-209.

Bennetzen, J. L., and Freeling, M. 1993. Grasses as a single genetic system: genome composition, collinearity and compatibility. Trends Genet. 9:259-261.

Bennetzen, J. L., and Hulbert, S. H. 1993. Organization, instability and evolution of plant disease resistance genes. Plant Mol. Biol. 20:575-578.

Bent, A. F., Kunkel, B. N., Dahlbeck, D., Brown, K. L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B. J. 1994. *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. Science 265:1856-1860.

Bentolila, S., Guitton, C., Bouvet, N., Sailland, A., Nykaza, S., and Freyssient, G. 1991. Identification of an RFLP marker tightly linked to the *Ht1* gene in maize. Theor. Appl. Genet. 82:393-398.

Bubeck, D. M., Goodman, M. M., Beavis, W. D., and Grant, D. 1993. Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci. 33:838-847.

Burnham, C. R. 1982. The location of genes to chromosome by the use of chromosomal interchanges. Pages 65-70 in: Maize for Biological Research. W. F. Sheridan, ed. Plant Molecular Biology Association, Charlottesville, VA.

Byrne, P. F., McMullen, M. D., Snook, M. E., Musket, T., Widstrom, N. W., and Wiseman, B. R. 1995. Quantitative genetic analysis of loci controlling synthesis of maysin, a corn earworm resistance factor, in maize silks. Maize Genet. Coop. Newsl. 69:53-54.

Chang, R.-Y., and Peterson, P. A. 1994. Chromosome labeling with

transposable elements in maize. Theor. Appl. Genet. 87:650-656.
Coe, E. H., Jr., Neuffer, M. G., and Hoisington, D. A. 1988. The Genetics of Corn. Pages 81-258 in: Corn and Corn Improvement. G. F. Sprague and J. W. Dudley, eds. Crop Science Society of America, Inc., Madison, WI.
Coulture, R. M., Routley, D. G., and Dunn, G. M. 1971. Role of cyclic hydroxamic acids in monogenic resistance of maize to *Helminthosporium turcicum*. Physiol. Plant Pathol. 1:515-521.
Dellaporta, S. L., and Moreno, M. 1994. Gene tagging with Ac-Ds elements in maize. Pages 219-233 in: The Maize Handbook. M. Freeling and V. Walbot, eds. Springer Verlag, New York.
Freymark, P. J., Lee, M., Martinson, C. A., and Woodman, W. L. 1994. Molecular-marker-facilitated investigation of host-plant response to *Exserohilum turcicum* in maize (*Zea mays* L.): Components of resistance. Theor. Appl. Genet. 88:305-313.
Freymark, P. J., Lee, M., Woodman, W. L., and Martinson, C. A. 1993. Quantitative and qualitative trait loci affecting host-plant response to *Exserohilum turcicum* in maize (*Zea mays* L.). Theor. Appl. Genet. 87:537-544.
Gardiner, J., Melia-Hancock, S., Hoisington, D. A., Chao, S., and Coe, E. H. 1993. Development of a core RFLP map in maize using an immortalized-F2 population. Genetics 134:917-930.
Grotewold, E., Drummond, B. J., Bowen, B., and Peterson, T. 1994. The *myb*-homologous *P* gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. Cell 76:543-553.
Guthrie, W. D., Wilson, R. L., Coats, J. R., Robbins, J. C., Tseng, C. T., Jarvis, J. L., and Russell, W. A. 1986. European corn borer (*Lepidoptera: Pyralidae*) leaf-feeding resistance and DIMBOA content in inbred lines of dent maize grown under field versus greenhouse conditions. J. Econ. Entomol. 79:1492-1496.
Helentjaris, T. 1993. Implications for conserved genomic structure among plant species. Proc. Natl. Acad. Sci. USA 90:8308-8309.
Helentjaris, T. 1995. Atlas of duplicated sequences in maize. Maize Genet. Coop. Newsl. 69:67-81.
Helentjaris, T., Weber, D., and Wright, S. 1988. Identification of the genomic locations of duplicated nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. Genetics 118:353-363.
Hoisington, D. A., and Coe, E. H., Jr. 1989. Methods for correlating RFLP maps with conventional genetic and physical maps in maize. Pages 19-24 in: Development and Application of Molecular Markers to Problems in Plant Genetics. T. Helentjaris and B. Burr, eds. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Hu, G., and Hulbert, S. 1994. Evidence for the involvment of gene conversion in meiotic instability of the *Rp1* rust resistance genes of maize. Genome 37:742-746.
Hulbert, S. H., Richter, T. E., Axtell, J. D., and Bennetzen, J. L. 1990. Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proc. Natl. Acad. Sci. USA 87:4251-4255.
Johal, G. S., and Briggs, S. P. 1992. Reductase activity encoded by the *HM1* disease resistance gene in maize. Science 258:985-987.
Jones, D. A., Thomas, C. M., Hammond-Kosack, K. E., Balint-Kurti, P. J., and Jones, J. D. G. 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266:789-793.
Jung, M., Weldekidan, T., Schaff, D., Paterson, A., Tingey, S., and Hawk, J. 1994. Generation-means analysis and quantitative trait locus mapping of anthracnose stalk rot genes in maize. Theor. Appl. Genet. 89:413-418.
Khavkin, E. E., and Coe, E. H. 1995a. Organization of growth-regulating genes in maize: 1. The functional clusters of genes. Russ. J. Plant Physiol. 42:408-420.
Khavkin, E. E., and Coe, E. H. 1995b. Organization of growth-regulating genes in maize: 2. Quantitative trait loci. Russ. J. Plant Physiol. 42:558-574.
Kyetere, D., Ming, R., McMullen, M., Pratt, R., Brewbaker, J., Musket, T., Pixley, K., and Moon, H. 1995. Monogenic resistance to maize streak virus maps to the short arm of chromosome 1. Maize Genet. Coop. Newsl. 69:136-137.
Louie, R., Findley, W. R., Knoke, J. K., and McMullen, M. D. 1991. Genetic basis of resistance in maize to five maize dwarf mosaic virus strains. Crop Sci. 31:14-18.
Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Frary, A., Ganal, M. W., Spivey, R., Wu, T., Earle, E. D., and Tanksley, S. D. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432-1436.
Martin, G. B., Frary, A., Wu, T., Brommonschenkel, S., Chunwongse, J., Earle, E. D., and Tanksley, S. D. 1994. A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. Plant Cell 6:1543-1552.
McMullen, M. D., Jones, M. W., Simcox, K. D., and Louie, R. 1994. Three genetic loci control resistance to wheat streak mosaic virus in the maize inbred Pa405. Mol. Plant-Microbe Interact. 7:708-712.
McMullen, M. D., and Louie, R. 1991. Identification of a gene for resistance to wheat streak mosaic virus in maize. Phytopathology 81:624-627.
Mindrinos, M., Katagiri, F., Yu, G.-L., and Ausubel, F. M. 1994. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucelotide-binding site and leucine-rich repeats. Cell 78:1089-1099.
Ming, R., McMullen, M. D., Brewbaker J. L., Pratt, R. C., Moon, H. G., Musket, T., and Holley, R. 1995a. RFLP mapping of genes conferring resistance to *Erwinia stewartii*. Maize Genet. Coop. Newsl. 69:60.
Ming, R., McMullen, M. D., Brewbaker J. L., Pratt, R. C., Moon, H. G., Musket, T., and Kyetere, D. T. 1995b. RFLP mapping of maize mosaic virus resistance gene. Maize Genet. Coop. Newsl. 69:60-61.
Pè, M. E., Gianfranceschi, L., Taramino, G., Tarchini, R., Angelini, P., Dani, M., and Binelli, G. 1993. Mapping quantitatvie trait loci (QTL) for resistance to *Gibberella zeae* infection in maize. Mol. Gen. Genet. 241:11-16.
Prince, J. P., and Tanksley, S. D. 1992. Restriction fragment length polymorphism in plant breeding and genetics. Proc. Royal Soc. Edin. 99B:23-29.
Sanz-Alferez, S., Richter, T. E., Hulbert, S. H., and Bennetzen, J. L. 1995. The *Rp3* disease resistance gene of maize: mapping and characterization of introgressed alleles. Theor. Appl. Genet. 91:25-32.
Schön, C. C., Lee, M., Melchinger, A. E., Guthrie, W. D., and Woodman, W. L. 1993. Mapping and characterization of quantitative trait loci affecting resistance against second-generation European corn borer in maize with the aid of RFLPs. Heredity 70:648-659.
Simcox, K. D., and Bennetzen, J. L. 1993. The use of molecular markers to study *Setosphaeria turcica* resistance in maize. Phytopathology 83:1326-1330.
Simcox, K. D., McMullen, M. D., and Louie, R. 1995. Co-segregation of the maize dwarf mosaic virus resistance gene, *mdm1*, with the nucleolus organizer region in maize. Theor. Appl. Genet. 90:341-346.
Simcox, K. D., and Weber, D. F. 1985. Location of the benzoxazinless (*bx*) locus in maize by monosomic and B-A translocational analyses. Crop Sci. 25:827-830.
Snedecor, G. W., and Cochran, W. G. 1967. Statistical Methods. The Iowa State University Press, Ames. p. 235.
Springer, P., Edwards, K., and Bennetzen, J. L. 1994. DNA class organization on maize *Adh1* yeast artifical chromosomes. Proc. Natl. Acad. Sci. USA 91:863-867.
Sudupak, M. A., Bennetzen, J. L., and Hulbert, S. H. 1993. Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region of maize. Genetics 133:119-125.
Whitman, S., Dinesh-Kumar, S. P., Choi, D., Hehl, R., Corr, C., and Baker, B. 1994. The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. Cell 78:1101-1115.
Zaitlin, D., DeMars, S., and Gupta, M. 1992. Linkage of a second gene for NCLB resistance to molecular markers in maize. Maize Genet. Coop. Newsl. 66:69-70.
Zaitlin, D., DeMars, S., and Ma, Y. 1993. Linkage of *rhm*, a recessive gene for resistance to southern corn leaf blight, to RFLP marker loci in maize (*Zea mays*) seedlings. Genome 36:555-564.

Molecular technologies for biodiversity evaluation: Opportunities and challenges

New technologies for detecting variation in DNA complement traditional methods in biodiversity.

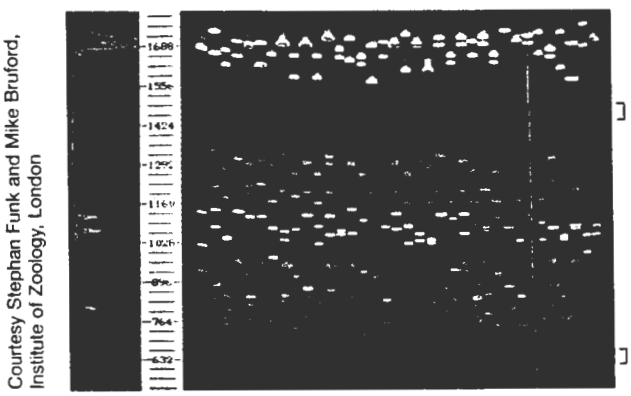
Angela Karp[1], Keith J. Edwards[1], Mike Bruford[2], Stephan Funk[2], Ben Vosman[3], Michele Morgante[4], Ole Seberg[5], Antoine Kremer[6], Pierre Boursot[7], Peter Arctander[8], Diethard Tautz[9], and Godfrey M. Hewitt[10]

Better information on the degree and distribution of genetic variation is essential for developing more efficient ways of evaluating and conserving biodiversity. At present, an array of molecular techniques is available to detect diversity at the DNA level[1], but the application of these techniques—so that they provide useful information and not simply data—depends critically on the analysis method employed (see "Analytical tools for molecular data"). In general, questions of genetic diversity can be addressed at the species, population, and within-population levels[2].

The species level

The identification of taxonomic units and the determination of the uniqueness of species is essential information for conservation. Questions at this level include: Does a particular isolate represent a species, subspecies, or race? Is it a hybrid? If it is a species, how unique is it? Molecular techniques are potentially relevant to all these questions. They can provide information that helps in defining the distinctiveness of species and their ranking according to the number of close relatives and their phylogenetic position[3]. Molecular markers also have much to offer to the resolution of problems concerning hybridization and polyploidy. Sequence data provide the most accurate information for questions of this type, as sequences are the only molecular markers that contain a record of their own history. In addition to revealing the groupings of individuals into different classes, appropriate analyses based on sequence data (or restriction site data) can provide hypotheses on the evolutionary relationships between the different categories. One important caveat regarding the interpretation of such data is that the information it provides relates to the evolutionary history of the sequence (gene) in question, which may be separate from that of the organism carrying it. A straightforward, but time-consuming way to avoid this difficulty would be to collect information on the genealogies of many independent sequences. Fortunately, studies so far suggest that data from mitochondrial (mt) DNA analysis, and 1–2 nuclear sequences from critical taxa, may suffice, as most species comparisons reveal quite high levels of divergence.

Although arbitrary, semiarbitrary, and other multilocus profiling techniques have been (and are) used to provide information for answering questions at the species level, we would argue strongly against this because of data limitations in allelic assignment, dominance and homology. In principle, these limitations are not insurmountable, provided that sufficient preliminary pedigree analysis is carried out to determine independence and mode of inheritance, and that sample sizes are large enough[4]. But in many biodiversity studies this is not possible due to sampling problems or financial and time constraints. Sequence tagged microsatellites (STMS) and minisatellites, in contrast, constitute a single locus with (usually) many different, codominant alleles. Identity and assignment of alleles is thus not a problem. Their high mutation rate does mean, however, that the accuracy with which true homology can be inferred for different genotypes becomes questionable over large genetic distances because of the increasing possibility of homoplasy. Although the presence or absence of a given STMS locus can be used as phylogenetic information, it is otherwise difficult to envisage the use of STMS in the reconstruction of phylogenies.



Courtesy Stephan Funk and Mike Bruford, Institute of Zoology, London

Figure 1. Automated single sequence repeat (SSR) genotyping allows rapid DNA fingerprinting of organisms. Such markers are highly informative for characterizing plant and animal genetic resources.

The population level

Below the species level, we are concerned with identifying how many different classes are present, determining the genetic similarities among the classes and their evolutionary relationships with wild relatives, and identi-

[1]Department of Agricultural Sciences, IACR-Long Ashton Research Station, Bristol, BS18 9AF, UK. [2]Institute of Zoology, Regents Park, London NW1 4RY, UK. [3]CPRO-DLO, PO Box 16, Wageningen 6700 AA, The Netherlands. [4]University of Udine, Via Fragagna, 209, I-33100 Udine, Italy. [5]Botanisk Laboratorium, Gothersgade 140, DK-1123 Copenhagen, Denmark. [6]Laboratoire de Genetique et Amelioration, INRA Station de Recherches de Bordeaux-Cestas, BP 45 Gazinet, Pierroton, France. [7]CNRS URA, University of Montpellier II, 34095 Montpellier Cedex 5, France. [8]Institute of Population Biology, University of Copenhagen, Universitetesparken 15, DK-2100 Copenhagen, Denmark. [9]Zoologisches Institut der Universität München, Luisenstr 14, 8000 München 2, Germany. [10]Biological Sciences, University of East Anglia, Norwich, NR4 7JJ, UK.

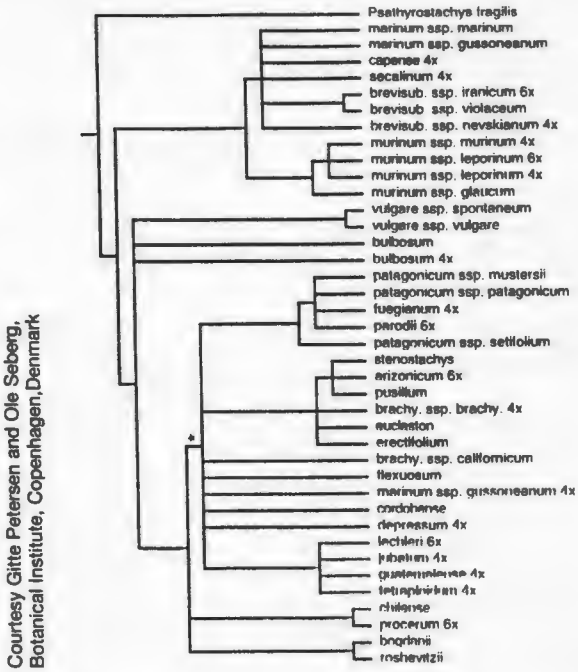
fying specific traits of interest. Much ex situ conservation, germplasm and breeding line management involves questions of this kind.

A variable number of tandem repeats (VNTR) fingerprints, amplified fragment length polymorphisms (AFLPs), and all arbitrary primed approaches [RAPDs, ISSR, DAMD, etc; see "Lexicon of molecular marker technologies"] produce multilocus profiles that are good for distinguishing between closely related genotypes. Their major applications are thus in establishing identities, determining parentage, fingerprinting genotypes, and in distinguishing genotypes below the species level. The difficulty of achieving robust profiles in arbitrary primed approaches such as RAPD does, however, make their reliability for "typing/fingerprinting" questionable. For the same reasons, band profiles are problematic for use in databases.

Questions concerning how many different classes are present and the estimation of genetic distances between them could, in principle, be tackled using any of the molecular techniques outlined in "Lexicon of molecular marker technologies." The choice will depend upon such factors as the anticipated level of polymorphism (e.g., where diversity is low, highly polymorphic markers are required, whereas the choice is wider for more diverse material) and the operational and financial resources available (e.g., RAPDs are less resource intensive than AFLPs). Caution should always be exercised, however, if information on the distribution of the markers is not known. Estimates of genetic distance between individuals (similarity or distance) may be affected by several factors: First, the number of markers used; second, the distribution of markers in the genome; and third, the nature of the evolutionary mechanisms underlying the variation measured.

Genome coverage is expected to affect the variance only in the presence of linkage disequilibrium, in which case equally spaced markers will give a better estimate than randomly distributed ones. In the case of linkage equilibrium, marker distribution is less important. This is true for most natural populations of outcrossing organisms (animals, trees, etc.), but may not be the case for selfing species, or those under strong selective pressure because of breeding. Further caution is required if classes are to be ranked in terms of evolutionary history, for reasons outlined previously.

For the location of specific traits, molecular markers that are widely distributed in the genome are required. The development of dense genetic maps, and strategies such as bulked segregant analysis, have greatly facilitated the identification of markers linked to agronomic traits. Although restriction fragment length polymorphisms (RFLPs) are attractive because of their robustness and codominance, PCR-based assays are necessary for application to the extensive sample sizes that need to be screened. Whatever the marker, it will only be of use as long as the linkage to the trait is maintained when changing from one genetic background to another. The limited extent to which genetic maps can, in detail, be transferred among crosses portends the difficulties that may have to be faced. STMS could provide the means to produce "index maps," in which the markers are easily transferable between crosses and their map position is unambiguously defined.



Courtesy Gitte Petersen and Ole Seberg, Botanical Institute, Copenhagen, Denmark

Figure 2. A dendogram constructed from molecular marker data reveals diversity patterns within resource collections (barley is shown here), facilitating both management of the collection and user access. See "Analytical tools for molecular data."

Natural populations

Population questions are fundamental to in situ conservation and include the following: How are populations of given species distributed? Are they widespread or isolated in small patches? Are they genetically distinct from one another? How much genetic variation is there? Is there gene flow among them, and how is the genetic variation distributed among populations?

Although many molecular techniques

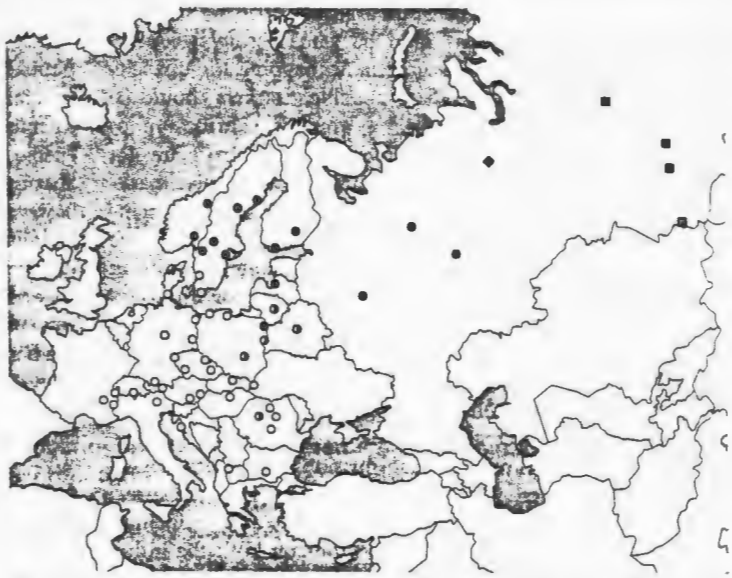
Lexicon of molecular marker techniques

Molecular marker techniques can be grouped into general categories, depending upon whether or not the assays are PCR-based and whether arbitrary/semiarbitrary primers for unknown sequences, or specifically designed primers for known sequences, are used. Non-PCR-based methods include RFLP analysis, in which DNA is digested with restriction enzymes and the resulting fragments are separated by gel electrophoresis, transferred to a filter by Southern blotting, and probes are hybridized to the filter. Hybridization to genomic DNA with probes for hypervariable regions composed of tandem repeats (known as "microsatellites" or simple sequence repeats (SSRs)—see Fig. 1) where the basic repeat unit is around 2–8 base pairs in length, and "minisatellites" for longer repeat units of approximately 16–100 base pairs) gives multilocus patterns that can resolve variation at the levels of populations and individuals[1]. This last technique is often referred to as VNTRs or oligonucleotide fingerprinting.

With the development of PCR, the necessity for probe hybridization steps could be avoided. Multiple arbitrary amplicon profiling (MAAP)[2] uses single arbitrary primers (purchasable from commercial companies) in the PCR reaction and results in the amplification of several discrete DNA products. MAAP includes RAPDs (random amplified polymorphic DNA), in which the amplification products are separated on agarose gels in the presence of ethidium bromide and visualized under ultraviolet light, and AP-PCR (arbitrary primed PCR) and DAF (DNA amplification fingerprinting), which differ from RAPDs in primer length, stringency conditions, and fragment detection. The newer technique of amplified fragment length polymorphism (AFLP)[3] is essentially intermediate between RFLP and RAPD and involves restriction digestion of the genomic DNA, followed by selective PCR amplification of the restricted fragments. The amplified products are usually separated on a sequencing gel and can be visualized after exposure to X-ray film, or by flourescent labeling. In directed amplification of minisatellite region DNA (DAMD), VNTR core sequences, such as M13, are used as primers in PCR reactions. In single-primer amplification reaction (SPAR)[4], primers are based on microsatellite core motifs. Another technique, intersimple

have been applied to questions of this kind, the most useful are codominant, single locus markers.

Information on the extent and distribution of diversity will assist in the development of efficient collecting and sampling strategies and in the identification of centers of diversity. For effective conservation, management principles have to be established[5,6] (see Fig. 3). Here, information on genetic diversity is needed to define appropriate geographical scales for monitoring and management, to establish gene flow mechanisms, and to identify the origin of individuals (e.g., to determine the role of migration). A prerequisite for conservation is the identification of populations with independent evolutionary histories and the ability to assess the conservation value of populations from an evolutionary or phylogenetic perspective. Furthermore, in the management of populations, demographic factors, such as mating systems, inbreeding depression, effective population size, and population subdivision, may be of equal importance to genetic factors[7]. Because the demographic history of a population is reflected in its genetic composition, molecular markers can provide important information on demography, provided that the data quality of different markers are taken into account[8]. STMS and sequences (haplotypes) are the markers of choice here, although the levels of polymorphism detectable in some sequences may be insufficient to yield useful information for other than the most divergent populations.



Courtesy Christoph Sperisen, Urs Büchler, and Gabor Mátyás, Swiss Federal Institute of Snow and Landscape, Switzerland.

Figure 3. The use of molecular markers enables the structure and history of diversity of a species (in this case, Norway Spruce) to be tracked. This knowledge is important for the management of populations to maintain diversity and for understanding the processes, dynamics, and biological function of biodiversity in natural and agricultural ecosystems.

Population diversity

Information on who breeds with whom and on the identity of individuals with respect to their parents is important for the management of small numbers of individuals in ex situ collections. Multilocus profiling approaches can provide extremely useful information for questions of this kind[8]. Provided the analysis is carried out properly (i.e., it is known that the bands in the fingerprint occur independently and there is no linkage disequilibrium) relatedness can be accurately estimated from band sharing coefficients for the identification of individuals (e.g., in forensics) or relatives (e.g., in mating behavior and paternity exclusion).

The future

Although molecular techniques are already available for application to biodiversity evaluation, the current technologies all suffer some technical and theoretical limitations. There is a tradeoff between different types of marker with regard to their use for diversity assessments. Techniques that generate multilocus profiles provide information on numerous (presumably) dispersed loci, although the information on a single locus is low. Conversely, sequencing and STMS are limited in loci coverage, but they are extremely informative for the locus concerned. Methods based on random (anonymous) markers have proved useful in restricted and specific applications, such as relatedness analyses or cultivar/strain identification. Even in these cases, however, more accurate answers to the same questions can be obtained with reliable markers at individual loci.

Importantly, molecular methods are useful, not only in biodiversity measurement, but also in biodiversity management. Their use makes it possible to obtain an unprecedented understanding of the processes and dynamics of biodiversity, its evolution, and

sequence repeat amplification (ISSR), involves the anchoring of designed primers to a subset of microsatellites and results in the amplification of the regions between two closely spaced, oppositely oriented, SSRs. Primers based on microsatellite (random amplified microsatellite polymorphism), transposon or interspersed repeat sequences (REP-PCR) may also be used.

To generate diversity data from specific sequences, such as genes, it is necessary to have knowledge of the sequence surrounding the target to design specific primer pairs. There are three sources of potential sequences for a PCR-targeted approach: The chloroplast (cpDNA), mitochondrial (mtDNA)[9,10] and nuclear[11] (nDNA)[12] genomes. These differ in their mode of inheritance, evolutionary rates, and recombination, all of which have important consequences in terms of their use in diversity studies. A targeted PCR approach is applicable to minute amounts of DNA from extremely small samples, e.g., single pollen grains, tiny leaf fragments, and even fossils. Sequencing the amplified fragment will potentially resolve all possible differences and the data from the aligned sequences of different individuals can then be compared. Gel systems, such as TGGE (thermal gradient gel electrophoresis), DGGE (denaturing gradient gel electrophoresis), single-strand conformational polymorphism (SSCP), and heteroduplex formation, provide sensitive assays for detecting variations down to a single base-pair and can be used to reduce the number of samples that need to be sequenced although, in practice, the generation of good TGGB and DGGE gels may be more laborious than sequencing. If the technically simpler method of PCR-RFLP, or cleaved amplified polymorphic sequence (CAPS), the amplified product is digested with a restriction enzyme and the products visualized on an agarose gel.

If SSR loci are cloned and sequenced, primers to the flanking regions can be designed to produce STMS[13,14]. STMS provide attractive markers because each primer pair usually identifies a single locus which, because of the high mutation rate of SSRs, is often multiallelic. It is common to run STMS on sequencing gels, where single repeat differences can be resolved and thus all possible alleles detected. Minisatellites are generally very difficult to clone by virtue of their size, however, if they can be isolated with sufficient flanking sequence for primer design, they provide single locus markers similar to STMS but even more polymorphic[15].

natural preservation—provided the right markers are chosen. All major advances in the field of population genetics and evolution have come from detailed studies of specific markers with well-known properties in terms of transmission, position in the genome, and mode of mutation. Of the current technologies, the marker systems that contribute most to this are STMS and sequences.

Current limitations lie in the number of well-defined markers available. Three points are relevant: First, random, or arbitrary amplification can be used as a first step toward the identification of single locus markers; second, considerable progress has been made in the field of genome mapping and sequencing of entire genomes, and a wealth of information of new gene and genomic sequences is thus being gathered; third, efficient retrieval systems for the isolation of large numbers of microsatellites from plant and animal genomes are now available.

Much could be gained from a convergence between genetic mapping and diversity studies. Where possible, markers should be chosen according to their distribution to ensure that marker sampling errors are not committed. Thus far, most molecular markers have been used in an anonymous manner—often it is not known where they are located in the genome, whether they are in coding or noncoding regions, or linked to major genes, or even sometimes whether they are in the nuclear or cytoplasmic genomes. Clearly, more information is needed to enable the classification of markers into different categories, for example, on the basis of mode of transmission, or evolution with respect to different selective pressures. Research in this area needs also to include theoretical investigations on both the influences of different marker properties and considerations of effective sampling strategies within genomes as well as at the individual, population, and geographic scales.

Finally, more facilities need to be devoted to microsatellite cloning and sequencing to enable researchers with access to the best data. Sharing and compilation of such data will, however, require the development of new bioinformatics methods adapted to the specific nature of polymorphism data. An interesting and useful byproduct of data from genome sequencing projects would be the preparation of a bank of primers of various types of organisms that would be accessible (at low or no cost) to anyone interested in applying molecular technologies to biodiversity.

Analytical tools for molecular data

It is essential to understand the way in which molecular data are analyzed[2]. Shared bands are scored as presence/absence and converted into similarity (or dissimilarity) measurements depending on the statistical method used (e.g., simple matching, Jaccard, Dice etc.). Such measures of "genetic distance" are an important way of expressing divergence (or similarity) between sequences, individuals, or taxa. All possible comparisons between the entities screened are used to construct a matrix of pairwise distances that are analyzed using clustering algorithms such as UPGMA (unweighted pair-group method using arithmetic averages) and neighbor-joining, or principle coordinate analysis (PCO). The results are presented as phenograms or principle coordinate plots that, respectively, provide graphic representations of the similarity between groups of entities or operational taxonomic units (OTUs). Nucleotide and restriction site data can be analyzed using such phenetic approaches (based on measures of overall distance/similarity) but they are also appropriate for cladistic analysis, which places samples together, not because of their high genetic similarity but because they share a given marker(s). The resultant dendograms (called cladograms) are, in fact, reconstructions of phylogeny and parsimony or maximum likelihood approaches can be used to select the best tree. The OTUs in such studies are usually selected representatives of higher taxa or species, but cladistic methods may be used where recombinations between the OTUs is negligible—for example, for well-defined non-specific populations and for nonallelic genes. Arbitrary-primed data (e.g., RAPDs) are not usually regarded as suitable for cladistic analyses because of problems of allelic assignment of bands and homology.

For investigation of diversity in natural populations, information on gene and allele frequencies is more relevant. Here, data are computed using population genetic statistics, the most common of which are F-statistics, which describe correlations between alleles at different levels of sampling (in the whole population, in subpopulations and between subpopulations)[18]. Estimates from these statistics are based on a form of hierarchial analysis of variance of allele frequencies at these different levels of subdivision of the sample. The parameter of intersubpopulation correlation (F_{st}) expresses the level of differentiation between subpopulations per generation and can serve to calculate the average number of migrants between subpopulations per generation that would give the same level of differentiation in an ideal population with isotropic migration between subpopulations. This approach poses serious problems in the definition of the parameters when numerous alleles are present at each locus and when high mutation rates cause frequent homoplasy (i.e., when identically scored alleles are not identical by descent), as in the case of microsatellites. A new approach has been proposed that incorporates information on the degree of similarity of the alleles. Several other indices, including R_{st} and N_{st}, which are analogous to Wright's F-statistics are available for use with sequence data[19] and recently, cladistic analytical approaches have been used on sequence data to reveal other population processes, including historical patterns of colonization and bottleneck events[20].

Acknowledgments

The authors thank Arnd Hoeveler, European Commission, DGXII Biotechnology, for helpful discussions. The financial support of the following EC DGXII Biotechnology grants is acknowledged: BIO2-920476; BIO2-920486; BIO2-930373; BIO2-930295.

1. Karp, A. et al. Molecular techniques in the assessment of botanical diversity. *Ann. Bot.* 78:143–149.
2. Avise, J.C. 1994. *Molecular Markers, Natural History and Evolution.* Chapman & Hall, London.
3. Vane-Wright, R.I. et al. 1991. What to protect—Systematics and the agony of choice. *Biol. Conserv.* 55:235.
4. Clark, A.G. and Lanigan, C.M.S. 1993. Prospects for estimating nucleotide divergence with RAPDs. *Mol. Biol. Evol.* 10:1096–1111.
5. Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: A critical review. *Mol. Ecol.* 3:401–411.
6. Milligan, B.G. et al. 1994. Conservation genetics: Beyond the maintenance of marker diversity. *Mol. Ecol.* 3:423–435.
7. Lande, R. 1988. Genetics and demography in biological conservation. *Science* 241:14515–1460.
8. Bruford, M.W. et al. 1992. Single locus and multi locus DNA fingerprinting. In *Molecular Genetic Analysis of Populations: A Practical Approach.* Hoelzel, A.R. (ed). IRL Press. Oxford.
9. Helentjaris, T. et al. 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant Mol. Biol.* 5:109–118.
10. Jeffreys, A.J. et al. 1985. Hypervariable 'minisatellite' regions in human DNA. *Nature* 314:67–73.
11. Caetano-Anolles, G. 1994. MAAP—A versatile and universal tool for genome analysis. *Plant Mol. Biol.* 25:1011–1026.
12. Vos, P. et al. 1995. AFLP: A new technique for DNA fingerprinting. *Nucl. Acids Res.* 23:4407–4414.
13. Demesure, B. et al. 1995. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Mol. Ecol.* 4:129–131.
14. Hillis, D.M. and Dixon, M.T. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *Quart. Rev. Biol.* 66:411–453.
15. Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA. *Nucl. Acids Res.* 17:6463.
16. Morgante M. and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant J.* 3:175–182.
17. Hillis, D.M. and Mable B.K. 1996. Applications of molecular systematics. The state of the field and a look to the future; pp. 515–543, in *Molecular Systematics*, Edn 2. Hillis D.M., Moritz C., and Mable B.K. (eds). Sinauer Associates, Sunderland, MA.
18. Weir, B.S. and Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* 38:1358–1370.
19. Goldstein, D.B. et al. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139:463–471.
20. Lynch, M. and Crease, T.J. 1990. The analysis of population survey data on DNA sequence variation. *Mol. Biol. Evol.* 7:377–394.

Plant Molecular Biology Reporter
Volume 9(3) 1991

pages 208-218

Genetic Resources

Nuclear DNA Content of Some Important Plant Species

*K. Arumuganathan and E. D. Earle**

Department of Plant Breeding and Biometry, Cornell University,
Ithaca, NY 14853-1902, USA.

Key words: Genome size, flow cytometry, propidium iodide, *Arabidopsis*, rice, tomato.

Abstract: Nuclear DNA contents of more than 100 important plant species were measured by flow cytometry of isolated nuclei stained with propidium iodide. *Arabidopsis* exhibits developmentally regulated multiploidy and has a 2C nuclear DNA content of 0.30 pg (145 Mbp/1C), twice the value usually cited. The 2C value for rice is only about three times that of *Arabidopsis*. Tomato has a 2C value of about 2.0 pg, larger than commonly cited. This survey identified several horticultural crops in a variety of families with genomes only two or three times as large as *Arabidopsis;* these include several fruit trees (apricot, cherry, mango, orange, papaya, and peach). The small genome sizes of rice and the horticultural plants should facilitate molecular studies of these crops.

Nuclear DNA contents of over 100 important plant species, were estimated by flow cytometry (Arumuganathan and Earle, 1991), and are presented in Table I. This table includes the major crop plants, some of which are widely used in molecular studies (e.g., rice and tomato), as well as *Arabidopsis thaliana*. For some, one or more nuclear DNA values (usually from Feulgen microdensitometry) have been published; for many others no such values have been reported. The information listed may be helpful to plant biologists interested in genome analysis or in the relationship of nuclear DNA content to plant physiology and ecology (Bennett, 1985).

*To whom correspondence should be addressed.

Abbreviations: C, DNA content of the unreplicated haploid chromosome complement; CRBC, chicken red blood cells; CE, chemical extraction; FC, flow cytometry; FMD, Feulgen microdensitometry; RK, reassociation kinetics.

Material and Methods

For flow cytometric analysis, suspensions of intact nuclei prepared by chopping pieces of young leaf tissues in $MgSO_4$ buffer were stained with propidium iodide and treated with DNAase-free RNAase (for details, see Arumuganathan and Earle, 1991). The DNA content of 2C nuclei (those in the G_0/G_1 phase of the cell cycle) was calculated by using nuclei from chicken red blood cells (CRBC) as internal standards. For the absolute DNA content of CRBC nuclei, we used the value of 2.33 pg/2C, as estimated chemically by Galbraith et al. (1983) after extraction of total DNA from a sample of cells.

Results and Discussion

In Table I plant species are presented in alphabetical order together with the DNA amount in 2C nuclei; a calculation of the number of base pairs per 1C (the unreplicated haploid genome of the species) is also provided. For plants that are not diploid in the cultivated form, the ploidy level of the species is included to permit estimation of actual genome size. For some species, several varieties or cultivars were assayed; nuclear DNA content often varied somewhat among the different cultivars, as has been previously reported (Bennett, 1985). Table II presents data for the same species in increasing order of genome size.

Arabidopsis has been reported to have the smallest genome known among flowering plants (Leutwiler et al., 1984). In our survey *Arabidopsis* also had the lowest value; its 2C nuclei contained 0.30 ± 0.012 pg DNA (mean ± standard deviation, n = 36). This is, however, about twice the widely cited genome size (70 Mbp or ~0.0725 pg DNA per 1C genome), deduced from DNA reassociation kinetics (Leutwiler et al., 1984).

Arabidopsis showed multiploidy in many of its tissues. Nuclei from leaves produced four or five distinct peaks of fluorescence (Fig. 1A), representing the 2C, 4C, 8C, 16C, and 32C complements of the genome. Fewer than 30% of the nuclei were in the 2C peak. The fourth (16C) peak coincides with or falls very close to the peak for nuclei of CRBC, used as an internal DNA standard (Fig. 1B). Although nuclei from rosette leaves and roots showed up to six peaks, nuclei from leaves on the flowering stem showed three or four; those from the flowering stem itself had only three. Nuclei from flower buds fall into only two peaks of fluorescence (Fig. 1C). Similar results were seen with cv. Columbia and cv. Landsberg (*erecta* mutant) and with plants grown either in sterile nutrient agar medium or in soil. These results suggest that multiploidy in *Arabidopsis*

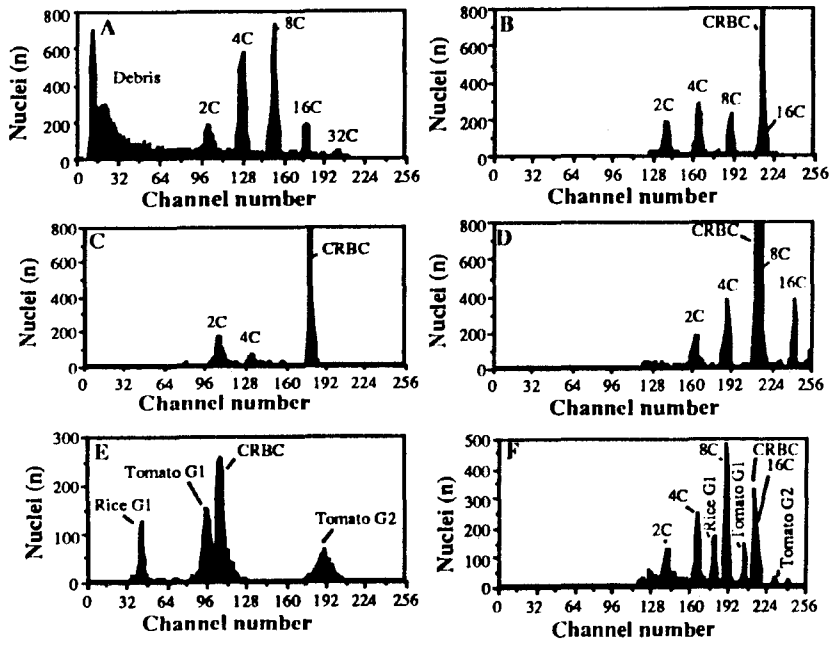


Fig. 1. Histograms of numbers of nuclei per channel as a function of relative fluorescence intensity. Signals from nuclei were gated to eliminate much of the debris from analysis (except in A). Histograms acquired on EPICS PROFILE flow cytometer are presented in Macintosh format (Cameron, 1990). The channel number is proportional to the log of the fluorescence intensity (except in E), such that a difference of 26 channels represents a doubling of fluorescence intensity. Nuclei from chicken red blood cells (CRBC) were included (except in A) to serve as an internal standard for the determinations of plant nuclear DNA contents. Nuclear DNA contents were calculated by comparing mean peak positions of the plant nuclei to the mean peak positions of the nuclei of CRBC. (A) Five fluorescence peaks representing five ploidy classes (2C, 4C, 8C, 16 C and 32C) for nuclei isolated from leaf tissues of *A. thaliana*. (B) The fourth nuclear peak from mature leaf of *A. thaliana* coincides with the nuclear peak of CRBC. (C) Only two peaks of fluorescence are present from nuclei (2C and 4C) from flower buds of *A. thaliana*. (D) The third nuclear peak from mature leaf of a transgenic *Arabidopsis* plant coincides with the CRBC peak, indicating that the plant is a tetraploid with at least four ploidy levels, as in normal plants. (E) The relative peak positions for the nuclei from leaves of rice and tomato, and CRBC. (F) Relative fluorescence peak positions of nuclei from *Arabidopsis*, rice, tomato and CRBC.

Table 1. Nuclear DNA content of a number of important plant species as determined by flow cytometry.

Scientific name	Common name	Family	Nuclear DNA content: This study pg/2C[a] (N)	This study ~Mbp[b]/1C	Previous work pg/2C (method used, ref)[c]
Aegilops squarrosa		Graminae	8.34	4024	7.2-11.6 (FMD, 1)
Allium ampeloprasum	Leek	Amaryllidaceae	50.27	24255	ND
Allium cepa	Onion	Amaryllidaceae	31.69, 32.74 (2)	15290, 15797	3.55 (FMD, 1)
Ananas bracteatus	Red pineapple	Bromeliaceae	0.92	444	ND
Ananas comosus	Pineapple	Bromeliaceae	1.09	526	ND
Arabidopsis thaliana	Arabidopsis	Cruciferae	0.30 (2)	145	0.5 (FMD, 1); 0.15 (RK, 2)
Arachis hypogaea (2n=4X)	Peanut/groundnut	Leguminosae	5.83	2813	3.5 (FMD, 5)
Asparagus officinalis	Asparagus	Liliaceae	2.71	1308	ND
Avena sativa	Oats	Gramineae	23.45	11315	26.5-27.5 (FMD, 1)
Beta vulgaris ssp. esculenta	Beet/beetroot	Chenopodiaceae	1.48	714	2.5 (FMD, 1)
Beta vulgaris ssp. saccharifera	Sugar beet	Chenopodiaceae	1.57	758	2.6 (FMD, 1)
Brassica campestris ssp. chinensis	Pak choi	Cruciferae	1.05	507	1.6 (FMD, 1)
Brassica campestris ssp. oleifera	Turnip rape	Cruciferae	0.97-1.07 (6)	468-516	1.6 (FMD, 1)
Brassica campestris ssp. rapifera	Turnip	Cruciferae	1.06	511	1.6 (FMD, 1)
Brassica hirta (=*Sinapis alba*)	White mustard	Cruciferae	1.02	492	1.0 (FMD, 3)
Brassica juncea	Brown mustard	Cruciferae	2.29	1105	3.1 (FMD, 1)
Brassica napus	Rapeseed	Cruciferae	2.34-2.56 (4)	1129-1235	3.2 (FMD, 1)
Brassica nigra	Black mustard	Cruciferae	0.97	468	1.6 (FMD, 1)
Brassica oleracea ssp. botrytis	Cauliflower	Cruciferae	1.30-1.37 (4)	628-662	1.8 (FMD, 1)
Brassica oleracea ssp. capitata	Cabbage	Cruciferae	1.25	603	1.8 (FMD, 1)
Brassica oleracea ssp. gemmifera	Brussels sprouts	Cruciferae	1.30	628	1.8 (FMD, 1)
Brassica oleracea ssp. italica	Broccoli	Cruciferae	1.24, 1.28 (2)	599, 618	1.8 (FMD, 1)
Brassica tournefortii		Cruciferae	1.64	791	ND
Capsicum annuum	Pepper/chili	Solanaceae	5.6-7.51 (7)	2702-3420	8.0-10.8 (FMD, 5); 5.52 (FC, 7)

Carica papaya	Papaya	Caricaceae	0.77	372	ND
Cicer arietinum	Chick pea	Leguminosae	1.53	738	1.9 (FMD, 1)
Citrullus vulgaris (= *lanatus*)	Watermelon	Cucurbitaceae	0.88, 0.90 (2)	425, 434	ND
Citrus sinensis	Orange	Rutaceae	0.76, 0.82 (2)	367, 396	ND
Crepis capillaris	Crepis	Compositae	3.87	1867	4.1 (FMD, 1); 3.55 (FC,8)
Cucumis melo	Cantaloupe	Cucurbitaceae	0.94, 1.04 (2)	454, 502	1.9 (FMD, 3)
Cucumis sativus	Cucumber	Cucurbitaceae	0.76	367	2.1 (FMD, 1)
Cucurbita pepo	Zucchini	Cucurbitaceae	1.04, 1.08 (2)	502, 521	5.6 (FMD, 3)
Datura stramoniun	Jimson weed	Solanaceae	4.11	1983	ND
Daucus carota	Carrot	Umbelliferae	0.98	473	2.0 (FMD, 1)
Dioscorea alata	Yam	Dioscoreaceae	1.15	555	ND
Diplotaxis erucoides		Cruciferae	1.31	632	ND
Eruca sativa		Cruciferae	1.16	560	ND
Glycine max (2n=4X)	Soybean	Leguminosae	2.31	1115	1.9 (FMD, 3)
Gossypium hirsutum (2n=4X)	Cotton	Malvaceae	4.39, 4.92 (2)	2118, 2374	6.1-6.5, 6.06 (FMD, 1, 6); 5.6 (FC, 6)
Helianthus annuus	Sunflower	Compositae	5.95-6.61 (3)	2871-3189	4.9-9.9, 7.22 (FMD, 3, 6); 3.67, 6.96 (FC, 7, 6)
Hordeum vulgare	Barley	Gramineae	10.10	4873	10.7-11.1 (FMD, 1)
Ipomoea batatas (2n=6X)	Sweet potato	Convolvulaceae	3.31	1597	ND
Lactuca sativa	Lettuce	Compositae	5.47	2639	5.32 (FC, 6); 6.06 (FMD, 6)
Lens culinaris (= *esculenta*)	Lentil	Leguminosae	8.42	4063	9.2 (FMD, 1)
Lycopersicon cheesemanii		Solanaceae	1.83	883	ND
Lycopersicon esculentum	Tomato	Solanaceae	1.88-2.07 (6)	907-1000	1.48, 1.9 (FC, 7, 6); 2.0-5.1, 2.05 (FMD, 1, 6)
Lycopersicon pennellii		Solanaceae	2.47-2.77 (3)	1192-1337	ND
Lycopersicon peruvianum		Solanaceae	2.27	1095	ND
Malus x domestica (2n=2X)	Apple	Rosaceae	11.54-1.65 (3)	743-796	ND
Mangifera indica	Mango	Anacardiaceae	0.91	439	ND

Manihot esculenta (= *utilissima*)	Cassava / manioc	Euphorbiaceae	1..43-1.72 (17)	690-830	ND
Medicago sativa (2n=4X)	Alfalfa / lucerne	Leguminosae	3.13	1510	3.5 (FMD, 1)
Medicago truncatula		Leguminosae	0.94-1.09 (6)	454-526	ND
Melilotus officinalis	Sweet clover	Leguminosae	2.25	1086	ND
Musa sp.	Banana	Musaceae	1.81	873	ND
Nicotiana plumbaginifolia		Solanaceae	4.74	2287	ND
Nicotiana tabacum (2n=4X)	Tobacco	Solanaceae	8.75-9.63 (4)	4221-4646	7.8 (FMD, 1); 9.67 (FC, 7)
Oryza longistaminata	African rice	Gramineae	0.78	376	1.5, 1.25 (FMD, 3, 4)
Oryza sativa ssp. Indica	Rice	Gramineae	0.87-0.96 (49)	419-463	1.2, 1.67, 2.1 (FMD, 1, 4, 3)
Oryza sativa ssp. Japonica	Rice	Gramineae	0.86-0.91 (20)	415-439	1.2, 1.55, 1.9 (FMD, 1, 4, 3)
Oryza sativa ssp. Javanica	Rice	Gramineae	0.88 (3)	424	ND
Passiflora menspermifolia	Passion fruit	Passifloraceae	4.54	2191	ND
Persea americana	Avocado	Lauraceae	1.83	883	ND
Petroselinum crispum	Parsley	Umbelliferae	3.96	1911	ND
Petunia hybrida	Petunia	Solanaceae	2.64	1274	3.1, 3.9 (FMD, 1, 3)
Petunia parodii		Solanaceae	2.53	1221	ND
Phaseolus acutifolius	Tepary bean	Leguminosae	1.34	647	ND
Phaseolus coccineus	Scarlet runner bean	Leguminosae	1.47	709	3.5, 1.9 (FMD, 1, 3)
Phaseolus lunatus	Lima bean	Leguminosae	1.29	622	2.5 (FMD, 1)
Phaseolus vulgaris	Common bean	Leguminosae	1.32	637	3.7, 2.7 (FMD, 1, 3)
Pisum sativum	Garden pea	Leguminosae	8.18, 9.11 (2)	3947, 4397	9.8-10.5, 9.75 (FMD, 1, 6); 7.72, 8.55 (FC, 7, 6)
Prunus armenaica	Apricot	Rosaceae	0.61	294	ND
Prunus avium	Sweet Cherry	Rosaceae	0.70	338	ND
Prunus avium x cerasus (2n=4X)	Cherry	Rosaceae	1.42	685	ND
Prunus cerasus (2n=4X)	Sour cherry	Rosaceae	1.24	599	ND
Prunus domestica (2n=6X)	Prune	Rosaceae	1.83	883	ND
Prunus persica	Peach	Rosaceae	0.54, 0.55 (2)	262, 265	ND
Pyrus communis	Pear	Rosaceae	1.03, 1.11 (2)	496, 536	ND
Raphanus sativus	Radish	Cruciferae	1.09	526	0.9 (FMD, 3)

Ricinus communis	Castor bean	Euphorbiaceae	0.67	323	ND
Rubus idaeus	Raspberry	Rosaceae	0.58	280	ND
S. barberi X S. spontaneum	Sugarcane	Gramineae	6.12	2953	ND
Saccharum barberi	Sugarcane	Gramineae	6.54, 8.34 (2)	3156, 4121	ND
Saccharum officinarum	Sugarcane	Gramineae	5.28-7.47 (3)	2547-3605	8.1- 8.7 (FMD, 1)
Saccharum robustum	Sugarcane	Gramineae	6.53	3151	7.7-8.6 (FMD, 1)
Saccharum sinense	Sugarcane	Gramineae	8.67	4183	ND
Sesbania rostrata		Leguminosae	2.46	1187	ND
Sinapis arcensis		Cruciferae	0.76	367	ND
Solanum berthaultii		Solanaceae	1.74	840	ND
Solanum melongena	Eggplant	Solanaceae	2.28, 2.48 (2)	1100, 1197	2.33 (FC, 7)
Solanum tuberosum (2n=4X)	Potato	Solanaceae	3.31-3.86 (3)	1597-1862	4.2 (FMD, 1)
Sorghum bicolor	Sorghum	Gramineae	1.55, 1.60 (2)	748, 772	1.56-1.74 (FMD, 5); 1.74 (FC, 6)
Spinacia oleracea	Spinach	Chenopodiaceae	2.05	989	1.9 (FMD, 3)
Trifolium pratense	Red clover	Leguminosae	0.97	468	ND
Trifolium repens	White clover	Leguminosae	2.07	999	ND
Tripsacum dactyloides	Gama grass	Gramineae	7.73	3730	ND
Triticum aestivum (2n=6X)	Wheat	Gramineae	33.09	15966	34.6 (FMD, 1, 6); 36.11 (FC, 6)
Triticum monococcum		Gramineae	11.92	5751	12.4-13.4 (FMD, 1)
Tulipa sp.	Garden tulip	Liliaceae	51.2, 63.6 (2)	24704, 30687	72.0 per cell (CE, 1)
Vanilla planifolia	Vanilla	Orchidaceae	15.90	7672	ND
Vigna mungo	Black gram / urud	Leguminosae	1.19	574	ND
Vigna radiata	Mung bean	Leguminosae	1.20 (2)	579	1.1 (FMD, 3)
Vigna unguiculata (= *sinensis*)	Cowpea	Leguminosae	1.27	613	ND
Vitis vinifera	Grape	Vitaceae	1.00	483	ND
X Citrofortunella mitis	Calamondin orange	Rutaceae	0.80	386	ND
Zea diploperennis		Gramineae	3.57	1723	5.28 (FMD, 5)
Zea mays	Corn	Gramineae	4.75-5.63 (6)	2292-2716	4.42-6.75 (FMD, 5); 5.99 (FC, 7)

Notes to Table I.

[a]Value for each cultivar was determined by two or more measurements of at least 2000 nuclei. N, number of cultivars or varieties examined.

[b]1 picogram (pg) = 965 million base pairs (Mbp) (1).

[c]FMD, Feulgen microdensitometry; FC, flow cytometry; ND, not determined; RK, reassociation kinetics; CE, chemical extraction; Ref., reference: 1, Bennett and Smith (1976); 2, Leutwiler et al. (1984); 3, Bennett et al. (1982); 4, Iyengar and Sen (1978); 5, Laurie and Smith (1982); 6, Michaelson et al. (1991). 7, Galbraith et al. (1983). 8, Galbraith (1990).

is under developmental control, with younger tissues showing less or no multiploidy, as in the succulent species with small genomes examined by De Rocher et al. (1990).

Analysis of nuclei from leaves of a transgenic *Arabidopsis* plant suspected of being tetraploid supported our estimation of the nuclear DNA value in this species. These nuclei produced four distinct peaks of fluorescence as in normal plants, but the third, rather than the fourth, peak coincided with the CRBC peak (Fig. 1D). This indicates that the plant is a tetraploid with four levels of ploidy. The nuclei with the lowest fluorescence intensity in the tetraploid plant had a DNA content of 0.60 ± 0.02 pg. These results are confirmation that the 2C value for diploid plants is indeed 0.30 pg; i.e., diploids have no smaller peak not visible on our flow histograms.

It is notable that several horticultural crops in a variety of families have genomes only two or three times as large as that of *Arabidopsis* (Table II). Among these are fruit trees (e.g. peach, apricot, sweet cherry) in the family Rosaceae (Table I). Thus low nuclear DNA content is not always associated with small size or short life cycles. No multiploidy was observed in any of these plants, so that phenomenon is also not a universal feature of small genome size.

The nuclear genome sizes of rice and tomato are of particular interest because these plants are being used extensively for mapping, isolation, and transfer of genes and are the most likely candidates for possible complete genome sequencing. Rice has a nuclear genome only about three times as large as that of *Arabidopsis*. The 2C genome size of *Oryza sativa* ssp. *indica* (cv. IR 36) is 0.90 ± 0.03 pg (mean ± standard deviation, n=31). The 2C value for tomato (cv. VFNT cherry) is 2.02 ± 0.13 pg (mean ± standard deviation, n=43), about 2.2 times as large as the rice genome (Fig. 1E). These values have been confirmed in many separate trials in our laboratory using other cultivars of rice (range of 0.86-0.96 pg/2C) and tomato (range of 1.88-2.07 pg/2C). The relative fluorescence of a mixture

Table II. DNA content of unreplicated haploid genome of various plants arranged according to genome size.

Plant species	DNA Mbp/1C	Plant species	DNA Mbp/1C
Arabidopsis thaliana	145	*Diplotaxis erucoides*	632
Prunus persica	262, 265	*Phaseolus vulgaris*	637
Rubus idaeus	280	*Phaseolus acutifolius*	647
Prunus armenaica	294	*Prunus avium x cerasus* (2n=4X)	685
Ricinus communis	323	*Manihot esculenta* (= *utilissima*)	690-830
Prunus avium	338	*Phaseolus coccineus*	709
Cucumis sativus	367	*Beta vulgaris ssp. esculenta*	714
Sinapis arvensis	367	*Cicer arietinum*	738
Citrus sinensis	367, 396	*Malus x domestica* (2n=2X)	743-796
Carica papaya	372	*Sorghum bicolor*	748, 772
Oryza longistaminata	376	*Beta vulgaris ssp. saccharifera*	758
X Citrofortunella mitis	386	*Brassica tourneforthii*	791
Oryza sativa ssp. Japonica	415-439	*Solanum berthaultii*	840
Oryza sativa ssp. Indica	419-463	*Musa sp.*	873
Oryza sativa ssp. Javanica	424	*Persea americana*	883
Citrullus vulgaris (= *lanatus*)	425, 434	*Prunus domestica* (2n=6X)	883
Mangifera indica	439	*Lycopersicon cheesemanii*	883
Ananas bracteatus	444	*Lycopersicon esculentum*	907-1000
Cucumis melo	454, 502	*Spinacia oleracea*	989
Medicago truncatula	454-526	*Trifolium repens*	999
Trifolium pratense	468	*Melilotus officinalis*	1086
Brassica nigra	468	*Lycopersicon peruvianum*	1095
Brassica campestris ssp. oleifera	468-516	*Solanum melongena*	1100, 1197
Daucus carota	473	*Brassica juncea*	1105
Vitis vinifera	483	*Glycine max* (2n=4X)	1115
Brassica hirta (=*Sinapis alba*)	492	*Brassica napus*	1129-1235
Pyrus communis	496, 536	*Sesbania rostrata*	1187
Cucurbita pepo	502, 521	*Lycopersicon pennellii*	1192-1337
Brassica campestris ssp. chinensis	507	*Petunia parodii*	1221
Brassica campestris ssp. rapifera	511	*Petunia hybrida*	1274
Raphanus sativus	526	*Asparagus officinalis*	1308
Ananas comosus	526	*Medicago sativa* (2n=4X)	1510
Dioscorea alata	555	*Ipomoea batatas* (2n=6X)	1597
Eruca sativa	560	*Solanum tuberosum* (2n=4X)	1597-1862
Vigna mungo	574	*Zea diploperennis*	1723
Vigna radiata	579	*Crepis capillaris*	1867
Prunus cerasus (2n=4X)	599	*Petroselinum crispum*	1911
Brassica oleracea ssp. italica	599, 618	*Datura stramoniun*	1983
Brassica oleracea ssp. capitata	603	*Gossypium hirsutum* (2n=4X)	2118, 2374
Vigna unguiculata (= *sinensis*)	613	*Passiflora menspermifolia*	2191
Phaseolus lunatus	622	*Nicotiana plumbaginifolia*	2287
Brassica oleracea ssp. gemmifera	628	*Zea mays*	2292-2716
Brassica oleracea ssp. botrytis	628-662	*Lactuca sativa*	2639

Table II, continued.

Plant species	DNA Mbp/1C	Plant species	DNA Mbp/1C
Capsicum annuum	2702-3420	*Lens culinaris* (= *esculenta*)	4063
Arachis hypogaea (2n=4X)	2813	*Nicotiana tabacum* (2n=4X)	4221-4646
Helianthus annuus	2871-3189	*Hordeum vulgare*	4873
Saccharum officinarum	2547-3605	*Triticum monococcum*	5751
S. barberi X S. spontaneum	2953	*Vanilla planifolia*	7672
Saccharum robustum	3151	*Avena sativa*	11315
Saccharum barberi	3156, 4121	*Allium cepa*	15290, 15797
Saccharum sinense	4183	*Triticum aestivum* (2n=6X)	15966
Tripsacum dactyloides	3730	*Allium ampeloprasum*	24255
Pisum sativum	3947, 4397	*Tulipa sp.*	24704, 30687
Aegilops squarrosa	4024		

of nuclei from *Arabidopsis*, rice, tomato and CRBC is shown in Fig. 1F; the peak for rice G_1 nuclei falls between the 4C and 8C nuclear peaks of *Arabidopsis*; the tomato G_1 peak falls between the 8C and CRBC peaks.

Comparisons of our nuclear DNA values to those from other methods and/or studies are included in Table I. For most plant species, including ones with diverse genome sizes (e.g. mung bean, sorghum, corn, pea, barley, oats, onion, wheat), our results from flow cytometry are in good agreement with those obtained in the laboratory of M. D. Bennett by Feulgen microdensitometry (although not with some of the values he compiled from earlier work elsewhere). The *Arabidopsis* values estimated by reassociation kinetics (0.0725 pg/1C) and by Feulgen microdensitometry (0.2 pg/1C) are substantially different; our results (0.15 pg/1C) are intermediate between the two but closer to the latter (Table I). Our values for rice (0.45/1C) are somewhat lower than those from microdensitometry (0.6 pg/1C or more). The widely used value of 0.74 pg/1C for tomato, obtained by flow cytometry (Galbraith et al., 1983), is about 25% lower than ours. This may be due to their use of mithramycin, a fluorescent stain that preferentially binds to G-C nucleotide pairs. Because the tomato genome has a G-C content of only 37% (Messeguer et al., 1991), staining with mithramycin is likely to underestimate actual genome size. Variation in nuclear DNA contents of different cultivars may also have contributed to the differences seen.

The flow cytometry protocol used in our study (Arumuganathan and Earle, 1991) should permit easy determination of nuclear DNA values for additional plants not listed in Table I or previously published reports. We would be glad to learn of such results for compilation of a more complete table.

Acknowledgments: We thank S. Tanksley, R. Last, R. Wu, M. Hanson, M. Ganal, G. Martin, R. Wing and J. Slattery for reviewing early versions of the manuscript; M. Rangel-Lugo for providing chicken blood; L. A. Batts for technical assistance; R. Last for the tetraploid transgenic *Arabidopsis* plant; H. Kikuchi and C. Martinez for some of the rice and cassava values; D. Heather for some of the values for *Brassica* and other Cruciferae. This work was supported by a grant from the Cornell Biotechnology program (which is sponsored by the New York State Science and Technology Foundation, a consortium of industries, US Army Research Office and the National Science Foundation), and by the United States Department of Agriculture.

Note added in proof: A recent manuscript [Galbraith, D. W, K. R. Harkins and S. Knapp. 1991. Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol. (in press)] also reports the presence of multiploidy in the vegetative tissues of *Arabidopsis* and estimates the 2C nuclear DNA content as 0.32, 0.33 and 0.34 pg using mithramycin, Hoechst 33258 and DAPI, respectively.

References

Arumuganathan, K. and E.D. Earle 1991. Estimation of nuclear DNA content of plants by flow cytometry. Plant Mol. Biol. Reporter 9:221-231.

Bennett, M. D. 1985. Intraspecific variation in DNA amount and the nucleotypic dimension in plant genetics. In: *Plant Genetics* (ed. M. Freeling), pp. 283-302. Alan R. Liss, New York.

Bennett, M. D. and J. B. Smith. 1976. Nuclear DNA amounts in angiosperms. Phil. Trans. R. Soc. Lond. B 274: 227-274.

Bennett, M. D, J. B. Smith and J. S. Heslop-Harrison. 1982. Nuclear DNA amounts in angiosperms. Proc. R. Soc. Lond. B216: 179-199.

Cameron, M. J. 1990. Macintosh graphics for the EPICS flow cytometer user. Cytometry 11:916-918.

De Rocher, E. J., K. R. Harkins, D. W. Galbraith and H. J Bohnert. 1990. Developmentally regulated systemic endopolyploidy in succulents with small genomes. Science 250: 99-101

Galbraith, D. W., K. R. Harkins, J. M. Maddox, N. M Ayres, D. P. Sharma and E. Firoozabady. 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220:1049-1051.

Galbraith, D. W. 1990. Flow cytometric analysis of plant genomes. In: *Methods in Cell Biology* (eds. Z. Datzynkiewicz and H. Crissman), pp. 549-563. Academic Press, New York.

Iyengar, G.A.S and S. K. Sen. 1978. Nuclear DNA content of several wild and cultivated *Oryza* species. Environ. Exp. Bot. 18: 219-224.

Leutwiler, L. S., B. R. Hough-Evans and E. M. Meyerowitz. 1984. The DNA of *Arabidopsis thaliana*. Mol. Gen. Genet. 194:15-23.

Laurie, D. A. and M. D. Bennett. 1985. Nuclear DNA content in the genera *Zea* and *Sorghum*. Intergeneric, interspecific and intraspecific variation. Heredity 55: 307-313.

Messeguer, R., M. W. Ganal, J. C. Steffens and S. D Tanksley. 1991. Characterization of the level, target sites, and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol. Biol. 16:753-770.

Michaelson, M. J, H. J. Price, J. R. Ellison and J. S. Johnston. 1991. Comparison of plant DNA contents determined by Feulgen microspectrophotometry and laser flow cytometry. Amer. J. Bot. 78: 183-188.

Erratum

Birgit Bochenek and Ann M. Hirsch. 1990. *In-situ* hybridization of nodulin mRNAs in root nodules using non-radioactive probes. Plant Mol. Biol. Report. 8(4):237-248.

p. 243, third line from bottom: "25%" acetic anhydride should read "0.25%".

***Cis and Trans*: Plant Genomes: Databases, Sizes, and Nomenclature** continued from p. 189

two opposing effects: it confers a high degree of redundancy into the names of genes (should all genes from *Lupinus luteus* start with *lu*?), and it complicates computer searches by implying differences among genes whose sequences may in fact be very similar.

In an article starting on p. 220, Dure proposes a numerical system of gene nomenclature based on the IUB system for classifying enzymes.

The Plant Journal—A New Publication

Last month saw the debut of a new journal, published by Blackwell Scientific Publications of Oxford, that will focus on the "plant molecular sciences". Promising rapid publication, the editors of *The Plant Journal* (Dianna Bowles, Michel Caboche, Liz Dennis, Dick Flavell, Atsushi Komamine, Jeff Schell, Chris Somerville, and Lothar Willmitzer) note that "advances in the development and application of powerful molecular techniques, particularly when combined with classical approaches, have largely overcome traditional barriers in the study of plants." Six of the eight editors are long-time members of the ISPMB.

Ω

crease the on-going self-reliance of individuals, families, and mmunities.
mote a program balance between learning processes (enduring skills) d recipes of technical information (perishable skills).
vide a source of objective information concerning environmental, nomic, and social impacts of change and our responsibilities to ture generations.
ng about voluntary positive change that addresses the challenges of future.
lp people understand why regulation is sometimes necessary.
ntribute to the development of a more innovative, and competitive ivate sector.
ild a collaborative and responsive public sector through essential liances.
ove citizens through its programs, giving them more control over their ves and greater ability to manage changing circumstances.

sion should avoid:

coming an "entitlement" or a continuing "free service" to clientele.
coming an advocate for, or a defender of, a particular customer or ient group.
roviding services that do not involve learning for Extension or the client.
oing for people what they can do for themselves or which is done by ther organizations or agencies.

t do You Think?

m very interested in your reaction to the ideas presented in this paper. e send your thoughts and suggestions to:

Peter Bloome, Assistant Director
Illinois Cooperative Extension Service
214 Mumford Hall
1301 West Gregory Drive
Urbana, IL 61801

Phone: 217-244-9407
FAX: 217-333-1952
E-Mail: p-bloome@uiuc.edu

Tools to Determine the Function of Genes

Robert Meeley, Katherine Canada, Arnold Oliphant, Steven Briggs
Trait and Technology Development
Pioneer Hi-Bred International, Inc.
Johnston, IA

Mutants are the Key

How would a novice in the realm of classical music answer the task of describing an oboe's contribution to a particular symphony? He could try to play it, or ask to hear it played. He could take the oboe apart and study how the sound is produced, and learn what range it has, but all of these details lack information about how the oboe's sound interacts with the other elements of the orchestra. Perhaps the best "tool" at the novice's disposal would be the following experiment: ask that the piece be played two ways—once with the full orchestra, and again without the oboes playing. The novice could then compare the two performances and describe whether the oboe plays a major or minor role in the composition.

This is an allegory for how mutants are useful for the study of gene function, or more precisely, for determining the role of a given gene in an organism. The consequent phenotype of mutant gene is like a window into normal gene function. There are many types of mutation, reflecting any change to a gene's DNA sequence, thus the consequences of mutation can vary greatly. In terms of functional analysis of genes, the type of mutation that is most useful is the *null mutation*—a change to a gene that precludes the expression of a functional gene product. This is the biological equivalent of our silent oboe.

ROBERT MEELEY

Controlling, inducing, and exploiting null mutations are common tools of the trade for molecular biologists, geneticists, and physiologists. Plant breeders and seedsmen, too, make masterful use of these genetic tools, so perhaps I belabor the point. For the purpose of clarity, consider some simple cases where null-

mutations have been used to influenceseed quality. For example,breeders have known for some time that silencing a single gene can have dramatic effects on seed quality. The mutations waxy (*wx1*) and amylose-extender (*ae1*) have been used to convert elite dent lines to high amylopectin and high amylose germplasm, respectively. Exploitation of the sugary-1 (*su1*) and shrunken-2 (*sh2*) genes for sweet corn development has also occupied a significant niche in commercial breeding efforts.

The Growing DNA Sequence Database

Just as the above mutations have drastic effects on kernel phenotype, we can learn many things about each maize gene's contribution to the whole plant if a simple and efficient method was available to select for null mutations in any gene of interest. Technical advances made over the last several years now make such a method feasible. Gene library construction, DNA sequencing, and database technologies have each progressed to the point where a focused effort, over as little as five years, could reveal the sequences for all expressed maize genes. This type of "genome science" is arguably as important for the future of agriculture as it is to human medicine. With so much DNA sequence information coming so fast, the sequence database will grow far in advance of functional information about the role of each gene in the plant. This is akin to having all musical instruments at our disposal, but knowing nothing about what their possible contributions may be. In many cases, the functional significance of a gene may be inferred by homology to known genes from other organisms; a type of analysis performed by computers within a database. The next decade will involve an explosion of this type of interspecies "cross-talk". However, many maize sequences will be found that have no significant interspecies homology, or at the least, their functional significance in maize will have to be investigated directly. A resource "bank" of maize mutants would provide valuable genetic materials to study the phenotypic contributions of specific genes. The benefit of this information is manifest—mutants may provide insight into agronomically important genes, assist efforts for gene cloning and genetic engineering, and promote basic research efforts by integrating DNA sequence information with biological relevance.

What is necessary for being able to assemble such a resource? Generally, three elements: 1) a method to effectively saturate the maize genome with tagged mutations, 2) a molecular pointer to search the collection for mutations in specific genes of interest, and 3) a genetic source for propagation and phenotypic analysis of mutant alleles. The following sections describe how such a resource has been assembled to facilitate the functional analysis of

Transposable Elements

Mutagenesis in maize has at its disposal a most useful genetic feature: transposable elements. Transposable elements (TE), or transposons, occur in many prokaryotic and eukaryotic organisms, but were first described in maize. These are the so-called "jumping genes" that were reported by Dr. Barbara McClintock. In 1950, McClintock published a description of the Activator/Dissociator (Ac/Ds) class of elements, and documented cases of genetic instability (mutation) that were attributable to the transposition of discrete genomic elements (1). Her work in this area was largely ignored until transposon systems were noted in bacteria and yeast; she then received the Nobel Prize for her work in 1983. Since McClintock's discoveries, other TE systems in maize have been characterized. Our gene function resource relies on one of these TE systems—the *Mutator* family of maize transposable elements.

Mutator is a family of transposable elements broken down into five major classes based on the relatedness of their DNA sequences. The *Mutator* family was described by Dr. Don Robertson in the late 1970's, when he noted an abnormally high frequency of germinal and somatic mutation in specific genetic backgrounds (2). The first *Mu* element to be cloned was accomplished by Dr. Jeff Bennetzen in 1984, who was able to "trap" *Mu1* by selecting for *Mu* disruption at the *Adh1* locus of maize (3). The cloning of *Mu1* allowed research to progress on the cloning and sequence characterization of other *Mu* elements. This information enabled the development of the gene isolation technique known as transposon "tagging," and now makes sequence-based selection of transposon-disrupted maize genes feasible.

Transposition of *Mutator* is controlled by the autonomous regulatory element known as *Mu-DR* (in honor of Dr. Robertson) (4). In genetic backgrounds where *Mu* is "on," transcripts encoded by *Mu-DR* are expressed, perhaps triggered by a cellular signal produced at or near the time of meiosis. New germinal mutations are produced in both the male and female gametophytes by the transposition of member elements into new locations—predominantly the coding regions of genes. Progeny plants from these gametes inherit these new transposon-insertion events. When the transposon inserts into a gene, this disrupted allele frequently behaves as a recessive null mutation. The high germinal mutation frequency, known sequence composition, and highly conserved DNA sequences of *Mu* element termini are several of the features that distinguish this family as the best transposable element system to exploit for the creation of our mutant bank.

Saturation of the Maize Genome

As shown in Figure 1, the genetic method required to disperse *Mutator* throughout the maize genome is quite straightforward. For the most part,

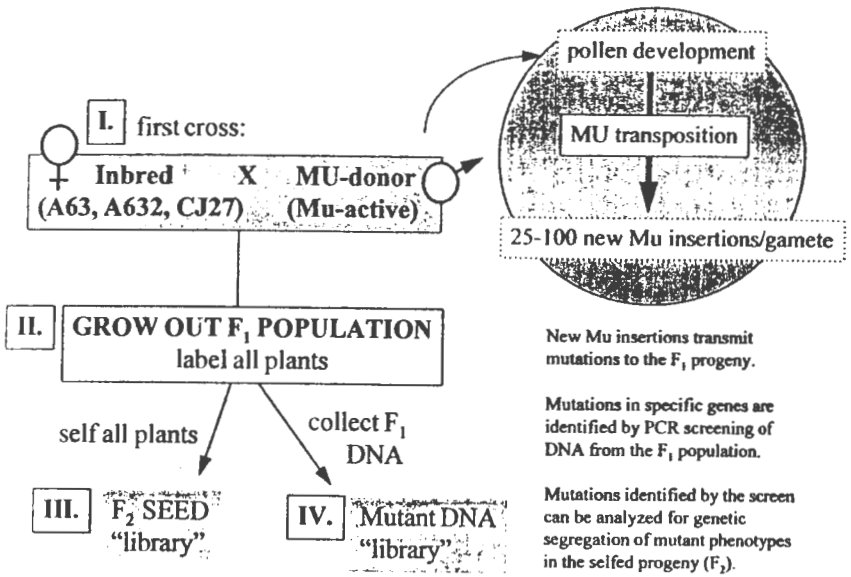


Figure 1 Genetics for saturation mutagenesis.

which contain many copies of *Mu* elements, and at least one copy of *Mu-DR*, can be maintained as field stocks. Transposition activity can be monitored by performing a variety of genetic tests, developed by Robertson, which determine if transposition is "on" within these lines. Insertion mutations are transmitted to progeny plants by performing crosses between the *Mu*-active line, and a recipient plant, such as an inbred. Since the greatest number of gametes is produced in the tassel, *Mutator* activity is typically transmitted to the recipient by using the *Mu* donor as the male parent. By indiscriminately crossing a number of *Mu* donors to recipient inbreds, tens of thousands of TE-mutagenized gametes can be fixed in the resultant F_1 population. The mutant copies of genes that *Mu* creates in the pollen will be transmitted to wild-type egg sacs in the recipient ear shoot. The resulting ear will consist of a collection of F_1 kernels, each containing a set of heterozygous mutations caused by *Mu* insertion.

Simple Mendelian genetics is then exploited to create segregating populations of these insertion mutations. Each F_1 plant is selfed. The resulting F_2 genotypes will segregate 1:2:1 for each insertion event that occurred in the original *Mutator* pollen grain. If we regard most insertion mutations as null alleles, where normal gene expression is silenced, the genetic segregation of mutations in the F_2 thus provides both mutant and wild-type genetic backgrounds. It is from this F_2 segregating family that functional analysis of maize genes is allowed to proceed. In addition, since these *Mu* active lines

contain many copies of *Mu* elements (from 25–100 copies), each pollen grain will contain a theoretical 25–100 new *Mu* insertions. If we assume an average of 50 new insertion events per gamete, and create a population of 40,000 F_1 individuals, the total number of novel insertion events approaches two million. Such a large number of independent transposon-insertion events represents theoretical saturation of all maize genes.

Accessing the Mutant Bank

How does a glut of DNA sequence information feed into the genetic scheme of saturating the maize genome with transposon-induced gene disruptions? The missing piece of the puzzle is how to identify a mutation in a specific gene of interest. The bridge that links DNA sequence information to the mutant bank is achieved by exploiting another Nobel Prize-winning discovery—the Polymerase Chain Reaction.

Polymerase Chain Reaction (PCR) was invented by Dr. Kerry Mullis, who was working on DNA technologies for the Cetus corporation during the 1980's. Mullis developed a method to amplify specific DNA sequences using a simple reaction that could be performed in a test tube (5). The features of this process are shown in Figure 2. Once bound to its partner (or parent) strand of maize DNA, an oligonucleotide "primer" can be extended by a polymerase enzyme that adds complementary nucleotides to the free end of the bound primer. The result is a nascent strand of DNA that is complementary to the original parent strand. The process becomes geometric when two oligonucleotide "primers" are pointed toward one another, each one binding

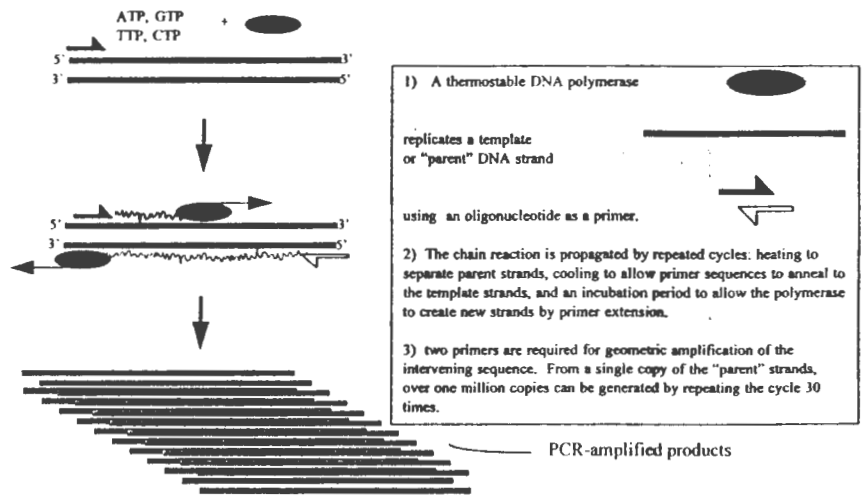


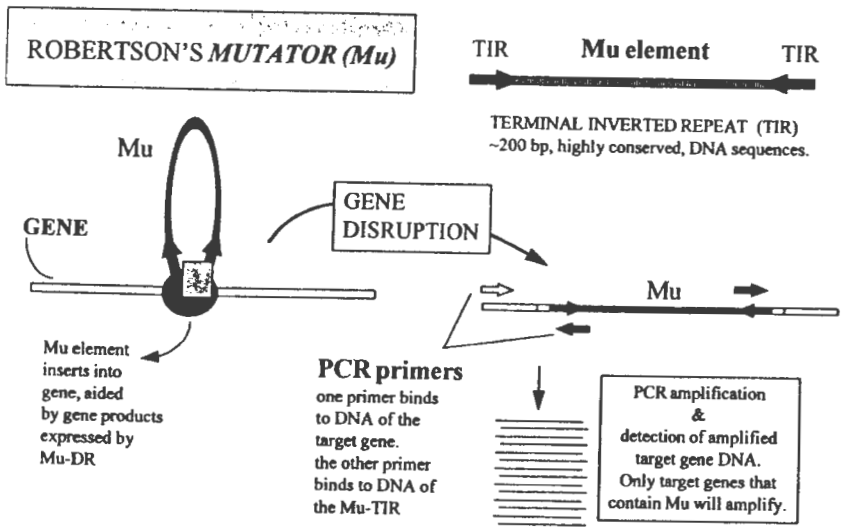
Figure 2 The Polymerase Chain Reaction (PCR)

to opposite parent strands. Repeated cycles of extension by polymerase, followed by melting of the DNA strands at high temperature, produces an amplification process that behaves like a chain reaction: by repeating the cycle 30 times, a single, double-stranded parent strand can be amplified over one million-fold. The result is a high copy number of the targeted sequence; a disproportionate amount that can be readily detected using molecular "probes" that bind specifically to the amplified DNA.

Specific sequences can be targeted for PCR by careful design of the oligonucleotide primers. Primers can be made synthetically by an automated machine which regulates the primer sequence, and produces them quickly, accurately, and cheaply. Oligonucleotide primers can be designed for any DNA sequence. In 1993, Dr. Mullis was presented with the Nobel Prize for his discovery of PCR. This point marks our entry into merging transposon mutagenesis with PCR-detection of specific transposon-disrupted target genes.

Mutator PCR and Specific Mutant Selection

The sequences of the *Mutator* elements are known, so PCR primers can be designed specifically to amplify from *Mu* elements. A consistent feature, shared by each member of the *Mu* family, is a 200 base pair left- and right-border sequence that delimits the ends of each *Mu* element. These border sequences, called terminal-inverted-repeats (TIR), are nearly identical across the entire *Mutator* family. As shown in Figure 3, a single oligonucleotide



can be designed which binds to opposite strands of the TIR. Primer binding occurs with an "outward" orientation, pointing into the DNA that flanks each *Mu* element. This arrangement is critical to the success of this system. Since geometric amplification *via* PCR is a process that requires a pair of primers that face one another, the reaction is facilitated by introducing a second primer that binds to the flanking DNA.

Because *Mu* elements prefer to "hop" into genes rather than repetitive regions of the chromosome, the maize DNA that flanks TIR sequences is typically part of a gene's promoter or coding region. Genomic DNA prepared from *Mu*-containing plants serves as our template for PCR. By controlling primer design for a target gene of interest, PCR can be manipulated to select for specific F_1 plants in which *Mu* elements have inserted in or near the target gene. Amplification of this nature can be detected simply by using a radiolabeled probe specific for the target gene (see Fig. 3).

The Trait Utility System for Corn (TUSC)

The combination of transposon mutagenesis *via Mutator* and PCR-based selection of target-specific *Mu* insertions makes up Pioneer's Trait Utility System for Corn (TUSC). This resource consists of the following: 1) genomic DNA prepared from 42,000 plants (the products of inbred × *Mu*-donor crosses), each containing approximately 50 copies of *Mu* elements; 2) F_2 seed from each of these plants, wherein *Mu* insertion alleles genetically segregate; 3) a set of equipment and technologies to perform high-throughput PCR on all 42,000 DNA samples simultaneously; 4) specialized computer software to speed data analysis and reporting.

The idea is to choose a target gene, design oligonucleotide primers and a probe for that target, and to pair the target primer with the *Mu*-TIR primer in PCR reactions. The products of this screening process serve to identify which individuals in the collection have *Mu* elements inserted into the target of choice. Once identified, the corresponding F_2 seed can be accessed for each PCR-positive individual. These materials provide a genetic starting point to investigate the effect of *Mu* insertion on target gene expression, and the phenotypic consequences of disrupting the target gene on the plant as a whole.

As a directed example, let us assume that we are interested in a gene that is highly induced in maize plants that are exposed to periods of drought. Is induced expression of this gene actually important for plants to survive drought stress? The question can be addressed with a mutational approach. Knock this gene out, and see if mutant plants are impaired in their capacity to endure such stress. By examining its DNA sequence, PCR primers can be designed for our drought-induced gene. Together with a *Mu* primer, each drought-gene primer can be used individually to screen our collection of DNA to identify individuals with *Mu* insertions in our stress-induced target

seed samples for phenotypic analysis. Molecular technologies can be used to identify which F_2 plants are homozygous for the *Mu* insertion in our stress-induced gene. The consequences of the homozygous mutant allele can be confirmed. Let us assume that in the homozygous condition, expression of the drought-induced gene has been nullified. Now, these mutant plants can be compared to non-mutant plants for their capacity to endure drought stress. Does susceptibility to drought cosegregate with the homozygous mutant genotype? If so, this gene and its gene product are likely to be important for drought tolerance. In this way, a functional role for this gene has been inferred by targeting it for disruption. The knowledge gained by mutational analysis permits a researcher to address further aspects of this gene in relation to drought tolerance. For example, what is the endogenous allelic variation at this locus when comparing elite inbreds with different drought tolerances? Is there a difference in the primary DNA sequence of this gene that influences the level of gene expression, or do tolerant alleles contain changes that produce protein products with altered properties? If no natural allelic variation for this gene can be uncovered, is there an aspect of its expression that may be amenable to genetic engineering for drought tolerance? Insights gained by determining the phenotypic roles for genes may perhaps allow research and breeding efforts to be focused in productive directions. This is a large part of the value of mutational analysis.

The main research benefit of TUSC is that it provides a major resource for determining the role that specific genes play in the life of a corn plant. In addition to this, there are three other benefits that TUSC brings to plant research:

I) TUSC assists gene isolation efforts. Transposon tagging relies on cosegregation of a transposon-disrupted allele with a mutant phenotype. The process of confirming the identity of a tagged gene usually relies on the description of more than one mutant allele. TUSC provides an easy way to obtain additional *Mu*-disrupted alleles for any candidate DNA sequence. This was the technique used to confirm a candidate clone for the anther ear 1 gene (*an1*) of maize (6).

II) TUSC may help develop genetic tools. Genetic tools such as male sterility may greatly increase the efficiency of hybrid production. TUSC is a way to address a complex process such as pollen development in order to define genes important for male fertility. Insights gained by dissecting the genetics of pollen development will point researchers to important genes, and promote greater flexibility for engineering male sterility in maize.

III) TUSC may help develop specialty grain traits. Simple genetic variation and null mutations can be exploited to remove undesirable traits from the kernel. Such use of mutation is widely in use now, as illustrated by the *ae1* and *wx1* examples. Waxy-converted lines achieve higher levels of amylopectin by removing the undesirable trait of starch de-branching from the kernel.

Similar approaches can be undertaken to effect altered starch, oil, or protein qualities for specialty product development.

TUSC is an exciting new tool with immediate short-term benefits, and promising long term potential. We have made this resource accessible to the greater research community. University-based researchers can simply request us to identify *Mu* insertion alleles for any target gene sequence of interest to them. At Pioneer, we run the high throughput PCR reactions for them, and send them back samples of the F_2 seed for their follow-up analysis. These services are performed in exchange for Pioneer's option to license important discoveries that come out of this approach. The business value of TUSC is a long-term investment, guided by the principle that mutational analysis of certain genes will identify important opportunities for genetic engineering, breeding practices, and product development. My thanks to the meeting organizers for this opportunity, and I would like to emphasize the availability of more information about access to this resource—simply contact me directly (meeleyrb@phibred.com).

REFERENCES

1. McClintock, B. (1950) *Proc. Natl. Acad. Sci. USA* **36**:344–355.
2. Robertson, D. S. (1978) *Mutation Research* **51**:21–28.
3. Bennetzen, J. L. *et al.* (1984) *Proc. Natl. Acad. Sci. USA* **81**:4125–4128.
4. Chomet, P. *et al.* (1991) *Genetics* **122**:447–457.
5. Mullis, K. B., Faloona F. A. (1987) *Methods Enzymol.* **155**:355–350.
6. Bensen, R. J. *et al.* (1995) *The Plant Cell* **7**:75–84.

Transgenic Technology for Crop Protection

The New "Super Seeds"

by Gerald Carlson, Michele Marra, and Bryan Hubbell

A new biotech revolution is sweeping through agriculture. With it comes enormous potential benefits, but also some potential problems, as a quick reading of the popular press readily shows:

> The new maize contains a gene from a widespread soil bacterium, *Bacillus thuringiensis* (Bt), which makes it resistant to the European corn borer, an insect pest which damages millions of hectares of the crop each year. It has been approved by regulators in a number of countries, including America, Canada and Japan, but in the European Union (EU), where environmental lobby groups have decried it as a "mutant," regulators have so far refused to allow it to be used. ...the EU will be faced with an awkward decision: either to ban all American maize imports (and start a trade war), or to permit the sale of a product which cannot be grown locally.
>
> (*The Economist,* 14–20 September 1996, p. 82).

> This year, for the first time, farmers have planted millions of commercial acres of genetically altered cotton, soybeans, corn and potatoes. The technology has worked surprisingly well, promising a new era of higher yields at lower cost. It is also triggering a stampede for Monsanto's so-called Roundup Ready soybeans and pest-resistant cotton, vindicating the company's years of investment in biotechnology.
>
> (*Wall Street Journal,* 24 October 1996, p. 1).

> Adoption of Roundup Ready soybeans has the potential to decrease herbicide use on soybeans by as much as one-third.
>
> (Monsanto press release, 24 September 1996, pp. 1–2)

> The bollworm dealt an unexpected blow to *Bacillus thuringiensis* cotton this season. Initial reports of bollworm infestations in the premium-priced cotton came from the Brazos Bottom area of Texas.... Did Monsanto promise more than it could deliver with Bollgard?
>
> (*Progressive Farmer Online,* 18 September 1996)

Here we summarize the latest information on the new super seeds, focusing on increased productivity, adoption potential, organization issues for the seed/agri-biotechnology/pesticide industry, and potential external effects of the new technologies. Finally, we offer some economic implications of this new biotechnology for agriculture.

Productivity effects

Bt corn

Overview. Bt corn is a plant which has *Bacillus thuringiensis* (Bt) toxins in many of its cells. This naturally occuring soil bacterium will control insect pests that feed on the plant. The new transgenic corn seed increases yield in regions where European corn borer (ECB) would reduce output. This insect is not well controlled by conventional insecticides because it is sporadic over time and space, and insects are shielded from sprays by boring into stalks. Experiments and field tests have shown that Bt corn plants will reduce ECB damage by about 95 percent, and this will translate into a 4–8 percent increase in yield (Koziel et al.), depending on ECB density.

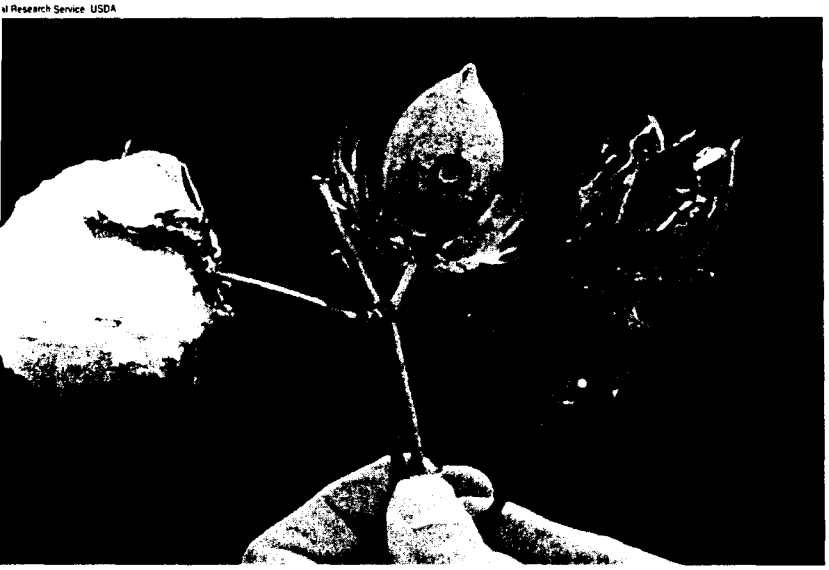
The 1996 experience. Mycogen and Ciba Seeds sold all seed permitted to be sold by the Environmental Protection Agency (EPA) (500,000 acres) at a cost premium of about $10 per acre, or a 30 percent increase in seed cost. ECB levels were lower than normal in many areas, so although Bt corn provided insurance in a year with relatively high corn price, protection values were probably lower

than for the average year.

Prospects for 1997 and beyond. There are four genetically different Bt corn products being sold in 1997. EPA approval of sales by Monsanto, Northrup King, and DeKalb has quickly changed the availability of Bt incorporated into desirable corn hybrids. Seed supply may be sufficient for 6–8 million acres in 1997, but the price premium ($11 per acre) may prevent all available seed from being sold. Universal acceptance of the technology for a 30 percent seed price premium seems unlikely. The 1994 analysis by Ciba Seeds projected adoption on about 50–70 percent of all corn acreage with a seed cost premium of about 20 percent over the course of about five years of full seed supply.

Roundup-Ready soybeans

Overview. In contrast to Bt corn, the glyphosate-tolerant soybeans (Roundup Ready, RR) will not provide pest control without pesticides. This combination pesticide/seed provides a change in weed management from several applications of several active ingredients to a single application of one broad-spectrum herbicide. This technology will lower weed control expenditures, but will slightly increase seed costs. Farmers who use the new soybean seeds must, by agreement, not save and use their own seed and use certain herbicide practices.



Mature cotton boll at left was protected by a gene for Bt; other bolls show damage from cotton pests.

The 1996 experience. The 1996 use of RR soybeans was slightly more extensive than Bt corn because seed was available for about one million acres. Monsanto, the patent holder, charged a technology fee of $5 per acre. There is considerable interest in this technology because it reduced average weed control costs from about $25–$30 for conventional herbicides to $18 per acre—$13 for Roundup and $5 for the technology fee (Fritsch and Kilman). The ability to spray soybeans with Roundup allows farmers to reduce preemergence herbicide applications, avoid some cultivations, and plant soybeans in narrower rows, thereby further crowding out weeds.

Prospects for 1997 and beyond. Seed companies rapidly expanded RR seed production during 1996. Seed for 8–10 million acres, about 20 percent of the soybean crop, was available in 1997 from seed companies working under licenses or agreements with the patent holder (Fritsch and Kilman). Acceptance of the soybeans in export markets has been approved, except in small parts of the European market. (However see the discussion below on possible export restrictions.) The higher cost associated with farmers not being able to save and use their own soybean seed and some additional costs related to the requirement that Roundup must be applied to RR soybeans may restrict adoption somewhat.

Tolerance for other herbicides, including a glyphosate substitute, glufosinate (Liberty), is being pursued rapidly in corn, soybeans, cotton, and other crops. A combination of Bt and herbicide tolerance in the same crop seed has been in process by Dekalb and other seed companies and is available in some cotton varieties this year (Monsanto). Achieving desired trait expression in commercial varieties now proceeds at a pace many times faster than development of the first transgenetic changes.

Bt cotton

Overview. The Bt insecticide, as that contained in the corn seed, is used to reduce insect damage from bollworm, pink bollworm, and budworm. This technology may sharply reduce conventional insecticide use. However, because the Bt toxin is highly effective, insect resistance may develop in a short period, rendering Bt less useful for some insect species of cotton and other crops. To prevent new resistance, the EPA has mandated that a resistance management program be put in place. For 1996, this took the form of either 3.85 percent of each field planted to non-Bt cotton and left untreated with insecticides, or 20 percent of the field planted with non-Bt cotton and use of insecticides other than foliar Bt, along with insect resistance monitoring.

The 1996 experience. Because of resistance development to conventional cotton insecticides in the 1990s, there was considerable interest in Bt cotton. In 1996, Bt cotton seed was planted on over 5,700 farms, or 1.8 million acres (Barton). Farmers paid the $32 per acre technology fee to the patent holder, Monsanto; a seed price premium of about $1.50 per acre; and the opportunity costs of providing the resistance management areas. There were no restrictions on cotton fiber or cottonseed sales. As the *Progressive Farmer* quote above indicates, there were problems with this technology in some re-

gions, as high bollworm and budworm populations led to cotton boll losses. Other areas experienced lower insecticide use and good levels of insect control (Fritsch and Kilman). A survey of eighty-nine Bt cotton users showed an average yield increase of 7 percent compared with conventional cotton (Barton). In addition to the direct insecticide savings, Bt cotton decreased insecticide use which increased beneficial insect numbers, thereby reducing the costs of controlling other pest types (Smith).

Prospects for 1997 and beyond. Like past insect control technologies, this one will require management and understanding. The cotton fruit are exposed for a long period to many potential insect types. The technology fee is high relative to that for soybeans and corn, but the potential insecticide reduction is also larger. The seed technology fees will partially support development of second-generation products which are introduced as pests become resistant, or improvements are made. Bt cotton was not widely adopted in some areas. North Carolina's cotton farmers used it on only 3% of cotton acreage in 1996. Consequently, discounts of $10 per acre on the technology fee were offered on the first 50 acres per farm at the beginning of the 1997 season (J.R. Bradley, professor of entomology, North Carolina State University, personal communication, April 1997).

Adoption issues

Who will adopt?

Following previous work on adoption, the early adopters should be farmers with high pesticide costs, those suffering from higher pest damage, and those who can better utilize other inputs which are complementary to the transgenic crops.

Usually, we think of early adopters of agricultural production technologies as those having more human capital, but this result may not hold in the case of transgenic crops. The transgenic crops seem to be easier to manage than the current crop/pest control methods. The pesticide spray decision is either irrelevant or simpler in the case of the Bt crops, and glyphosate-tolerant crops require only one spray and have a wider application window in most cases. So early adopters may include those with lower-than-average skills for managing variable pest populations. However, skilled managers can be expected to assess the profitability of the technology better in marginal cases, assimilate more information on local suitability of seeds, and process more complex information on multiple (transgenic and other) traits of seeds. Likewise, farmers with larger crop acreage will appreciate the lower management time requirements and may be among the early adopters.

Where will adoption occur?

It seems reasonable to expect areas with higher pest infestations and more severe resistance problems to be early targets for seed development and farmer adoption. For cotton, bollworm and budworm resistance to pyrethroid insecticides has been particularly acute in the Midsouth region. The Colorado potato beetle is resistant to many types of insecticides used in potato production in the eastern states. These were areas of widespread adoption in 1996.

European corn borer populations are sporadic over years and regions; however, they are more frequent pests in the western Corn Belt. There are usually two generations of this pest per season, and Bt corn seems effective against both generations. Crop value protected is highest in the high-yield, irrigated corn regions of the Plains states. Spatial availability of Bt corn seed is affected in 1997 since the EPA limited total sales in cotton-producing counties to help prevent resistance development for the corn earworm. In addition, European restrictions on Bt corn imports, brought on by consumer safety concerns, may limit adoption in some regions.

Agricultural Research Service, USDA

Entomologist Hollis Flint compares an insect-ravaged cotton leaf from a control variety with one that has been genetically engineered with a protective gene from *Bacillus thuringiensis*.

Glyphosate is an effective broad-spectrum herbicide against both broadleafs and grasses. Roundup Ready soybeans should be adopted first in production zones where this weed combination requires a relatively high degree of control, such as in most areas of the South. Glyphosate and glufosinate can be used as burn-down herbicides in reduced-tillage systems. Farmers practicing no-till are likely adopters. Likewise, broad-spectrum herbicides with low cost may fit the weed spectrum on land coming out of CRP, and Roundup Ready crops may be widely adopted on these lands.

What crops and technologies are being tested?

The crops and genetic traits now being field tested provide an indirect picture of what to expect in the next three to five years. Figure 1 shows USDA Animal and Plant Health Inspection Service (APHIS) data on numbers of separate field trials of genetic products by transgene category. These trials are approved by APHIS and conducted by the private seed/pesticide companies or by university contractees. Herbicide tolerance, insect resistance, and product quality account for most of the growth in tests. However, the growing work on virus and fungus resistance is an important new trend. Even testing of pharmaceutical and industrial properties (such as plastic producing cotton) began to appear in 1994.

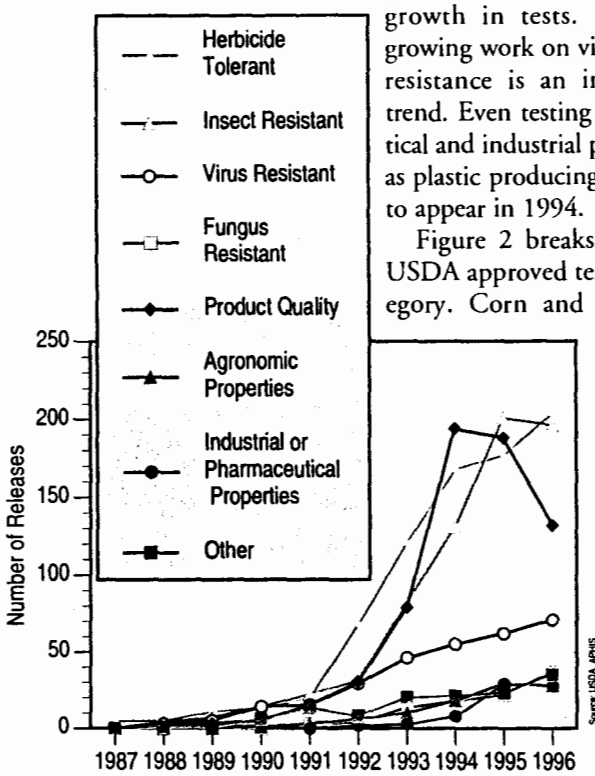


Figure 1. Trends in field releases by transgene category

Figure 2 breaks out the same USDA approved tests by crop category. Corn and vegetables account for most of the tests since 1993. Interestingly, the number of tests on cotton, corn, and soybeans declined in 1996. Almost half of all field tests for herbicide tolerance was conducted on corn during 1996 (USDA).

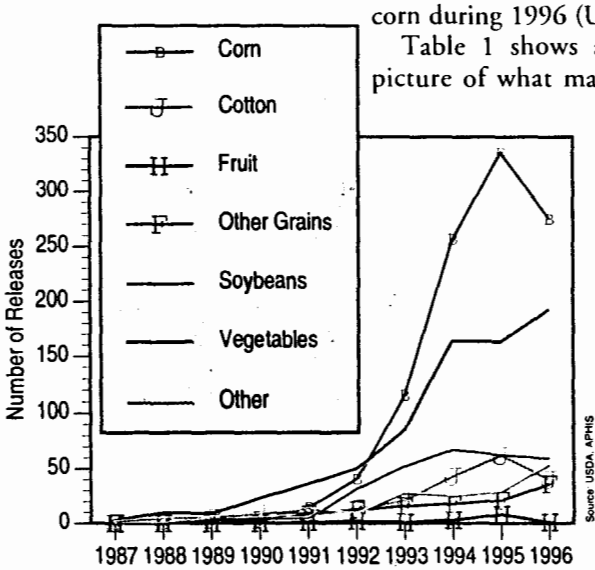


Figure 2. Trends in field releases by crop category

Table 1 shows a more direct picture of what may lie ahead in transgenic technology. Monsanto plans to increase its offerings of crops and genetic traits. These new products will use multiple genes for herbicide tolerance and insect control in the same plant, and second generation Bt proteins to combat insect resistance.

Aggregate supply and adoption

Roundup Ready soybeans may prove to have a comparative advantage over other crops such as corn. The pest control cost savings with little change in yield could make these beans more profitable than corn at prevailing output prices. Therefore, some regions may experience an acreage shift from corn to soybeans, particularly in the South where corn yields are much lower than in the Midwest.

The availability of transgenic seeds for other growing regions of the world will surely come. Monsanto, with Delta and Pine Land Seed, will sell Bt cotton seed in China in 1998, and Australia will slowly increase its Bt cotton over the next few years (Monsanto). U.S. farmers may have a few years of lead time in the biotech crops over their competitors and may see a short-run gain in crop export shares.

Organization of the seed/pesticide industry

Different industrial structures usually evolve when dramatic technological changes occur. Crop germplasm and pesticide supply need closer coordination with the transgenic crops than with conventional pesticides and crop varieties. Biotech firms will take organizational and marketing steps to improve scale economies, expand sales of the new technologies, and increase profits.

Pesticide and seed industry firms are making many different organizational and marketing changes. Some examples are listed in table 2, with few firms involved with the actions at the top of the list and more market-oriented strategies at the bottom. Acquisitions and mergers are self-explanatory, but the purchase of genetic resources may involve purchase of patents, research expertise, or access to gene libraries. In some cases seed companies make the purchases, and in others the pesticide firm has been the purchaser.

Exclusive agreements restrict the spread of genetic resources more than nonexclusive licenses or agreements. However, the exclusive agreement can allow access to marketing and other resources as well as genetic ones. Access to local seed producers who tailor varieties to local growing conditions will remain important. The "low-cost seeds" strategy is novel for agriculture, but analogous to the strategy of software companies giving away products to developers so that a product can become widely used and seen as an industry standard. The technology fee charged to final users is relatively new in agricultural input industries but is used so farmers know the component prices of the genetic traits. The combining of technology fee and seed price repre-

sents the traditional pricing of genetic improvement. The tie-in sales strategy is only possible if a firm has some control over the two products—in this case the herbicide and the herbicide-tolerant crop seeds. Regional price discrimination promotes early adoption by charging a lower price where demand is more elastic and can be used effectively only with region-specific varieties.

The above actions tend to convert separate seed and pesticide industries into a combined industry (Seghal). Not all of the agreements and relationships are friendly. Conflicts arise as private firms attempt to rapidly enter these markets and keep others out. Currently, there are eight major lawsuits involving use of Bt. A recent edition of the *Information Systems for Biotechnology News Report* outlines the major issues and parties involved in these suits (Klein).

The primary role of the public sector in the deployment of the new seeds has been to oversee public safety. USDA's APHIS must approve initial field tests of new genes. The EPA under its pesticide authority must approve commercial use of "plant pesticides," and it has assumed the role of approving the commercial release of transgenic seeds. For example, the EPA has limited Bt corn sales and placed geographical restrictions on deployment of the Bt technologies to prevent resistance development. Reduction or delay of onset of negative external effects is the most credible rationale for these government interventions.

Potential externalities

Transgenic crops present two major types of negative external effects: (*a*) more rapid deployment of resistance in pests which damages both the target crop and surrounding crops, and (*b*) possible toxins in food produced from these genetically altered crops. Other safety concerns include development of herbicide tolerance in weedy relatives of the transgenic crops, spread of pesticide tolerance to bacteria or other human pests, and spread of allergens.

Observers often overlook the positive externalities associated with the new seeds and associated pest management. The new seeds will reduce resistance development to conventional pesticides thereby reducing the future dosages required to achieve acceptable control and the cost of replacing these now less effective chemicals. This affects pest control on both the transgenic crop and on other crops which these pests attack. The new seeds may also reduce negative external effects as amounts, movement, exposure, and toxicities of conventional pesticides are reduced. There is some evidence that Bt toxins will reduce crop problems, such as microtoxins on corn, in addition to the ECB. In some Bt cotton fields in 1996, beneficial insect populations increased because of farmers' reduced insecticide sprays (Smith, Wilkins). Transgenic crops may also reduce the run-off from foliar-applied relative to soil-applied herbicides. Transgenic crops may encourage the adoption of complementary, environmentally friendly techniques, such as conservation tillage, that reduce run-off problems. And finally, transgenic crops may enhance existing area-wide pest management strategies such as the boll weevil eradication program (Smith).

Europe has the most pronounced commitment, presumably to protect consumers, against geneti-

Table 1. Planned transgenic crop introductions by Monsanto Corporation

Expected Launch Year	Transgenic Crop
1997	Insect-protected tomatoes Insect- and Y-virus-protected potatoes RR cotton Bt corn RR and Bt cotton
1998	Insect- and virus-protected potatoes RR corn BXN herbicide-tolerant and Bt cotton
1999	RR oilseed rape Second generation Bt cotton Virus-protected tomatoes
2000	Insect-protected corn (corn rootworm) RR sugar beets
2001	Disease-controlled potatoes (fungal diseases)
2002	Boll weevil-protected cotton Disease-controlled strawberries
2003+	Higher-yielding corn Improved-quality potatoes Naturally colored cotton

Source: 1996 Monsanto Annual Report
Note: Commercialization depends on the successful completion of such factors as research, field trials, and regulatory approval.

Table 2. Organizational and marketing changes in the pesticide/seed industries related to transgenic crops

Strategies	Examples (Affected Transgenic Crops[a])
Organizational	
Acquisitions	Monsanto acquires Calgene, Asgrow, Agricetus (1, 2, 3)
Mergers	Ciba and Sandoz become Novartis (Ciba Seeds and Northrup King) (1)
Purchase genetic resources	Pioneer acquires gene libraries from Mycogen (1)
Marketing	
Exclusive sales agreement	Monsanto and Delta and Pine Land (2)
Nonexclusive licenses	Mycogen and Cargill and other seed companies (1)
Distributing seeds at low cost	Agro Evo and most corn seed companies (4)
Separate technology fee	Monsanto (1, 2, 3), DeKalb (1)
Combined seed and technology pricing	Novartis (1)
Tie-in sales	Monsanto and farmer licensees (3, 5)
Regional price discrimination	Monsanto (2)

[a] 1 = Bt corn, 2 = Bt cotton, 3 = Roundup Ready soybeans, 4 = Liberty Link corn, 5 = Roundup Ready and Bt cotton.

cally altered crops. Consumer and farm interest groups and government leaders have discussed mandatory labeling and import bans. Limited exports of RR soybeans to the EU were approved in the summer of 1996, but Bt corn is facing stiff opposition. Environmental groups, including the Union of Concerned Scientists, Jeremy Rifkin, the Consumers Union, and Greenpeace, have been vocal in their criticism of transgenic crops.

Several approaches might be used to address food safety concerns, including import bans, mandatory labeling, voluntary labeling of unaltered products, deployment in limited quantities, information on relative toxicity, and product price discounts for genetically altered foods. In the United States, mandatory labels are required by the Food and Drug Administration only if the product contains a known allergen or has a composition significantly different than the standard crop (such as high-laurate canola).

Approaches for reducing externalities and transaction costs require more information on risks and benefits, and this seems to be a shortcoming at this time. Some of the approaches will have substantial costs; for example, mandatory labels necessitate separate marketing channels for biotech crops. Mandatory resistance management programs impose costs on farmers and seed/pesticide firms. Resistance management costs may exceed benefits if new transgenic versions of crops are forthcoming. On the other hand, the Consumers Union claims that Monsanto's resistance management plan (a high dose of Bt in the plant plus the refugia set-asides) failed against the cotton bollworm in 1996. They argue further that the EPA must act immediately to restrict plantings of Bt cotton to only the modest experimental plots required to continue research on resistance management (Benbrook and Hansen).

Gerald Carlson and Michele Marra are professor and associate professor, respectively, in the Department of Agricultural and Resource Economics at North Carolina State University. Bryan Hubbell is assistant professor in the Department of Agricultural and Applied Economics at the University of Georgia.

Summing up

These seed technologies, while not silver bullets in pest control, do expand the pest management arsenal through both substitute and complement relations to existing pest control approaches. The crop protection experience with the new traits will provide information about the technical and economic potential for other traits, for other crops, and for many regions of the world. The organizational structures being adopted in the seed/pesticide industries are helpful for rapid tailoring and marketing of these products to heterogeneous production areas. The external effects related to transgenic crops are both positive and negative and point to the potential demand for more consumer education and economic evaluation. Because these technologies are output expanding, and may lower costs of production for export crops, there may be lower food and fiber costs without decreasing U.S. farm income.

For more information

Barton, G. "Bollgard™ Cotton Update." Monsanto Corp., St. Louis MO, March 1997. Available online: http://www.monsanto.com/MonPub/InTheNews/Articles/97-03-20BollgardUpdate.html/

Benbrook, C., and M. Hansen. "Return to the Stone-Age of Pest Management." Remarks before EPA public meeting on behalf of the Consumers Union, March 1997. Available online: http://www.pmac.net/stoneage.htm/

Fritsch, P., and S. Kilman. "Seed Money: Huge Biotech Harvest Is a Boon for Farmers—and Monsanto." *Wall Street Journal.* New York: Dow Jones and Co., 24 October 1996.

Klein, J. "Litigating Bt Patents for Market Share." Information Systems for Biotechnology (ISB) News Report, April 1997. Available online: http://www.nbiap.vt.edu/news/1997/news.rpt

Koziel, M.G., et al. "Field Performance of Elite Transgenic Maize Plants Expressing an Insecticidal Protein Derived from Bacillus thuringiencis, 1993." *Bio/Technology* 4(1993):194–200.

Monsanto Corp. 1996 Annual Report, St. Louis, MO.

Seghal, S. "Patents Are Driving the Restructuring of the Seed Industry." *Seeds and Crop Digest,* November 1996.

Smith, R. "Alabama's Bollgard Gamble a Good Bet." *Southeast Farm Press,* Raleigh NC, 18 September 1996.

U.S. Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS). Biotechnology Permit Database. Washington DC, various years. Available online: http://www.aphis.usda.gov/bbep/bp/

Wilkins, T. "Performance of Bt Cotton: Status Report." *Cotton Rev.* 41(1997). Available online: http://www.uckac.edu/danrscr/Vol.41.htm/

Critical Reviews in Plant Sciences, 16(3):231–251 (1997)

Scientific and Public Perception of Plant Genetic Manipulation — A Critical Review

D. Boulter

Department of Biological Sciences, University of Durham, South Road, Durham DH1 3LE, U.K.

Referee: Dr. Phil Dale, John Innes Centre, Colney, Norwich NR4 7UH England

ABSTRACT: Scientific views on the importance and potential risks of plant genetic engineering are summarized and the ecological and economic consequences of the large scale release of engineered crops is discussed. The difference between scientific risk assessment and the public's perception of risk due to modification of the former by outrage factors is explained. Polls, media, and literature sources show that the public supports plant genetic engineering activity, but has some concerns. These concerns fall within two categories: (1) it is risky, (2) it is morally wrong. Polls show a lack of knowledge about GE by the public and that outrage factors play a major role; their effects on perception are described in detail. Ethical considerations also exercise a powerful influence over GE's acceptance, but the complementary roles of Religion and Science are emphasized. Different attitudes to nature are described and it is concluded that if the world is to be fed and life support systems sustained, active management of nature is required.

Lastly, the need for scientists to communicate with the public and the basis of effective presentation is discussed.

KEY WORDS: risk, perceived risk, "outrage reaction", ethics, eco-systems, scientists' role.

1. INTRODUCTION

Unlike in some periods in the past when science and technology were seen as the source of endless blessings, the public now view new science and technology with caution. The benefits that a new technology may deliver are not yet apparent and there is a lack of knowledge about the scientific facts and uncertainty of the potential risks involved (Macdonald, 1991); the public are now aware, from the examples of nuclear power and the chemicals industry, that new technologies can bring dangers that were not fully realized at the time of their introduction.

In a modern democratic technological society, therefore, government and industry are expected by the public to show that the means adopted to protect them from danger are more than adequate to keep pace with technological innovation; when they are perceived not to do so, the result is a complete loss of public confidence (e.g., the recent BSE "crisis" in the U.K. and Europe generally). Regulations should be enforced that reflect the public's wishes so that new technology can develop at the optimum pace, but with socially and ethically acceptable targets (C.E.C., 1993). Modern plant genetic engineering (PGE)(manipulation) is a case in point, being a new technology that dates from the early 1970s. Surveys (see Zechendorf, 1994) indicate a measure of support for PGE by the public, but also some

0735-2689/97/$.50

concerns. It is important, therefore, that scientists should be aware of the public's perception of their activities and of the need for them to communicate effectively with the public, because they are best placed to present the scientific facts and likely developments accurately; scientists and their employers should now realize that scientists have, as individuals, a social obligation in this regard. Otherwise, the worth of their work could be disparaged and acceptable progress inhibited by unreasonable regulations and views (Miller, 1993), or, alternatively, delays could occur in curbing undesirable applications. A classical example of the role of the public in this latter regard is the curtailing of the indiscriminant use of chemical pesticides following the publication of Rachel Carson's book *The Silent Spring.*

This article, addressed to plant scientists, presents those aspects of PGE and some of the likely developments that relate to the potential risks of this new biotechnology. It also documents the public's perceptions of the dangers of PGE and examines briefly the bases for these views in order to suggest how and what scientists should do about these issues. In covering this topic, it is necessary to include apart from the science, some information and ideas from various other disciplines such as religion, psychology, and literature.

II. WHAT IS PLANT GENETIC ENGINEERING?

The broadest definition of PGE is, changing the genetic make-up of plants to provide plants, plant products, and processes for our needs. In this sense, PGE has been around for a very long time. By the beginnings of agriculture, more than 10,000 years ago and even before, man has been changing the genetic make-up of plants. At first by unconscious selection, then more recently in this century by conventional plant breeding made more effective by applying the knowledge of genetics and by developing a seed industry (older PGE). In the last two decades, these technologies have been complemented by using laboratory-based recombinant DNA technology (modern PGE). The public often does not appreciate the long, safe history of older PGE in producing the food crops and drinks we consume and without which modern civilization would not have been possible.

In order to change the genetic make-up of a plant, genes have to be added and/or inactivated. Modern PGE (abbreviated to GE) uses a different method of gene transfer as compared to conventional plant breeding (see Table 1 in Boulter, 1995). The main differences between the two methods are the greater technological input, use of a wider gene pool, and random chromosome location of the transferred gene(s) in GE compared to conventional breeding. A transferred gene incorporated into a chromosome in an indirect way (i.e., other than by homologous recombination) may inactivate or activate other genes, giving additional effects to those expected from the product of the gene alone. In this connection, a distinction is made between coding and regulatory sequences, the former producing a specific gene product (protein of known function) and the latter determining the activity of a gene. Normally, a hybrid gene is transferred with its own regulatory sequence replaced by one of a few well-characterized regulatory sequences designed for expression in the particular host. However, if the transferred gene encodes or affects a regulatory molecule (transcription factor, hormone), the activity of other genes could be affected (pleiotrophic effects). The inactivation or enhanced activation of individual genes can and does occur on crossing in conventional breeding and so these processes occurring from GE are not generally expected to produce phenotypes hitherto

novel to the species, but the incorporation of novel promoters with expression patterns unusual for the species may induce patterns of gene expression that would not be expected to happen by conventional plant breeding. The products of genes can also interact with one another (epistasis), but there is no biological reason to suppose that epistatic effects would be greater with engineered plants. However, with GE, DNA sequences foreign in the history of the species are being introduced into the chromosomes, and novel fusion genes and gene products could be created by random insertion to give novel, unpredicted properties. Another source of increased variation is the technological operations (e.g., a tissue culture step) of GE, but the variation seen so far due to both these causes is not great compared with the offspring variation encountered on crossing nonidentical parent plants, and there is no reason to presume that unpredictable variation from GE is any more likely to pose a hazard than variation emerging from sexual crosses. In most applications, the known identity of the genes inserted by GE and the ability to analyze their structure, activity, and stability make predictions of new variation introduced by a transgene an easier and more accurate process leading to a shortening in the length of the associated breeding program to generate a new variety compared with conventional breeding. Just as genetically engineered modification of crops results in the production of "novel" individuals, genetically modified organisms (gmos), so does sexual crossing through recombination of genetic variation. In both processes, unwanted novel plants can be identified and eliminated in early (small-scale) field trials, but not necessarily all unwanted genotypes. The process of gene transfer itself, therefore, while different, would appear to pose no foreseen additional risks and it has generally been accepted that genetically engineered organisms should be evaluated and regulated on their phenotype (by product not process) (Tiedje, 1989). However, it is important that this difference in process should be borne in mind with each new gene application and it is important that this aspect of GE has been questioned because when a new technological method is substituted for an older one, occasionally risk occurs.

Apart from these considerations, the consequences of releasing any novel organism has to be considered. These consequences are not predictably precise in all environments and the need to monitor any undesirable side effects should be considered (see later discusson). Genetic engineering can now be applied to virtually all crops (Fraley, 1992; Kung, 1993; Dale, 1994), and gene mapping and sequencing of many major crop plant genomes are in progress, promising to identify many additional useful genes in the near future and to facilitate marker-assisted breeding (Sobral, 1996).

III. WHY IS GE IMPORTANT?

The most far reaching importance of GE will be its use as a powerful new research tool to help understand biological processes. It impacts strongly on all classical biological disciplines from taxonomy and ecology to physiology, genetics, and biochemistry. Furthermore, it has great applied significance.

Politicians (House of Lords, 1993), industrialists (Fraley, 1992), farmers (National Farmers Union, 1995), and scientists (O.E.C.D, 1992a) all identify GE as a means of facilitating the provision of more nutritious, healthier, and safer foods, pharmaceuticals, increased agricultural productivity in poor environments, agriculture sustainability, and help to reduce damage to the environment (Best, 1995). Its diffusion is predicted to be widespread in the agrochemical, seeds, and food industries (O.E.C.D. 1992a). In 1996, one can point to an established, global

plant biotechnology industry in which many companies have products entering or about to enter the market place. These involve a wide range of crops, and in general set out to provide better quality, healthier, and better value foods thereby also reducing the costs of therapeutic and health care. Some examples are herbicide-resistant cotton; engineered cotton fibers; increased flavor vine-ripened (Flavr Savr) tomatoes; tomato pastes; high laurate oil rapeseeds; bioproduced chymosin; biofungicides, biopesticides, and biobacteriacides of many crops; high solids potatoes and tomatoes; and sweeter peppers.

IV. HOW IS GE REGULATED?

With the advent of GE in the early 1970s, scientists appreciated the power of gene transformation and called for a moratorium until their safety concerns could be investigated under strictly controlled conditions. When the worst scenarios were not realized, GE went ahead world wide under a set of guidelines that are now embodied in legislation. For example, in Europe (following a series of reports from the O.E.C.D., 1986, 1992b, 1993), gmos are regulated by the European Directives 90/219 (contained use) and 90/220 (deliberate release), which are implemented, for example, in the U.K. by the Genetically Modified Organisms (Deliberate Release) Regulations 1995, under the Environmental Protection Act, 1990. (Similar regulations exist in the U.S., Japan, and many other countries [12].) Safety is sought by the application of risk/safety analysis and risk management. Risk/safety analysis is based on the characteristics of the organism, the introduced trait, the environment, the interactions between the organism and the environment, and the application (for example, see Tiedje, et al. 1989; D.N., 1991). Information about any of these factors provides "familiarity," which is an important part of the assessment. Familiarity is defined as having enough information to judge the safety of the use, or to indicate ways of managing the risk; low levels of familiarity may be compensated for by appropriate management. Sufficient information may be known of the science and/or from previous experience of the organism/trait/environment that a risk assessment suggests little risk. On the other hand, this may not be the case and monitoring of specific possible effects is called for. In yet other applications, more data are required before proceeding with large-scale trials.

Regulations, although adopted in many countries, are nevertheless designed for the national situation, as have been the scale of performance trials and small-scale release experiments. These days, plants move across national boundaries on a global scale (e.g., an engineered potato may not be a risk in the U.K. where there are no sexually compatible wild relatives, but may not necessarily be risk-free in Peru where there are). It is appropriate, therefore, that the second Conference of the Parties (COP) to the Convention on Biological Diversity will consider, among other things, the need for global biotechnology regulation (Lex, 1995).

V. THE SCIENTIFIC DEBATE ABOUT POSSIBLE RISKS OF GE

Before considering the public's concerns about GE, it is important to summarize the views of scientists.

Apart from the process itself leading to unintended genetic and phenotypic variability, scientists (Tiedje et al., 1989; Karieva and Parker, 1994; Rogers and Parkes, 1995; Gates, 1995) have identified the following concerns: the possibility of engineered crops becoming weeds by seed transfer; the escape of transgenes by cross-pollination with wild relatives and related species giving rise to

superweeds that cause damage to the environment and costs to agriculture; loss of biodiversity; changed agricultural practice with attendant social implications (e.g., dependence on F1 hybrids); and the inadvertent production of toxins/allergens in new plants used as food.

A large body of scientific information related to these risks is available from the performance trials of conventional agriculture, studies on the ecology of natural plant populations, the behavior of equivalent genes to the "transgenics" of GE in conventional breeding, the results of large numbers of small-scale field trials specifically designed to examine some safety concerns, land use and sociological effects of high-tech agriculture, and data on food toxins and allergens. In attempting to analyze this database, the approach generally has been to identify a hazard (e.g., the generation of a superweed by cross-pollination), and to calculate risk (i.e., the chances of damage from the hazard occurring). In this particular case, the risk would depend on the likelihood of pollen transfer occurring "multiplied" by the consequences if it did. The likelihood of pollen transfer would depend on the reproductive characteristics of the crop, the existence, if any, of potentially sexually compatible wild relations and conspecifics, the population of these in the vicinity, and the presence of vectors for pollen and seed persistence. The consequences of the transfer would then depend on the biological properties conferred on the host, for example, by the type of gene (see Rogers and Parkes [1995] for comments on the use of many transgenes), whether a recipient plant would show enhanced competitiveness or reproduction success as a result, and, if so, would this lead to "weeds" that could cause economic loss or environmental damage. An example of this approach is that of Ahl Goy and Duesing (1996), who analyzed the results of 391 field trials conducted in Europe up to the end of 1993. They concluded that the results of 91% of the trials carried out in Europe with various genes/crops were likely to have a minimal, if any, potential impact on the environment; the remaining 9% of the trials show a low potential environmental impact (but see later discussion).

Most experiments have been done on the possibilities for transfer of pollen or seeds, either with performance trials of conventional agriculture or using specially designed small-scale trials, although the latter can be criticized in that the methods used would often not detect the rare event and that many trials were designed under containment regulations (see Regal, 1994). Information exists on many crops at least growing in some ecogeographic locations (Economidis, 1992; Jones, 1994; Darmency, 1994; McPartlan and Dale, 1994; Bright et al., 1996, Roger and Parkes, 1995). Similarly for many crop plant/locations the occurrence of wild relatives and conspecifics is known and classification into high, medium, or low of the chance of hybrid formation proposed (Raybould and Gray, 1993). Thus, for the U.K., Raybould and Gray (1993) divided crops into group 1, minimal probability of gene flow (e.g., potato, maize, wheat, tomato); group 2, low probability of gene flow (e.g., rape, flax, lettuce, barley), and group 3, likely probability of gene flow (e.g., sugar, beet, carrot, and forage grasses). Similar information exists for the Netherlands (de Vries et al., 1992) and Europe generally (Brown and Crawley, 1991). However, although the results so far are in the main reassuring, more data are needed from experiments designed and analyzed on the basis of ecological principles (Kjellson and Simonson, 1966).

Much less information is available on the "consequences of transfer" (Weverling and Schenkelaars, 1992). So far, release experiments have concentrated on genes for pest and disease resistance, herbicide tolerance, stress tolerance, and enhanced plant

food quality. What would be the influence of these genes in hybrids with regard to invasiveness and competitiveness in natural populations? In some cases, the same genes have also been used in conventional breeding programs and some useful information exists, but this is not always the case (e.g., Bt for insect resistance) and in the future the number of such "novel" genes will increase. Furthermore, there is still no agreement on what qualities and how many a plant needs to be weedy (Williams, 1993; Baker and Stebbens, 1965; Perrins et al., 1992a,b). Accordingly, experts cannot agree on the extent of the risks involved. Environmental impact of an "escaped" gene would depend not only on the gene, but also on a variety of abiotic and biotic factors, depending on the locality of the released gene. Once again, there is considerable disagreement by experts on the importance of these various aspects and more data are needed.

A detailed review of existing data and its interpretation is given in Rogers and Parkes (1995) and will not, therefore, be repeated here. It can be concluded that the use of engineered crops on a commercial scale will lead to some feral crop establishments and some hybrids being formed in some wild populations (i.e., there will be some risk that cannot be precisely quantified). Furthermore, we have to accept at this stage that there is an element of subjectivity in risk assessments. In 1989, the Ecological Society of America produced a comprehensive report that supported the use of advanced biotechnology for the development of environmentally sound products and concluded that the careful design of transgenic organisms, together with proper planning and regulatory oversight, would ensure that these new organisms would pose little or no ecological risk.

The question is, therefore, do existing regulation and management practices reduce this risk to an acceptable level? After all, the public accept that life carries some risk, but require that the risk be acceptable (i.e., would rather accept the risk than lose the benefit of an application and, as we have seen, most people accept that there are considerable benefits to be obtained from using GE). The public expect biologists to have performed sufficient proactive, precautionary research to avoid problems of the type subsequently found in the chemical and nuclear industries, which were not foreseen at the time the technology was introduction.

For many scientists, existing data (although incomplete), experience, and biological considerations suggest that, provided trials are properly managed, it would be acceptable to proceed with large-scale field trials (applications) in most cases under the latest regulatory systems, as exemplified by the U.S. and U.K. (Ward, 1996a), because: (1) only large-scale risks will give answers, and (2) considering the risk (loss of benefit) of not proceeding. Monitoring for any environmental effects, where appropriate, should be mandatory (NFU, 1995), which, because it is long term, will be costly. Farmers themselves could help in this activity.

Proper management is the key; in cases where no compatible wild or weedy relatives exist locally, safety is not at issue locally, but may be so globally. In other cases, where pollen transfer to sexually compatible plants would occur, even if rarely, or the nature of the trait to be transferred was of concern (e.g., insect resistance or herbicide resistance [see details in Rogers and Parkes, 1995]), then it would be necessary to ensure that management practices were used so as to reduce the risk to acceptable levels. A wide range of appropriate practices for this purpose already exist as a result of performance trials of conventional agriculture. In some cases, further work is required before proceeding to large-scale field trials (e.g., use of some virus nucleotide sequences for tolerance to viruses [McGarvey and Kaper, 1993; Tepfer, 1993]). A majority of present (near future) applications are to improve food

quality by switching off genes (antisense) or by modifying metabolic pathways (Ahl Goy et al., 1994). Regal (1994) has pointed out the potential in this type of application for inadvertent toxin or allergen formation, but this problem is also well known from conventional breeding. Antimetabolites in the past have had to have been bred out of many crop plants and the types of antinutritional or toxic compounds that need to be monitored in transgenic plants are known (Hahn et al. 1984; D'Mello et al., 1991).

In contrast to this governmental and scientific approach, is the attitude of some, but not all, environmental groups. Greenpeace, for example, who says that GE should be stopped as risky, without supplying specific detailed data to prove (or even suggest) the level of the risk. This is because they consider any risk unacceptable as they claim GE is unnecessary because alternative practices (e.g., multi/inter-cropping) could also feed the growing world population. However, where is the evidence for this assertion? Although this is not to deny that integrated agriculture has an important role to play and has, for various reasons, been neglected. More recently, Greenpeace considers that GE should be stopped until international GE regulations are in place for all countries, especially developing ones.

VI. PUBLIC PERCEPTION — WHAT ARE THE FEARS/CONCERNS ABOUT PLANT GENETIC ENGINEERING?

The public's views about PGE can be judged, to some extent, from letters published by newspapers, from media articles and programs, and from portrayals of science and scientists in the literature and in films. However, these views are highly selective and subjective and a more reliable indication is the results of polls. In 1994, Zechendorf undertook a review of the many polls (>20) on what the public thinks about biotechnology, which included views about PGE. It was not possible to compare one survey with another directly as the questions and methods used were not the same in different polls. Even in the same poll, different questions with some common content led to logical inconsistencies and contradictions (e.g., in a single poll, a majority could accept the production of pharmaceuticals in animals, yet would also consider genetically engineering animals unacceptable). Nevertheless, although unable to make quantitative comparisons, he was able to discern significant trends. There was a high level of interest in scientific matters, but a low level of knowledge of biotechnology and PGE. There was a large measure of support (a majority) for research on PGE, more so than for animal and human genetic engineering. Approval or otherwise was based on perception of products, not the process (also see Hallman, 1996). Concerns were expressed about some potential risks and some ethical issues. A recent poll (Hallman, 1966) of American residents in New Jersey has suggested that the public has not yet made up its mind fully about GE. This conclusion, if general, is of great importance because as products reach consumers in increasing numbers, the efforts of opponents to influence customers will intensify.

The public concerns about PGE are (1) it is risky, and (2) it is morally wrong; knowledge, risk perception, and ethical views all influence the degree of acceptability of PGE.

A. GE Is Risky

1. Actual vs. Perceived Risk

A hazard is a situation that poses a threat, whereas the risk associated with that hazard is the chance of it causing harm within a certain period of time. With every technology, there is a level of associated risk. Risks

are often expressed as the chance of causing some number of deaths or illnesses, or as quantified economic or quality of life loss, etc. These risk numbers are important in the management of hazards and are the main scientific tool used, but this is only a part of the story so far as the public themselves are concerned. For the public, the situation is often more influenced by perceived risk than by risk as defined above; perceived risk is risk modified by a so-called "outrage reaction" (Sandman, 1992). Whereas risk is expressed in numerical and technical terms, perceived risk and outrage involve the many social and cultural dimensions of hazards as well as the qualitative aspects of the hazards. Confusion between these two different concepts of risk has led to claims (and counterclaims) by scientists that risk is a scientific concept and by sociologists that it is a sociological one. The influence of the calculation of risk in that of perceived risk varies, and it is often difficult to assess this, as well as its part in the achievement of safety.

The components of the outrage reactions are many, but the main ones are familiar vs. unfamiliar, more knowable or less knowable, diffuse or concentrated, nondreaded or dreaded, voluntary or involuntary, nonmemorable or memorable, natural or artificial, fair or unfair, controlled or uncontrolled, morally irrelevant or relevant (see Sandman [1992] for definitions and explanation of these terms). In general, due to an outrage reaction, the public underestimate the risk or perceive a high level of risk tolerable if the first of the pair of opposites applies, and overestimate it if the second of the pair applies. For example, most members of the public "underestimate" the risks associated with car travel and overestimate the risks associated with air travel, similarly they underestimate the risk from radiation when sunbathing and overestimate the risk of irradiated food.

Polls show that the public are not well informed about GE and in the absence of knowledge, outrage factors play an important role in public perception, and their effects are discussed in some detail.

2. The Influence of the "Outrage Reaction" on the Perception of the Risk of GE

a. Familiar vs. Unfamiliar

Familiarity in the present context is having lived with a risk with little going wrong. Because new GE products are only now becoming available to a limited public, the perception of risk is presently overestimated. As the flow of products increases, this component should diminish, but the efforts of opponents may intensify.

b. Knowable vs. Unknowable

Knowability is based on the level of understanding of the science involved, the extent to which experts agree about the risk, and the degree of trust by the public in the sources of the information.

At present, GE is relatively unknowable for the following reasons:

1. As is the case with science generally, the level of understanding by the public is low (due to the complexity of the subject matter and its specialist position in education); the public obtain their knowledge about biotechnology mainly from the media (C.E.C., 1993), which in turn obtains it from the scientific community and from special interest groups (e.g., customer and environmental groups). Each of these sources has its own agenda and constituency.
2. There is considerable disagreement among experts as to the extent of the risks involved in specific aspects of GE. For example, experts disagree about the

risks involved in releasing engineered plants into the environment. This disagreement arises because not enough useful data exist and experts therefore project from their differing experience and concept bases

3. The public have a relatively high degree of trust in the information supplied by customer and environmental groups and by scientists in universities and institutes, but less in that supplied by commercial companies or the media (Hoban and Kendall, 1992; C.E.C., 1993; Zechendorf, 1994).

c. Nonmemorable vs. Memorable

Memorability in this context means how easy is it to envisage something going wrong (i.e., of losing control of the application of the results of scientific endeavors). Memorability is a key outrage factor for GE, whether plant, animal, or human, with images strongly rooted in our culture, for example, in literature, the Frankenstein myth (the contemporary parable of perverted science [Warner, 1994; Butler, 1993]) and in films such as *Jurassic Park.* Popular belief is often influenced more by images than facts so that writers' opinions of the scientists and science of their day can contribute strongly to the public's evaluation of these scientists and their work. This is borne out by surveys conducted with various groups representative of the general public about their image of scientists, which show that few actual scientists have contributed to this image, whereas fictional creations (e.g., Dr. Frankenstein and Dr. Strangelove) have exerted a strong influence (Haynes, 1994). It is important, therefore, for scientists to be aware of how they have been represented over time in Western literature, and more recently films, because these fictional scientists are the authors concerned, response to the role of science and technology in the society of their time. A majority of fictional scientists have been portrayed with overall negative qualities so a study of this literature enables us to identify what the concerns are and to enquire into the basis of these often unacknowledged concerns by society about science. If science is to be socially relevant, scientists must be able to relate their field of science to the wider general culture, which requires some knowledge of relevant history and of literature. The importance of this can be neatly seen in a comparison of Bateson's and Bauer's attitude to eugenics. From the same scientific standpoint, Bateson came to a different conclusion from Bauer due to the cultural influences on him (see Harvey, 1995).

In the following survey of the scientist in Western literature and films, this author has drawn heavily on the wealth of information in Haynes (1994).

Fictional scientists are portrayed as one or another of six recurring stereotypes (Haynes, 1994):

1. The alchemist who was, in medieval times, at the cutting edge of experimental research into the secrets of nature. He is portrayed in the literature as secretive, mysterious, obsessed or mad, and pursuing, ineffectively, esoteric intellectual goals, that are inherently evil in that they attempt to overreach God's limitation of man's knowledge and the divine prerogative, in trying to create life from matter (Marlowe, 1604; Jonson, 1610). Although in the seventeenth century alchemism decayed and Baconian scientific endeavor superseded it, literature has, to this day, repeatedly called up this stereotype of the scientist (e.g., the secretive, evil biologist attempting to create new species by near-magical genetic engineering techniques even to the extent of the similarity between the alchemist trying to turn base metal into gold and mod-

ern genetic engineers who have not delivered so far the great benefits that they predicted).

2. The foolish specialist who is out of touch with reality and, therefore, ineffectual and a social and moral failure. He is out of touch with reality either in choosing to reject the messy real world for a mechanical, logical scientific one or because science necessitates it (i.e., to be successful in science you have to isolate yourself from the world of human relationships). Dr. Frankenstein (Shelley, 1818) had to isolate himself at first by frequenting lonely graveyards for organs and then in his laboratory to create the monster. Also, Frankenstein rejects Elizabeth because he can substitute marriage and potential procreation with Elizabeth by creating the monster instead (Butler, 1993).
3. The unfeeling impersonal scientist of Romantic literature. Romantics believed that understanding the world was a heroic game and that creative poetic experience transcended the limitations of the material world. This was achieved by insights gained from subjective feelings and emotions induced by personal relationships and by communicating with nature. They believed experimental, analytic, objective science that was materialistic and rational interfered and eroded this true, subjective approach (Blake, 1966). They represented the scientist accordingly as someone who suppresses all feeling and eschews personal relationships as a necessary price for engaging in scientific activities (Shelley, 1818). This view of science and scientists includes the amoral scientist as a subspecies. Whereas the Romantic's representation of the scientist referred to isolation from the scientist's family and friends, the amoral scientist affects the whole of society and humanity by his disregard of ethics (see Schirmbeck, 1960). This stereotype is often identified as a respected, powerful, and very successful scientist, advising industry and governments; there are many fictional works on nuclear weapons, robots, and environmental pollution based on this stereotype (see Haynes [1994] for examples).
4. The helpless scientist who loses control of his discoveries, which become unexpectedly monstrous themselves or spawn applications whose implementation he cannot control (e.g., Frankenstein [Shelley, 1818]).
5. The scientist as a hero who uses his towering intellectual prowess to open up new possibilities or to save mankind from disasters (see Daumann, 1940).
6. The scientist as idealist, creator of a scientific utopia of plenty and human fulfillment (e.g., Wells, 1926).

A majority of these stereotypes and fictional works portray scientists (and science) in negative terms, yet these representations provide an understandable means for the public to view science and scientists when compared with the facts, which are difficult to understand. These stereotypes are extreme representations of public concerns and the main point is not that the average man in the street necessarily thinks all scientists are evil, but that scientists, when presenting their field of science to the public, must be aware of these often buried concerns (see later discussion).

d. Nondreaded vs. Dreaded

Some hazards have associated with them a dread component so that perceived risk is high (e.g., fear of snakes or mice). Perception of dread may arise from personal or

nonpersonal experiences. An example of where a personal experience is responsible is when environmental polluting waste has an association with childhood experiences. An example of a nonpersonal source is recall of images of media coverage of past events; the recalled image causing dread has often no direct connection with the present hazard. As pointed out by Gell-Mann (1994), the impact of scientific discovery on the literary world and popular culture often gives rise to words or phrases so that connotations other than those of the original technical meaning are attributed to the scientific discovery (e.g., popular use of the scientific meaning of the term "chaos"). Dread can become attached to some hazards by dread-associated words: the term biotechnology, for example, is less dreaded than genetic engineering (C.E.C., 1993). GE has a large dread component, with images of the possible pollution of air, water, food by uncontrolled, "escaped" invisible genes that cannot be recalled.

e. Voluntary vs. Nonvoluntary

In an advanced technology, the opportunity for the choice exists and voluntary risks are much more acceptable than nonvoluntary risks. Hang-gliding as a pastime carries a risk of fatal accident, but those participating in hang-gliding accept this risk voluntarily (i.e., set a higher-than-average risk tolerance) because of the benefits it affords them. Voluntariness is usually associated with a perceived benefit, which can vary from individual to individual (i.e., a perceived benefit increases the acceptable risk tolerance level as seen in the above example). By a similar process, the existence of possible alternatives reduces the acceptable risk-tolerance level. This can be seen in the next example where some environmental groups have argued that PGE, even if of low risk, is unacceptable because satisfactory alternatives (e.g., new cropping methods) exist (sometimes stated as the fourth hurdle, i.e., demonstration of need).

How voluntary then is GE? In the 1970s, once scientists were convinced of its safety (at least under containment conditions), GE as part of science progressed at a rapid pace. The aim of science is to understand the world and its prime motivation is curiosity and serendipity. Plant biotechnology has come into being by industry applying the findings from science. Technology sets out to make "things," and uses not only the finding of science, but also other practical experiences. Technology's development is mainly driven then by the ideas of scientists, initially funded mainly by the public purse, but less so recently, by the chosen objectives of commercial companies, and by government regulations, which are influenced by the viewpoints of scientists, industry, concerned special interest groups, and the general public. Thus, the introduction of GE has an element of perceived nonvoluntary risk because, although regulated (controlled), the public were not consulted over its introduction, although a moratorium on some GE activities is in place (e.g., germline gene therapy on humans).

Whether or not to label food produced by GE is now a major debate (Cooperative Wholesale Society, 1994). Labeling would make any risk involved voluntary and many believe that this would boost public confidence, and in the case of those sections of the community (e.g., Muslims) who should not eat food containing genes from some sources, ethically correct to do so (HMSO, 1996). However, not every component of a food can be labeled, and methods of food production ranked biotechnology fifth out of seven factors when compared with other types of information; more interest was shown in fat content and levels of pesticides and food additives (Hoban and Kendall,

1992). The approach of the U.S. Food and Drug Administration (FDA, Rockville, MD) is that labeling must be accurate and that the material be relevant and not mislead, (see Miller, 1995). In the end, whether or not to label is a decision of government after receiving advice from appropriate committees (see Ward, 1995), but it is important that these committees include a wide representation of the public. Politicians often rely only on scientists to advise them about risk, but there will always be disagreement among scientists, especially when there are few firm data, so outrage factors (lack of trust, unknowability, involuntariness) may predominate so that the scientific assessment of risk then becomes irrelevant. A good example of that happening was the public perception of risks due to BSE in the U.K. in 1996.

f. Diffuse vs. Concentrated

A distinction is made between chronic (less severe but recurring) and catastrophic risks (severe but infrequent). For example, the car is perceived partly as less risky than the airplane, because airplane disasters, although they occur less often, are highly concentrated in the sense that many may die at once. Concentration also refers to whether future generations will be affected. For example, a chemical spill is a single event whose consequences can be minimized, whereas mistakes with genetically engineered changes would be inherited and perpetuated. Scientists disagree as to whether biotechnology is a concentrated risk. Some ecologists claim that although the chance of GE causing serious ecological damage is small, if it occurred, the consequences could be disastrous. This aspect of an outrage reaction can be reduced by having publicly acceptable regulations in place and contingency plans ready if anything should go wrong.

g. Controlled vs. Uncontrolled

Whereas voluntariness refers to who decides if something will happen, control refers to who regulates its implementation after it has. Most biotechnological projects have been initiated without prior discussion with the public (i.e., involuntarily), but if the range of interests of the general public are represented on the membership of statutory national and local regulatory bodies, then implementation of biotechnological projects will be, in this sense, controlled.

h. Natural vs. Technological Risk

Relative to natural risks, technological risks are exaggerated by the public. Some "genetic engineering" processes do take place in nature, mainly among microorganisms, but it is misleading to claim that GE is natural as many of the operations of GE are laboratory based. However, most scientists view GE as conceptually no different from conventional breeding. Furthermore, GE is subject to regulations that some say err on the side of caution (see House of Lords Select Committee Report, 1993).

B. GE is Morally Wrong

GE is regarded by some members of the public as being blasphemous or unnatural or disrespectful, or unfair, on ethical grounds (Straughan, 1992); they therefore oppose the development or use of GE products. Unlike in the outrage reaction described above, the extent of the risk is not at issue, just that GE is unacceptable. There is no doubt that ethical considerations can exercise a powerful influence over the acceptance of GE (Hoban and Kendall, 1992; Straughan, 1992).

1. PGE Is Blasphemous

Based on the religious belief that God created a perfect, natural order, some mem-

bers of the public consider it wrong to manipulate DNA and cross species boundaries (i.e., play God). This moral concern also applies equally forcibly to traditional animal and plant breeding and other human activities that interfere with "created order." However, not all religious believers have this view of creation; many accept that species have changed during evolution and recognize continued interference as reasonable. Still other religious objectors who are not creationalists and accept evolution nevertheless believe that man should carry out God's purpose by acting as a steward of nature and not interfere with nature by using GE. They may even argue that although in nature the question of individual responsibility does not apply (i.e., it is amoral), man is considered to be outside nature in this regard and has moral duties of care. In contrast, others think that in applying GE we are using gifts given to man by God in order to adapt to the environment.

Recently, the discussion in this area has been complicated by assertions that science is antireligious and undermines man as a transcendental being, thereby destroying life's meaning. If we follow Durkheim's secular view that religion has its origins in the demands of life in a society and fulfills the practical need of sociability and the intellectual need to understand the world (Giddens, 1978), we can see that religion and science (understanding of the world), although using different means, have some aims in common and have sometimes indeed come up with different answers to the same question. However, the role of religion in relation to moral conduct, stressing something more important than self and giving coherence of social outlook, cannot be delivered by science in its stead.

Belief in God, a matter of faith, gives meaning and solace to many peoples' lives, whereas science is needed to help to provide the wealth (breathing space) to allow free exercise of the intellect. In this way, religion and science are complementary and many scientists are practicing Christians, etc.

2. GE Is Unnatural

An historical perspective of man's attitude to nature gives a sense of balance to present day reflections on nature. Mankind has seen nature at different times very differently (e.g., as the abode of gods or of devils). He has feared nature and tried to placate it, or seen it as the source of all true feelings (a paradise he longs for), has hated it as dirty and germladen, or as a purifying influence; he has seen it as a resource to be exploited and recently as in need of our help to preserve it (Thomas, 1983). Man's view of nature is deeply rooted, but ambivalent.

Some people today view GE as wrong because it is artificial; logically, this implies that all that is natural is good and all that is unnatural is bad (Straughan, 1992). Not everyone would agree; "nature red in tooth and claw" is an alternative view. In any case, there is no clear-cut distinction between natural and unnatural activities (Straughan, 1992; Boulter 1995). Criticisms of artificiality are closely related to the idea of an imagined Arcadian wilderness where without human tampering, balance and harmony reign and all will be well (i.e., nature knows best). This myth of a golden age where man lived in harmony with nature is deeply rooted in many cultures (e.g., see Romantic literature), but it is doubtful whether it ever existed; nostalgia is confused with reality. Modern ecological studies show that ecosystems, which from this standpoint must be conserved at all costs, are interwoven throughout with previous human intervention, thus weakening the force of the argument. From this viewpoint, nature is seen as a delicate web of interconnected species and any possible loss of biodiversity is met with alarm since nature, on her own, will evolve harmony and balance. Again, modern ecologi-

cal and geological studies do not confirm this interpretation of the past. It is by no means certain that nature, if left alone, will arrive at a perfect, balanced, harmonious climax ecosystem. Ecosystems change their compositions and the history of living organisms (plants, animals, and microbes) on this planet is full of examples of extinctions, several of which have been catastrophic for the existing flora and fauna but have led to the evolution of more complex organisms (Budianski, 1995).

Apart from genetic pollution, GE is also said by some to aggravate the depletion of biodiversity. Undoubtedly, high-tech agriculture has led to loss of land races (no longer planted) and wild and weedy relatives (removed so as not to decrease yields of highly bred crops) and GE, it is claimed, will accelerate this trend. Others counterclaim that GE is a powerful new tool that is essential if we are to understand complex ecosystems sufficiently to devise sustainable agricultural practices (the only renewable energy source available) and conserve the environment.

Thus, there is a conflict of concepts between some ecologists and molecular biologists as to how best to understand material processes, either holistically and subjectively (ecologically) or by what is perceived as reductive scientific analysis, (molecular biologically, but see later discussion) and there is a conflict of projected outcomes — GE leading to genetic pollution and genetic erosion on the one hand, or to sophisticated biological management and sustainability on the other. Ecologists have been instrumental in drawing attention to the importance of biodiversity and the undervaluation of biological wealth (Wilson, 1992); of course, conservation, preservation of biodiversity, and protection of the environment are essential if life-support systems are to be safeguarded. However, if the world is to be fed and life-support systems sustained, active management of nature will be required (Budianski, 1995). As science becomes more holistic in approach and as the conceptual conflict lessens, and because everyone accepts the desirability of sustaining resources, it should now be possible to have constructive multidisciplinary debate.

There are many for whom an isolated contemplation of nature brings joy, dignity, and spiritual peace and such persons have been forthcoming worldwide since time immemorial (e.g., see Tobias, 1995). Similarly, the artist, in his intense study of nature, is enriched by feelings of aesthetic delight. No one advocates that these nonscientific approaches to a study of nature are invalid or that science can substitute for them. The problem arises when such arguments are used as a basis to propose that no interference with nature should be allowed and that GE should be stopped forthwith as it "tampers" with natural processes causing changes that would never, or are extremely unlikely, to happen without human intervention, thus interfering with a better way of interpreting the natural world. This attitude is essentially the same as that of the Romantics for whom subjective gut feelings and poetic sensibilities are a better way of understanding the world than scientific analysis.

3. GE Is Disrespectful

The view that GE invokes the charge of disrespect is based on the suggestion that GE is reductionalist (i.e., sees life as merely a collection of chemicals [genes] available for manipulation without regard for the ends of others [see Straughan, 1992]). The vision is of a machine-like, efficiently engineered world lacking free will and compassion and where the material and social environments are seen as being of no importance.

This view misunderstands the role of genes, which only predispose organisms to

have character traits and whose expression is strongly dependent on the environmental and genetic background. Modern biology is reductionalist in that it does not invoke "vital forces" over and above the physicochemical laws, but nearly all biology is concerned with finding regularities and laws at its own level (organisms and populations) with little attempt to relate to the fundamental level. The view also ignores the fact that modern science teaches us that the laws of nature are probabilistic, governed by quantum uncertainty and "chaotic," a far cry from a Newtonian mechanistic world view.

4. GE Is Unfair

It is a common perception that the distribution of the risks from GE will not correlate with who will benefit. Industry is perceived by some to benefit most, perhaps unfairly (who gets most benefit from genetically engineered tomatoes?). In the worst scenario, industry is accused of cornering the market by upping the entry price with high technology and by patenting living organisms (see Busch, 1995; Ward, 1996b). Another concern is the possible adverse effects of GE on Third World economics and environments, by product substitution, by owning IPR to germplasm, by involuntary adverse effects on their environments due to global distribution of GE crops, or by technological innovation disturbing rural demographic and social change (Meister and Mayer, 1994; Gates, 1995). The alternative view (Gressal, 1992) is that the Third World will be likely to benefit most from GE as it will provide a means to solve the problems of plant production under the stress conditions of Third World agriculture.

Unfairness to organisms, another area of concern, is not involved in GE as the important distinction in this connection is between sentient vs. nonsentient organisms (plants). Even this is contentious as the cultivation of arable crops has led in the past to much loss of bird and small mammal life (Girling, 1996). Although these important issues are not addressed in detail here, wide-ranging open debate of all of them is essential.

VII. WHAT CAN SCIENTISTS DO?

In the past, scientists, apart from a few notable exceptions, have not seen it as their responsibility to inform the public of the results of their work; popularization of science was usually defined as amateur science due to the entertainment context. Now it is becoming apparent that if science is to be fully accepted as part of our culture, scientists have a responsibility to ensure communication of their work to the public in a balanced, understandable, interesting, and relevant way (Boulter, 1995), and clauses to cover this responsibility are becoming more common in Contracts of Service. Because of the complexity of the wider value judgment issues, communication cannot be left to the scientists alone, and their employers need to provide specialist communicators as well (USDA, 1992; BBSRC, 1994).

Most scientists have learned their professional skills by training and experience and get on with their daily scientific endeavors without thinking very much about what it is scientists actually do. Before attempting to communicate their work to the public, it is necessary for scientists to take stock, first of the way science is done and second of the way to avoid/minimize outrage reactions.

Scientists undergo long training and science is a highly specialized activity. Therefore, the public often find science difficult to understand, even though polls show a great desire on their part to know more about science (Mervis, 1966). The results of consensus conferences, moreover, show that the nonscientist, man in the street, can under-

stand scientific issues when time and effort are expended (U.K., 1994).

In science, theories (generalization) ideally arise from a large number of observations (i.e., they are a condensed description of many publicly tested experimental observations [Gell-Mann, 1994]). Theories are tested and accepted or falsified (Popper, 1976) according to their generality, coherence, and ability to explain and predict new observations. In practice, scientific activity does not fit any model exactly of how science works (Medawar, 1969; Fayerabend, 1975), whether this model is positivist (Cole, 1996), Popperian (Popper, 1976), or Kuhnian (Kuhn, 1970), insofar as scientists allow selfishness, dishonesty, and prejudice to frustrate the scientific criteria of strict reproducibility and falsifiability. Even so, the requirement that data become part of science only if in the public domain means science tends to be self-correcting (Gell-Mann, 1994).

Scientific endeavor is then an extension and a limitation of the way the ordinary citizen goes about problem solving in practice. Dunbar (1995) has suggested that evolution has selected a more superficial, faster approach for everyday use that has served man well, bombarded as he is by scientific data and because of his need to operate as a social animal. Thus, the public weigh evidence much in the way a jury does when assessing expert testimony and use a much wider basis for consideration than the strictly scientific ones of empirical experimental data evaluation within an existing paradigm (Kuhn, 1970). The public's view of science as value-free depends on science placing all the emphasis on the data (not by whom or why it was obtained) and on science limiting itself to areas that are amenable to experimentation. This does not imply that scientists do not have values or that value judgments outside science are invalid. Nevertheless, the necessity to demonstrate the power of "reductive," "value-free" scientific analysis as a means of solving problems has led to a too-narrow focus on mechanisms to be very useful with complex adaptive systems; this, however, is an argument for broadening the scientific approach rather than disparaging scientific endeavor. Scientists also need to recognize the reasons why magic and mystification rather than scientific explanations are sometimes preferred by the public (Sagan, 1996).

In attempting to understand science, the man in the street has to accommodate to this change in pace and intensity and has often to grasp counter-intuitive ideas. It is important for scientists to realize the extent of this conceptual and knowledge gap and adjust presentations accordingly (see later discussion).

It has often been advocated that the main reason for a lack of public acceptance of science is due to factual ignorance on the public's part. Although polls show that increased factual knowledge reflects pollsters attitudes in general, more knowledge can lessen support for morally contentious research (Evans and Durant, 1995). However, a further complication is that, apart from what was learned at school, the public's perception of science comes from the media. Special interest groups are adept at getting across their viewpoints to the media. Although not disputing their right to do so, scientists need to collaborate with writers and journalists to counter unbalanced reporting when it occurs and to increase the amount of high quality science published in the media by overcoming the present conflicts between the aims of the media to entertain and the presentation of "hard science" (Sherwood, 1993).

Scientists should also be prepared to sit on national and local committees concerned with scientific issues and should argue for wide membership (nonscientific representation). When few data are available, as with a new technology, they should be particularly

careful to give a balanced view and separate opinion from scientific fact. Finally, scientists have a special duty to become involved, where appropriate, in helping to inform school children, and there are now many joint ventures between universities, research institutes, and industry with schools in many countries (Solomon, 1996).

VIII. HOW SHOULD SCIENTISTS PRESENT THEIR WORK TO THE PUBLIC?

When talking to the public, it is necessary to realize what is to be achieved. The first objective is not to convey the technical content of his/her work but to convince the public that he/she is a well-balanced person: "Can scientists shake off their mad media image?", reads the headline (Brown, 1966) and explain their aims and disappointments in carrying out their work. The next objective is to establish trust and to convince the public that he/she realizes scientific activities must be socially acceptable by showing a concern for possible risks and their support for regulatory controls and a consideration of contingency plans if anything were to go wrong. Past mistakes by scientists should be acknowledged and while being enthusiastic about scientific achievements, he/she must not appear arrogant and overplay the benefits of science; the public realize that many benefits can accrue without science (Irwin, 1995) (e.g., conventional breeding without the influence of genetics). Scientific articles and TV programs that fulfill these criteria are popular and successful. The scientific content of the talk should be concise, jargon-free, and be written bearing in mind that the choice of words is very important. On the one hand, the choice should not be seen as attempting to camouflage risk, but on the other, should not evoke "outrage" reactions. The overall objective of communication is to (1) emphasize the beneficial role of GE and its similarity with conventional agriculture so that, although new, it is in many ways familiar; (2) provide an opportunity for public debate to influence research objectives; and (3) consider the risk/benefit trade offs, including ethical and social issues.

IX. CONCLUSIONS

Whereas science and technology were formerly seen as separate (science dispassionately understanding nature, and technology applying science, crafts, and empirical experiences), the former only for good, the latter both for good or evil depending on choice, now science and technology are becoming increasingly merged (see recent funding trends worldwide). Whereas scientists were seen as being driven by curiosity, serendipity, and open publication of their findings, now sociologists have shown that scientists are rarely motivated purely by scientific criteria and science like other activities is strongly influenced by social, economic, and political considerations.

Rightly, therefore, the public are becoming more involved in setting the scientific agenda, but there is also an urgent need for increased public understanding of both science itself and the issues, including the cultural bases of so-called "subjective irrationality," which generally in the past has been in dynamic equilibrium with rationality. Scientists, sociologists, professional bodies, industry, and governments all have important roles to play in this education process.

Scientists would gain by talking to the public, since by doing so they gain their trust and also their understanding of the importance of GE science in wealth creation and for the public good. Failing to do so will mean the continued demonization of scientists (yuk factor), which results from the

public's lack of understanding and its sense of powerlessness (Duffy, 1995).

ACKNOWLEDGMENT

I would like to thank A. Richardson for help in preparing the manuscript.

REFERENCES

Ahl Goy, P., Chasseray, E., and Duesing, J. H. 1994. Field trials of transgenic plants: an overview. *Agro Food Ind. Hi-Tech*, **2**: 10–15.

Ahl Goy, P. and Duesing, J. H. 1996. Assessing the environmental impact of gene transfer to wild relations. *Biotechnology*, **14**: 39–40.

Baker, H. G. 1965. Characterisation and modes of origin of weeds. **In:** *The Genetics of Colonising Species*. Baker, H. G., Stebbings, G. L., eds., pp. 142–172. Academic Press, New York.

Blake, W. 1966. **In:** *The Complete Writings of William Blake*. Keynes, G., ed., Oxford University Press, London.

BBSRC. 1994. *Agricultural and Food Research Council Annual Report*, p. 13. Polaris House, Swindon, U.K.

Best, S. G. 1995. Developing sustainable agricultural production systems for the 21st century. *Gen. Eng. News*, p. 22. AGBIO, July 1995.

Boulter, D. 1995. Plant biotechnology. Facts and public perception. *Phytochemistry*, **40**: 1–9.

Bright, S. W. J. et al. 1996. Environmental impact from plant biotechnology. **In:** *Engineering Plants for Commercial Products and Applications*. Collins, G. B., ed., *Ann. NY Acad. Sci.*, in press.

Brown, S. L. and Crawley, M. J. 1991. *Environmental Risk Assessment Consideration for Certain Gene Modified Plants*. EC Contract Ref. B6614/90/011521.

Brown, G. 1996. Can scientists shake off their mad media image? *The Times*, January 1st, p.12. Times, London.

Budianski, S. 1995. *Nature's Keepers*. Butler and Tanner, London.

Busch, L. 1995. Eight reasons why patents should not be extended to plants and animals, p. 24, *Biotechnology and Development Monitor*, 24. University of Amsterdam, The Netherlands.

Butler, M. 1993. The first Frankenstein and radical science. *Times Literary Supplement*, April 9, pp. 12–14. Times, London.

C.E.C. 1993. Eurobarometer 39.1, INRA, *The European Commission*.

Cole, S. 1995. *Making Science*. Harvard University Press, Cambridge, MA.

Cooperative Wholesale Society Ltd. 1994. Leaflet. *The Right to Know: Your Guide to Genetic Engineering*. CWS, P.O. Box 53, New Century House, Manchester, U.K.

Darmency, H. 1994. The impact of hybrids between genetically modified crop plants and their related species: introgression and weediness. *Mol. Ecol.*, **3**: 27–40.

Daumann, R. H. 1940. *Protuberanzen*, Schützen-Verlag, Berlin.

de Vries, F. T., van der Meijden, R., and Brandenburg, W. A. 1992. Botanical files – a study of the real chances of spontaneous gene flow from cultivated plants to wild flora of the Netherlands. *Gorteria* (**Suppl.**)**1**: 1–100.

D'Mello, J. P. F., Duffus, C. M., and Duffus, J. H. 1991. *Toxic Substances in Crop Plants*. The Royal Society of Chemistry, Cambridge, U.K.

D. N. 1991. Report 7b. *Ecological risks of releasing genetically modified organisms into the natural environment*. Directorate for Nature Management, Trondheim.

Duffy, M. 1995. When the scientist plays the Devil. *The Independent Newspaper*, 28 February 1995, p. 23.

Dunbar, R. 1995. *The Trouble with Science*. Faber and Faber, London.

Economidis, I. 1992. Biosafety research in the European Community: results and perspectives. **In:** *Proc. 2nd Int. Symp. on the Biosafety Results of Field Tests of Genetically Modified Plants and Microorganisms*. pp. 135–140. Casper, R. and Landsmann, eds., Goslar, Braunscheig, Germany.

Evans, G. and Durant, J. 1995. The relationship between knowledge and attitude in the public

understanding of science in Britain. *Public Understanding Sci.*, **4:** 57–74.

Fraley, R. 1992. Sustaining the food supply. *Bio/technology*, **10:** 40–43.

Feyerabend, P. 1975. *Against Reason*. NLB, London.

Gates, P. 1995. The environmental impact of genetically engineered crops. *Biotechnol. Eng. Rev.*, **13:** 181–195.

Gell-Mann, M. 1994. *The Quark and the Jaguar*. Little, Brown and Co., London.

Giddens, A. 1978. *Durkheim*. Fontana Modern Masters, pp. 125. Fontana/Collins.

Girling, R. 1996. Cheese is Murder. *The Times Magazine*. January 13th 1996, pp. 9–15, *The Times* , London.

Gressel, J. 1992. Genetically engineered herbicide-resistant crops —a moral imperative for world food production. *Agro-Food Ind. Hi-Tech*, **November/December:** 3–7.

Hahn, D. H., Rooney , L. W., and Earp, C. F. 1984. Tannin and phenols of sorghum cereal foods. *World*, **29:** 776–779.

Hallman, W. K. 1996. Public perceptions of biotechnology: another look. *Bio/technology*, **14:** 35–38.

Harvey, R. D. 1995. Pioneers of genetics: a comparison of the attitudes of William Bateson and Erwin Bauer to Eugenics. *Notes Rec. R. Soc.* London, **49:** 105–117.

Haynes, R. D. 1994. *From Faust to Strangelove*. The John Hopkins University Press, Baltimore.

HMSO. 1996. Report of the Committee on the Ethics of Genetic Modification and Food Use. HMSO.

Hoban, T. J. and Kendall, P. A. 1992. Consumer attitudes about the use of biotechnology in agriculture and food production. Report to the USDA Extension Service, North Carolina State University, Raleigh, N.C.

House of Lords. 1993. Select Committee Report: Regulation of the United Kingdom Biotechnology Industry and Global Competitiveness.

Irwin, A. 1995. *Citizen Science: A Study of People, Experience and Sustainable Development*. Routledge, London.

Jones, D. D. 1994. Proc. 3rd Int. Symp. on biosafety results of field trials of generically modified plants and microorganisms. University of California, Division of Agricuture and Natural Resources, Oakland, CA.

Jonson, B. 1610. The alchemist. **In:** *The Complete Plays of Ben Jonson*. Wilkes, G. A., ed., Clarendon Press, Oxford, 1982.

Karieva, P. and Parker, I. 1994. Environmental risks of genetically engineered organisms and key regulatory issues. Independent report prepared for Greenpeace International, Canonbury Villas, London N1 2PN, U.K.

Kjellson, G. and Simonsen V. 1996. *Methods for Risk Assessment of Transgenic Plants. Competition, Establishment and Ecosystems Effects*. Birkhäuser Verlag AG, Basel, Switzerland.

Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*. 2nd ed. University of Chicago Press, Chicago.

Kung, S.-D. 1993. Introduction: from hybrid plants to transgenic plants. **In:** *Transgenic Plants*, Vol 1. Kung, S.-D. and Wu, R., eds. pp. 1–12. Academic Press, London.

Lex, M. 1995. *Eur. Biotechnol. Inform. Serv. Newsl.*, **5**. European Commission, Brussels.

Macdonald, J. F. (Ed) 1991. *Agricultural Biotechnology at the Crossroads*. N.A.B.C. Report 3, University of California.

Marlowe, C. 1604. The Tragical History of Dr. Fautus. **In:** *Marlowe's Plays and Poems*. Ridley, M. R., ed., Dent, London, 1963.

McGarvey, P. B. and Kaper, J. M. 1993. Transgenic plants for conferring virus tolerance: satellite approach. **In:** *Transgenic Plants*, Vol. 1. Kung, S.-D. and Wu, R., eds., pp. 277–296. San Diego, Academic Press.

McPartlan, H. C. and Dale, P. J. 1994. An assessment of gene transfer by pollen from field grown transgenic potatoes and related species. *Transgenic Res.*, **3:** 216–225.

Medawar, P. B. 1969. *Induction and Intention in Scientific Thought. Deficiency of Observation, Experimentation and Induction on Both Theory and Practice*. Methuen, London.

Meister, I. and Mayer, S. 1994. Genetically engineered plants: releases and impacts on less

developed countries. Submitted to First Conference of Parties of the Convention on Biological Diversity, Nassau, by Greenpeace International.

Mervis, J. 1996. Wanted citizen scientists. *Science*, **271:** 904.

Miller, H. I. 1993. Perceptions of biotechnology risks: the emotional dimension. *Bio/technology*, **11:** 1075–1076.

Miller, H. 1995 Labelling of "Biotech Foods" is no free lunch. *Bio/technology*, **13:** 1510.

National Farmers Union. (NFU) 1995. Report of the Biotechnology Working Report, N.F.U., London.

O.E.C.D. 1986. Recombinant DNA Safety Considerations for Industrial, Agricultural and Environmental Applications of Organisms Derived by Recombinant DNA Techniques. Paris, France.

O.E.C.D. 1992a. Biotechnology and Food. Paris, France.

O.E.C.D. 1992b. Good Development Principles (GDP) for Small-Scale Field Research. Paris, France.

O.E.C.D. 1993. Safety Considerations for Biochemistry: Scale-Up of Crop Plants. Paris, France.

Perrins, J., Williamson, M., and Fitter, A. 1992a. Do annual weeds have predictable characters? *Acta Oecol.*, **13:** 517.

Perrins, J., Williamson, M., and Fitter, A. 1992b. A survey of differing views of weed classification: implications for regulation of introductions. *Biol. Conservation*, **60:** 47–56.

Popper, K. 1976. *Unending Quest*. Fontana, London.

Raybould, A. F. and Gray, A. J. 1993. Genetically modified crops and hybridisation with wild relatives: a UK perspective. *J. Appl. Ecol.*, **30:** 199–219.

Regal, P. J. 1994. Scientific principles for ecologically based risk assessment. *Mol. Ecol.*, **3:** 5–13.

Rogers, H. J. and Parkes H. C. 1995. Transgenic plants and the environment. *J. Exp. Bot.*, **46:** 467–488.

Sagan, C. 1996. *The Demon-Haunted World: Science as a Candle in the Dark*. Headline, New York.

Sandman, P. M. 1992. Hazard versus Outrage. *I.O.M.A. Broadcaster*; 6–17.

Schürmbeck, H. 1960. *The Blinding Light*. (Transl. Denny, N.) Collins, London.

Shelley, M. W. 1818. *Frankenstein, or the Modern Prometheus*. Dowse, R. E. and Palmer, D. J., eds. Dent, London, 1963.

Sherwood, R. F. 1993. President's lecture: the need for scientific communication with the public. *Science*, **260.** 1571-1576.

Sobral, B. 1996. *The Impact of Plant Molecular Genetics*. Birkhäuser Verlag AG, Basel, Switzerland.

Solomon, J. 1996. School science and the future of scientific culture. *Public Understanding Sci.*, **5:** 157–165.

Straughan, R. 1992. *Ethics, Morality and Crop Biotechnology*. Occasional Publ. ICI Seeds, Bracknell, U.K.

Tepfer, M. 1993. Viral genes and transgenic plants. *Biotechnology*, **11:** 1125–1132.

Tiedje, J. M. et al. 1989. The release of genetically engineered organisms: a perspective from the Ecological Society of America. *Ecology*, **70:** 297–315.

Tobias, M. 1995. *A Vision of Nature*. The Kent State University Press, Kent, OH.

Thomas, K. 1983. *Man and the Natural World*. Penguin Books, London.

U.K. 1994. *Final Report National Consensus Conference on Plant Biotechnology*. Science Museum, London.

United States Department of Agriculture (USDA). 1992. Methods of Communicating Biotechnology with the Public. Final Report. United States Commission of the European Communities Workshop, Dublin, Ireland.

Ward, M. 1995. European food rules creep forward. *Bio/technology*, **14:** 1428–1429.

Ward, M. 1996a. Another push to revise Eurobiotech directives. *Bio/technology*, **14:** 133–134.

Ward, M. 1996b. Biopatents in Europe re-opened take account of EP's ethical concerns. *Bio/technology*, **14**:139.

Warner, H. 1994. *Managing Monsters*, Vintage, London.

Wells, H. G. 1926. *A Modern Utopia*. Unwin, London.

Weverling, J. and Schenkelaars, P. (Eds.). 1992. *Ecological Effects of Genetically Modified Organisms*. Netherlands Ecological Society. Drenthesingel 11, Arnhem, The Netherlands.

Williamson, M. 1993 Invaders, weeds and the risk from genetically manipulated organisms. *Experentia*, **49**: 219–224.

Wilson, E. O. 1992. *The Diversity of Life*. Harvard University Press, Cambridge, MA.

Zechendorf, B. 1994. What the public thinks about biotechnology. *Bio/technology*, **12**: 870–875.

Transgenic plants: An emerging approach to pest control

Juan J. Estruch*, Nadine B. Carozzi, Nalini Desai, Nicholas B. Duck, Gregory W. Warren, and Michael G. Koziel

Insect Control-Seeds, CIBA Agricultural Biotechnology, 3054 Cornwallis Rd., Research Triangle Park, Durham, NC 27709.
**Corresponding author (e-mail: estruchj@am.abru.cg.com).*

Received 4 September 1996; accepted 20 December 1996

Insect pests are a major cause of damage to the world's commercially important agricultural crops. Current strategies aimed at reducing crop losses rely primarily on chemical pesticides. Alternatively transgenic crops with intrinsic pest resistance offer a promising alternative and continue to be developed. The first generation of insect-resistant transgenic plants are based on insecticidal proteins from *Bacillus thuringiensis* (*Bt*). A second generation of insect-resistant plants under development include both *Bt* and non-*Bt* proteins with novel modes of action and different spectra of activity against insect pests.

Keywords: *Bacillus thuringiensis*, δ-endotoxins

Plant biotechnology has become a source of agricultural innovation, providing new solutions to age old problems. The insect control arena is one of its most active theaters of operation. The simple underlying reason for this interest is that worldwide crop damage inflicted by phytophagous insects is staggering despite the use of sophisticated crop protection measures, chiefly chemical pesticides. The cost associated with management practices and chemical control of insects approaches $10 billion[1] annually, yet global losses due to insects still account for 20% to 30% of the total production. Every year, individual farmers face the possible devastation of their crops by insect infestations. In view of this, genetic engineering of insect resistance into crops represents an attractive opportunity to reduce insect damage with several clear advantages over traditional chemical pesticides. The potential exists to transfer some of the most active insecticidal principles found in diverse organisms into crop plants. Over the past few years, the success in producing insect-resistant crops through gene transfer has been impressive, and the culmination of this process occurred in 1996 when the first generation of insecticidal plants, generically known as Bt plants, were introduced into the marketplace. The next generation of insect-resistant plants will use novel insecticidal principles to target important insect pests that have escaped the first strike of insect control biotechnology.

Insect control-driven technology. Although *Agrobacterium*-mediated transformation has been a well-established method to introduce foreign genes into plants[2], this method has been mostly restricted to dicotyledonous plants. Technology based on direct gene transfer[3] where DNA transformation is mediated by non-biological means such as polyethylene glycol-mediated transformation, electroporation, or microprojectile bombardment has been developed to transform monocotyledonous plants. Advances in tissue culture have been combined with improvements in transformation technology to increase transformation efficiencies[4]. In the case of maize[5] and cotton[6], two of the most important crops, the technology has been pushed even further to accomplish the integration of the insecticidal genes directly into elite cultivars.

Following transformation, a critical factor required for insecticidal plants is the proper expression of the insecticidal gene. Most of the insecticidal genes currently in use are of bacterial origin, in particular from *Bacillus thuringiensis* (*Bt*)[7], whose high A/T content reduced their expression in plants for a variety of reasons[8]. Improvements in the expression of insecticidal proteins have been accomplished by increasing the G/C content of their encoding genes and/or by using plant preferred codons[8,9]. A higher G/C content *per se* does not necessarily result in an increase of gene expression unless improper splice sites and/or improper polyadenylation signals are removed during the re-synthesis of the gene. Some groups have gone even further by synthesizing insecticidal genes using crop-specific preferred codons[5,8]. Because of these advances in transformation and gene expression of foreign genes, the technology is at a point where the potential exists to virtually transform any crop with a crop tailored gene.

First generation of insecticidal proteins

The *Bt* odyssey. The concept of creating insect resistant plants begins with identifying proteins with insecticidal properties. The soil microorganism *Bacillus thuringiensis* (*Bt*) has proven to be a rich source for insecticidal proteins and genes[7]. During the sporulation phase, *Bt* produces parasporal crystals that consist of about 130 kDa proteins known as δ-endotoxins[7]. δ-endotoxins are solubilized and processed in the insect midgut to active forms of the toxin of about 65–75 kDa comprising the N-terminus of the proteins. They exert their toxicity by binding to midgut epithelial cells and ultimately causing osmotic lysis through pore formation in the cell membrane[10], a highly effective mode of action. Known Bt strains contain a great diversity of δ-endotoxins encoding genes—a total of 96 (ref. 11) have been described so far and more are being reported routinely—and have proven to be the source *par excellence* of insecticidal principles to be used in transgenic plants. Bt has been used as an insecticide for more than 40 years, but it was the cloning and sequencing of the first insecticidal protein encoding genes in 1981 (ref. 12) that raised the prospect of using these

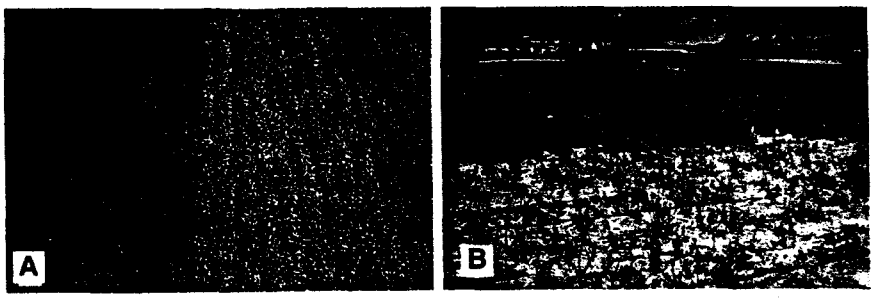


Figure 1. The first generation of built-in insect control plants, coined "*Bt* plants". (A) *Bt* cotton; the white area (right) represents the portion of the field planted with transgenic cotton expressing the Cry1A(c) protein, while the brown area (left) is planted with non-transformed cotton. (B) *Bt* potato; field performance of transgenic potatoes expressing the Cry3A protein, which confers tolerance to Colorado potato beetle (background) while non-transformed potatoes show heavy damage by the beetle (foreground). (C) *Bt* maize; side-by-side comparison in a field performance of transgenic maize expressing the Cry1A(b) protein, which confers high level of protection against European corn borer (left) versus non-transformed maize (right).

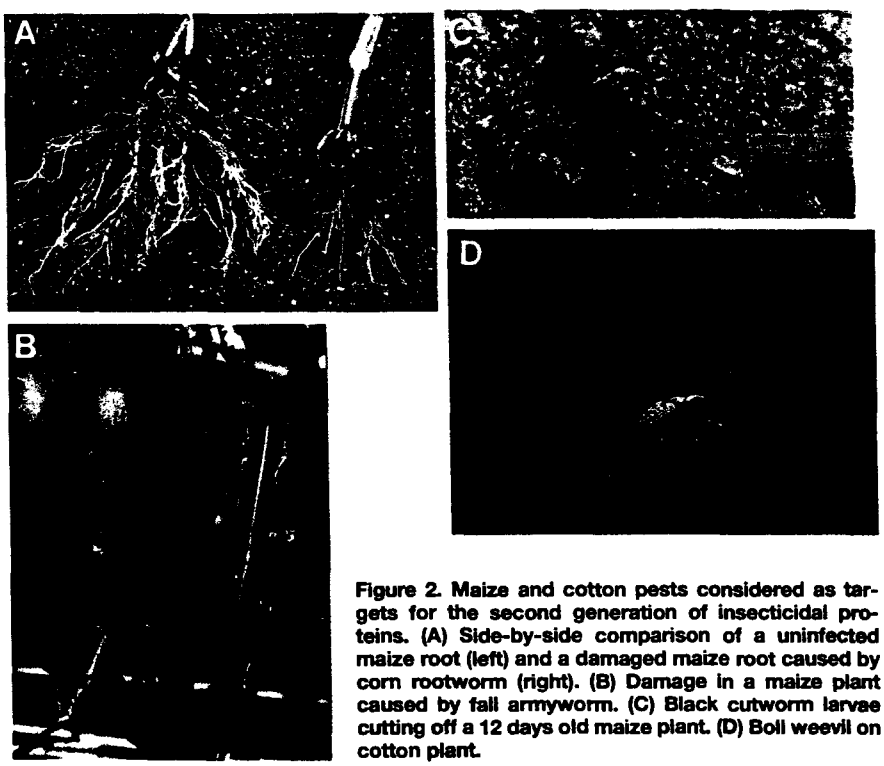


Figure 2. Maize and cotton pests considered as targets for the second generation of insecticidal proteins. (A) Side-by-side comparison of a uninfected maize root (left) and a damaged maize root caused by corn rootworm (right). (B) Damage in a maize plant caused by fall armyworm. (C) Black cutworm larvae cutting off a 12 days old maize plant. (D) Boll weevil on cotton plant.

insecticidal proteins in transgenic plants.

Bt microbials have a long history of effectiveness and safety against important agronomic insect pests[7]. This contributed to the development of *Bt* plants among the first plant biotechnology products with commercial relevance. The *Bt* genes *cry1Aa*, *cry1Ab* or *cry1Ac*, encoding δ-endotoxins isolated from the lepidopteran active *Bt* subsp. *kurstaki*, were introduced into tobacco[13], tomato[14] or potato[15]. Both the full-length and truncated forms of the toxins were introduced by *Agrobacterium*-mediated transformation, and their expression conferred some degree of protection against tobacco pests (*Manduca sexta*)[13], tomato pests (*Heliothis virescens* and *Helicoverpa zea*)[14] and potato pests (*Phthorimaea operculella*)[15]. By the end of 1990, at least seven research groups had conducted field trials with transgenic plants expressing *Bt* proteins[7]. These reports served at least two purposes. First, they demonstrated that genes encoding *Bt* insecticidal proteins could be expressed in transgenic plants. Second, they showed that the levels of insecticidal proteins (a few ng per mg of protein) attained in transgenic plants transformed with native *Bt* genes were relatively low and, with some exceptions[16,17], generally not sufficient to provide adequate field protection from agronomically relevant insect pests.

Efforts to enhance *Bt* gene expression in plants became a priority for many groups involved in insect control programs. *Bt* genes are A/T rich while plant genes tend to have a higher G/C content. Therefore, partial or full re-synthesis of the genes to contain a higher G/C content solved this problem and allowed the codon usage to be adapted for a particular crop[5]. The resulting partial or fully-synthetic gene would encode exactly the same protein as the native gene. In 1990, Perlak et al.[18] reported on the performance of cotton plants (var. Coker 312) transgenic for *cry1Ab* or *cry1Ac* with effective control of cotton pests such as cotton bollworm (*H. zea*). This was accomplished by increasing—by as much as 100-fold—the level of insecticidal protein in the transgenic plant (hundreds of ng per mg of protein) through a combination of powerful promoters (i.e. CaMV 35S with duplicated enhancers) and sequence modifications in areas of the gene with predicted mRNA secondary structures. Transformed cotton lines based on these improvements were field tested in 1992, and shown to provide good field protection against cotton bollworm and pink bollworm (*Pectinophora gossypiella*)[19] (Fig. 1A). A similar modified version of another δ-endotoxin-encoding gene, *cry3A*[20], isolated from the coleopteran active Bt subsp. *tenebrionis*, was transformed into potato plants conferring protection from damage by the Colorado potato beetle (*Leptinotarsa decemlineata*) under high levels of natural field infestations. Improved plant expression of the *cry3A* gene was achieved by increasing its overall G/C content from 36% to 49%[21]. Transformed potatoes producing Cry3A from the synthetic gene provided field protection against Colorado potato beetle larvae[20] (Fig. 1B). These initial accomplishments were restricted to dicotyledonous plants. Expression of *Bt* genes in monocotyledonous plants was advanced when Koziel et al.[5], transformed elite cultivars of maize with a truncated *cry1Ab* gene that completely replaced the bacterial codons with maize-preferred codons. As a consequence, the G/C content of the synthetic gene was increased from 37% to 65%. These transgenic maize plants provided excellent protection against European corn borer (*Ostrinia nubilalis*) even when challenged in the field with over 2,000 larvae per plant (Fig. 1C), an insect pressure several hundred fold higher than natural infestations. These plants combined the

potency of a maize-optimized gene with the accuracy of tissue-specific promoters, thus targeting Cry1A(b) production to maize tissues relevant for European corn borer control. These first generation insecticidal plants were introduced to the market during the 1996 growing season.

Managing insect resistance to transgenic *Bt* plants. Insects have demonstrated a high capacity to develop resistance to a wide array of chemical insecticides[22]. More recently, field populations have been shown to be equally adept at developing resistance to microbial sprays based on the *Bt* δ-endotoxins[23]. These cases are limited and have been associated with frequent and prolonged use of the microbial products on geographically isolated insect populations. Nevertheless, there is an industry-wide recognition of the need for research to address the potential for resistance to *Bt* δ-endotoxins.

The first step in resistance management is to establish the target pest's baseline susceptibility to the insecticidal protein. The baseline should be determined by geographical location, insect species and selective agent used. The selective agents, whether transgenic plants or microbial sprays, should be precisely characterized as to the nature and amount of δ-endotoxins produced. Baselines should be established on field populations, not in domesticated laboratory cultures derived from geographically distant populations. Once a baseline has been established, regular monitoring can be used to detect changes in susceptibility that may indicate early stages of resistance in local insect populations.

Pest population monitoring will most likely be combined with a variety of approaches aimed at minimize the potential for resistance to develop to transgenic plants. Transgenic plants producing very high levels of insecticidal proteins should be produced in an effort to eliminate heterozygotes carrying a resistance allele. Second, a refugia strategy should be adopted to provide a source of susceptible insects for mating with the selected population to prevent the fixation of resistance. Refuges can naturally occur when a certain percentage of an agricultural acreage is planted with non-transgenic plants. Refuges may also be established by design through the use of seed mixtures or susceptible border rows, or may occur outside the crop for highly mobile insect species with alternate host plants. The success of this targeted production will depend on the specifics of the insect-crop interaction and the nature of the product. Third, transgenic plants producing additional insecticidal proteins with either different modes of action, different targets, or both will be developed. Several options would be available that include the introduction of multiple genes with unique modes of action into the same plant or alternatively, independent plants each with a unique insecticidal protein could be used in rotation.

Other insecticidal proteins. Although most of the insecticidal plants have been based on *Bt* δ-endotoxins, many research projects are aimed at discovering non-*Bt* proteins to control insect pests. A number of these insecticidal proteins interfere with the nutritional needs of the insect. For instance, polyphenol oxidases[24] generate toxic compounds from dietary components; proteinase inhibitors[25] and α-amylase inhibitors[26] deprive the insect of nutrients by interfering with digestive enzymes of the insect. Proteinase inhibitors have received ample attention because their small size, abundance, and stability make them easy to work with[27]. They are usually highly specific for a particular class of digestive enzymes. There have been many reports of transgenic plants expressing proteinase inhibitors that act primarily as insect growth retardants[28]. However, insects have proven to be flexible enough to switch proteinase composition[29] in their guts to overcome transgenic plants expressing a particular proteinase inhibitor. The insecticidal properties of chitinases have also been explored. Expression of insect chitinases in tobacco plants resulted in some reduced damage by tobacco budworm[30]. After ingestion, the chitinases are presumed to target chitin structures such as the peritrophic membrane, a crucial sieve protecting the delicate midgut cells present in the insect gut lumen. Insects are continuously regenerating the peritrophic membrane[31]. Therefore it might take large amounts of chitinases to effectively inhibit this process.

Lectins constitute another large family of proteins with certain insecticidal properties[32]. Indeed, lectins captured the interest of many research groups as an alternative to *Bt* δ-endotoxins[33]. Although the mechanism of insect toxicity of lectins remains unclear, their ability to bind glycosylated proteins on the insect midgut is well characterized[34]. The most active lectins afford good insecticidal properties only when insects are exposed to μg levels in diet incorporation bioassays[33,35]. This is at least one order of magnitude higher than the activity of δ-endotoxins on susceptible insects[7]. Hence, although a number of lectin encoding genes have been expressed in transgenic plants, their insecticidal performance has been disappointing due to the level that must be produced to make them effective. Maize plants expressing wheat germ agglutinin, jacalin or rice lectin, three leading members of the lectin family, when tested for European corn borer and/or southern corn rootworm control (*Diabrotica undecimpunctata*) demonstrated modest larval growth inhibition, with very little mortality[36].

Insecticidal proteins with acute versus chronic bioactivity. The effectiveness of insecticidal proteins produced in transgenic plants is determined by their potency toward target insects. Potency is determined by measuring LC_{50} values using known quantities of the insecticidal protein in insect diet bioassays. The LC_{50} values obtained from in vitro experiments are excellent predictors of the insecticidal protein performance in transgenic plants. In these assays, while *Bt* δ-endotoxins provide activity in the range of 50–500 ng/ml of diet (ppb values)[7], proteinase inhibitors, chitinases and lectins show insecticidal properties at mg/ml (hundreds or thousands ppm)[37–39]. It appears that the potency of a given insecticidal protein can be explained by its mode of action. *Bt* δ-endotoxins are proteins that form pores in membranes causing cell lysis[10]. A few δ-endotoxin molecules will be sufficient to challenge the osmotic integrity of midgut epithelia cells leading to cell lysis and death. Therefore, we define the δ-endotoxins as having acute bioactivity. Proteinase inhibitors and lectins most likely bind to their respective targets stoichiometrically. Since gut cells and/or gut fluids are a source of multiple binding sites, and binding does not directly elicit cell death, a large amount of protein and a long-term, chronic exposure is needed to manifest any insecticidal property. Insecticidal proteins that only afford a chronic bioactivity may not be expressed in transgenic plants at high enough levels to render adequate insect resistance.

Second generation of insecticidal plants

Bt has been, and still is, a remarkable source of insecticidal proteins. However, some economically important insect pests such as northern and western corn rootworm (*D. longicornis barbieri*, and *D. virgifera virgifera* respectively) and boll weevil (*Anthonomus grandis*) are not effectively controlled by any known *Bt* δ-endotoxin. A large number of *Bt* proteins possess some insecticidal activity against specific insects but the quantities needed for effective control may not be attainable in transgenic plants, at least with current technology. Accordingly, there has been an active effort to find and characterize novel insecticidal proteins. One strategy is to screen bacterial isolates for production of insecticidal proteins in physiological stages of bacterial growth other than the typical sporulation stages where production of δ-endotoxins occurs. Another strategy is to screen new sources for potential insecticidal proteins in a random fashion. Sources for screening include plant

samples, particularly tropical plants with well-known insecticidal properties, microbial fermentation broths or proteins acquired from existing protein libraries.

VIPS, vegetative insecticidal proteins. With respect to destructive corn pests, species of corn rootworm top the list[40], particularly northern and western corn rootworm. Rootworm larvae cause most of the plant damage by feeding exclusively on corn roots (Fig. 2A), and an infestation, therefore, develops in a subtle way. Black cutworm and fall armyworm are important corn insect pests with sporadic but highly destructive outbreaks. During early stages, cutworms feed on weeds or other crops surrounding the corn fields. As larvae mature, they move into corn fields, severing young corn plants at ground level (Fig. 2B). Armyworms are voracious foliar feeders that will migrate in mass to fresh sources of food defoliating crops entirely (Fig. 2C). They damage crops so rapidly that it is too late when farmers realize their presence. Damages caused by rootworms and cutworms combined account for over $1 billion/year in crop loss and preventive insecticide treatment costs[41]. Unfortunately, rootworms and cutworms are quite recalcitrant to the known family of insecticidal proteins, including *Bt* δ-endotoxins.

In the search of novel insecticidal activities, it was discovered that clarified culture supernatant fluids collected during vegetative (i.e. log-phase) growth of *Bacillus* species are a rich source of insecticidal activities[42,43]. Supernatant fluids of certain *B. cereus* isolates possess acute bioactivity against Northern and Western corn rootworms[42]. The insecticidal properties are associated with a binary system, whose two proteins are termed Vip1 and Vip2. Similarly, supernatant fluids of certain *B. thuringiensis* cultures had potent insecticidal properties against black cutworm and fall armyworm[43]. The cloning and characterization of the gene encoding this activity yielded a novel insecticidal protein, Vip3A, that has no homology with any known protein. The N-terminal sequences of the Vip proteins possess a number of positively charged residues followed by a hydrophobic core region that are similar to other signal peptides described for *Bacillus*[42,43]. Thus, the molecular and biological properties of the Vip proteins, that they are secreted and produced during vegetative stages[42,43], establish them as distinct from insecticidal proteins that belong to the δ-endotoxin family. Potency studies performed on Vip proteins show that they afford acute bioactivity against susceptible insects in the range of nanograms per ml of diet[42,43], a level comparable to that exerted by δ-endotoxins[7]. In particular, Vip3A displays insecticidal activity against a wide-spectrum of lepidopteran insects, particularly black cutworm, fall armyworm and beet armyworm (*S. exigua*) with LC_{50} between 30–100 ng per cm^2 (ref. 43). An important aspect of the biology of insecticidal proteins has been the characterization of the tissue they target as well as the histopathology developed by susceptible insects upon their ingestion. Histopathology of susceptible insects, fed a diet supplemented with Vip3A, showed gut paralysis followed by complete lysis of the gut epithelium cells, which resulted in larval death[44]. The midgut epithelium cells, in particular the columnar cells, seem to be targeted by the Vip3A protein becoming distended and bulbous. A complete degeneration of the epithelium columnar cells was observed in second instar larvae of black cutworm fed on Vip3A-containing diet after 48 hours. Although the symptomatology developed upon Vip3A ingestion resembles that of δ-endotoxins, its timing seems to be delayed[10,44]. In vivo immunological studies have revealed specific binding of the Vip3A protein to gut epithelium cells of susceptible insects, particularly to columnar cells, supporting the histological observations.

Cholesterol oxidases, a new family of insecticidal proteins. Boll weevil is a major cotton pest worldwide[45] (Fig. 2D). In the US, boll weevil outbreaks are so devastating that they continue to shape the geographic distribution of US cotton production. A massive chemical pesticide program intended to eradicate boll weevil from the US cotton fields was started in the 1970's and was based upon blanketing cotton growing areas with chemical insecticides[46]. This program met with only partial success. Plants resistant to boll weevil would work well in concert with integrated pest management strategies. Boll weevil eggs are oviposited within the cotton flower buds, where the larvae undergo their complete development. Conventional insecticides have almost no access to larvae. Therefore a transgenic plant directing production of insecticidal proteins to the flower bud (where the pest feeds) would be ideal.

A random screen of filtrates from microbial fermentations uncovered an association between sterol oxidases and insects. A protein present in two *Streptomyces* culture filtrates showed acute toxicity to boll weevil larvae[47]. This protein turned out to be a cholesterol oxidase (CO), a member of a large family of acyl sterol oxidases. It was soon determined that its LC_{50} on first and second instar larvae of boll weevil was comparable to the LC_{50} reported for *Bt* proteins on *Bt* susceptible insects[47]. Histological studies performed on boll weevil larvae fed a diet containing CO showed alterations in their midgut epithelium cells[47]. The midgut epithelium of larvae exposed to CO exhibited cellular attenuation accompanied by local cytolysis. The earliest cytological symptoms seem to be localized at the tip of the midgut cells, which are in contact with the midgut lumen, suggesting that CO alters the cholesterol already incorporated into the membrane[48]. CO catalyzes the oxidation of cholesterol to produce ketosteroids and hydrogen peroxide[48]. Cholesterol is required for the integrity and normal function of virtually all cellular membranes and, therefore, any interference in its availability for incorporation into and maintenance of cellular integrity could lead to toxicity[49]. These observations suggest cholesterol (or a related sterol) in the membrane of the boll weevil midgut epithelium is accessible to CO, and its oxidation compromises the integrity of their membranes resulting in cell lysis and death. Again, the midgut epithelium seems to be the primary target for the CO. Considering that the midgut represents the organ in insects where critical functions like digestion and absorption take place, interference with the midgut function by disrupting the gut cells seems to be the strategy adopted by the most effective insecticidal proteins[10,44,47]. There are also studies on the effects of CO on boll weevil adult insects[50]. Although newly emerged adult boll weevils exposed to CO did not exhibit acute toxicity, they displayed severe oostatic effects (i.e., a reduction in fecundity and oviposition) due to poor development of ovaries in adult females[50].

Enzymatically active cholesterol oxidase was detected in extracts of tobacco protoplasts transformed with the native cholesterol oxidase gene[51,52] indicating that high G/C content genes (isolated from *Streptomyces*) can be readily expressed in plants.

Summary and future directions

The recent introduction of transgenic plants containing insect resistance genes has launched a new era in agriculture. The developing technology allows the transfer of a genetic trait into plants from vastly different species. The first genes, encoding insecticidal proteins of *Bacillus thuringiensis*, have served as a model for understanding gene expression in plants, developing plant transformation systems, realizing the importance of germplasm manipulation, and understanding the field efficacy and environmental impact of transgenic plants. The expression of very effective insecticidal proteins by plants delivers a remarkable level of insect control unsurpassed by any other insecticidal treatment. In the transgenic plant approach, concepts like timing, doses, scouting, directions of application, so often used in the conventional treatments, become less relevant.

We expect to expand our technology base by incorporating

novel and more powerful promoters, new and tighter tissue-specific promoters as well as a better understanding of the principles governing plant gene expression. Current and novel Bt δ-endotoxins are fully expected to be part of the transgenic plant approach to combat pests, simply because they work. Proteins such as cholesterol oxidases and members of the Vips family represent the second generation of insecticidal transgenes that will compliment the novel Bt δ-endotoxins.

Acknowledgments

We thank the members of the Insect Control Group and Registration Team of Ciba Seeds, whose work has contributed to the content of this review. We thank Brian Carr for the critical review of the manuscript. We thank Don Allemann (Ciba Crop Protection) for providing the photograph of boll weevil, and Jennifer Feldman (NatureMark Potatoes) for providing photographs of Bt *potato and* Bt *cotton.*

1. Oerke, E.C. 1994. Estimated crop losses due to pathogens, animal pests and weeds, pp. 72–88 in *Crop production and crop protection: estimated losses in major food and cash crops.* Oerke, E.C., Dehne, H.W., Schönbeck, F., and Weber, A. (eds.). Elsevier, Amsterdam.
2. Klee, H., Horsch, R., and Rogers, S. 1987. Agrobacterium-mediated plant transformation and its further applications to plant biology. *Annu. Rev. Plant Physiol.* **38**:467–486.
3. Potrykus, I. 1991. Gene transfer to plants: Assessment of published approaches and results. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42**:205–225.
4. Randolph-Anderson, B., Boynton, J.E., Dawson, J., Dunder, E., Eskes, R., Gillham, N.W., Johnson, A., Perlman, P.S., Suttie, J., and Heiser, W.C. 1995. Submicron gold particles are superior to larger particles for efficient biolistic transformation of organelles and some cell types. Bio-Rad US/EG Bulletin 2015.
5. Koziel, M.G., Beland, G.L., Bowman, C., Carozzi, N.B., Crenshaw, R., Crossland, L., Dawson, J., Desai, N., Hill, M., Kadwell, S., Launis, K., Lewis, K., Maddox, D., McPherson, K., Meghji, M.R., Merlin, E., Rhodes, R., Warren, G.W., Wright, M., and Evola, S.V. 1993. Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Bio/Technology* **11**:194–200.
6. Umbeck, P.F. 1992. Genetic engineering of cotton plants and lines. *United States Patent* 5,159,135.
7. Koziel, M.G., Carozzi, N.B., Currier, T.C., Warren, G.W., and Evola, S.V. 1993. The insecticidal crystal proteins of *Bacillus thuringiensis*: past, present and future uses. *Biotechnol. and Genet. Engineer. Rev.* **11**:171–228.
8. Koziel, M.G., Carozzi, N.B., and Desai, N. 1997. Optimizing expression of transgenes with an emphasis on post-transcriptional events. *Plant Mol. Biol.*, in press.
9. Perlak, F.J., Fuchs, R.L., Dean, D.A., McPherson, S.L., and Fischhoff, D.A. 1991. Modification of the coding sequence enhances plant expression of insect control protein genes. *Proc. Natl. Acad. Sci. USA* **88**:3324–3328.
10. Gill, S.S., Cowles, E.A., and Pietrantonio, P.V. 1992. The mode of action of *Bacillus thuringiensis* endotoxins. *Annu. Rev. Entomol.* **37**:615–636.
11. World Wide Web address: www.susx.ac.uk/users/bafn6/bt/index.html
12. Schnepf, H.E. and Whiteley, H.R. 1981. Cloning and expression of the *Bacillus thuringiensis* crystal protein gene in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78**:2893–2897.
13. Vaeck, M., Reynaerts, A., Höfte, H., Jansens, S., De Beuckeleer, M., Dean, C., Zabeau, M., van Montagu, M., and Leemans, J. 1987. Transgenic plants protected from insect attack. *Nature* **327**:33–37.
14. Fischhoff, D.A., Bowdish, K.S., Perlak, F.J., Marrone, P.G., McCormick, S.M., Niedermaeyer, J.G., Dean, D.A., Kusano-Kretzmer, K., Mayer, E.J., Rochester, D.E., Rogers, S.G., and Fraley, R.T. 1987. Insect tolerant transgenic tomato plants. *Bio/Technology* **5**:807–813.
15. Peferoen, M. 1992. Engineering of insect resistant plants with *Bacillus thuringiensis* crystal protein genes, pp. 135–153 in *Biotechnology in agriculture*. Gatehouse, A., Hilder, V., and Boulter, D. (eds.). CAB International, Wallingford.
16. Warren, G.W., Carozzi, N.B., Desai, N., and Koziel, M.G. 1992. Field evaluation of transgenic tobacco containing a *Bacillus thuringiensis* insecticidal protein gene. *J. Econ. Entomol.* **5**:1651–1659.
17. Delannay, X., La Vallee, B.J., Proksch, R.K., Fuchs, R.L., Sims, S.R., Greenplate, J.T., Marrone, P.G., Dodson, R.B., Augustine, J.J., Layton, J.G., and Fischhoff, D.A. 1989. Field performance of transgenic tomato plants expressing the *Bacillus thuringiensis* var. *kurstaki* insect control protein. *Bio/Technology* **7**:1265–1269.
18. Perlak, F.J., Deaton, R.W., Armstrong, T.A., Fuchs, R.L., Sims, S.R., Greenplate, J.T., and Fischhoff, D.A. 1990. Insect resistant cotton plants. *Bio/Technology* **8**:939–943.
19. Wilson, F.D., Flint, H.M., Deaton, R.W., Fischhoff, D.A., Perlak, F.J., Armstrong, T.A., Fuchs, R.L., Berberich, S.A., Parks, N.J., and Stapp, B.R. 1992. Resistance of cotton lines containing a *Bacillus thunngiensis* toxin to pink bollworm (Lepidoptera: Gelechiidae) and other insects. *J. Econ. Entomol.* **85**:1516–1521.
20. Perlak, F.J., Stone, T.B., Muskopf, Y.M., Petersen, L.J., Parker, G.B., McPherson, S.A., Wyman, J., Love, S., Reed, G., Biever, D., and Fischhoff, D.A. 1993. Genetically improved potatoes: protection from damage by Colorado potato beetles. *Plant Mol. Biol.* **22**:313–321.
21. Adang, M.J., Brody, M.S., Cardineau, G., Eagan, N., Roush, R.T., Shewmaker, C.K., Jones, A., Oakes, J.V., and McBride, K.E. 1993. The reconstruction and expression of a *Bacillus thuringiensis cryIIIA* gene in protoplasts and potato plants. *Plant Mol. Biol.* **21**:1131–1145.
22. Moberg, W.K. 1990. Understanding and combating agrochemical resistance, pp. 3–16 in *Managing Resistance to Agrochemicals*. Green, M.B., LeBaron, H.M., and Moberg, W.K. (eds.). ACS Symposium Series, Washington.
23. Shelton, A.M., Robertson, J.L., Tang, J.D., Perez, C., Eigenbrode, S.D., Preisler, H.K., Wilsey, W.T., and Cooley, R.J. 1993. Resistance of the diamondback moth (Lepidoptera: Plutellidae) to *Bacillus thuringiensis* subespecies in the field. *J. Econo. Entomol.* **86**:697–705.
24. Felton, G., Donato, K., Broadway, R., and Duffey, S. 1992. Impact of oxidized plant phenolics on the nutrional quality of dietary protein to a noctuid herbivore, *Spodoptera exigua*. *J. Insect. Physiol.* **38**:277–285.
25. Hilder, V., Gatehouse, A., Sheerman, S., Barker, R., and Boulter, D. 1987. A novel mechanism of insect resistance engineered into tobacco. *Nature* **330**:160–163.
26. Huesing, J.E., Shade, R.E., Chrispeels, M.J., and Murdock, L.L. 1991. α-Amylase inhibitor, not phytohemagglutinin, explains the resistance of common bean seeds to cowpea weevil. *Plant Physiol.* **96**:993–996.
27. Ryan, C. 1990. Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Ann. Rev. Phytopathol.* **28**:425–449.
28. Johnson, R., Narvaez, J., An, G., and Ryan, C. 1989. Expression of proteinase inhibitors I and II in transgenic tobacco plants: effects on natural defenses against *Manduca sexta* larvae. *Proc. Natl. Acad. Sci. USA* **86**:9871–9875.
29. Jongsma, M.A., Bakker, P.L., Peters, J., Bosch, D., and Stiekema, W.J. 1995. Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci. USA* **92**:8041–8045.
30. Ding, X. 1995. *Manduca* chitinase-mediated resistance to tobacco budworm (*Heliothis virescens*) and tobacco hornworm (*Manduca sexta*) larvae in transgenic tobacco plants. Ph.D. Dissertation, Kansas State University, Manhattan, pp. 73.
31. Chapman, R.F. 1985. Structure of the digestive system, pp. 165–205 in *Comprehensive insect physilogy, biochemistry and pharmacology*. Kerkut, G.A. and Gilbert, L.I. (eds.). Pergamon Press, Oxford.
32. Chrispeels, M.J. and Raikhel, N.V. 1991. Lectins, lectin genes and their role in plant defense. *The Plant Cell* **3**:1–19.
33. Cavalieri, A., Czapla, T., Howard, J., and Rao, G. 1995. Larvicidal lectins and plant insect resistance basased thereof. *United States Patent* 5,407,454.
34. Harper, S.M., Crenshaw, R.W., Mullins, M.A., and Privalle, L.S. 1995. Lectin binding to insect brush border membranes. *J. Econ. Entomol.* **88**:1197–1202.
35. Czapla, T. and Lang, B.A. 1990. Effect of plant lectins on the larval development of European corn borer (Lepidoptera: Pyralidae) and southern corn rootworm (Coleoptera: Chrysomelidae). *J. Econo. Entomol.* **83**:2480–2485.
36. Maddock, S.E., Hufman, G., Isenhour, D.J., Roth, B.A., Raikhel, N.V., Howard, J.A., and Czapla, T.H. 1991. Expression in maize plants of wheat germ agglutinin, a novel source of insect resistance. 3rd Int. Cong. Plant Mol. Biol., Tucson, Arizona.
37. Boulter, D. 1993. Insect pest control by copying nature using genetically engineered crops. *Phytochemistry* **34**:1453–1466.
38. Shah, D.M., Rommens, C.M., and Beachy, R.N. 1995. Resistance to diseases and insects in transgenic plants: progress and applications to agriculture. *Trends Biotechnol.* **13**:362–368.
39. Kanost, M. and Jiang, H. 1996. Proteinase inhibitors in invertebrate immunity, pp. 45–69 in *New directions in invertebrate immunology*. Soderhall, K., Iwanaga, S., and Vasta, G. (eds.). SOS Publications, New Jersey.
40. Luckmann, W.H. 1982. Integrating the cropping system for corn insect pest management, pp. 499–519 in *Introduction to insect pest management*. Metcalf, R.L. and Luckmann, W.H. (eds.). John Wiley & Sons, New York.
41. Metcalf, R.L. 1986. Forword, pp. vii–xv in *Methods for the study of pest Diabrotica*. Krysan, J.L. and Miller, T.A. (eds.). Springer-Verlag, New York.
42. Warren, G.W., Koziel, M.G., Mullins, M.A., Nye, G.J., Carr, B., Desai, N., Kostischka, K., Duck, N.B., and Estruch, J.J. 1996. Novel pesticidal proteins and strains. *World Intellectual Property Organization* WO 96/10083.
43. Estruch, J.J., Warren, G.W., Mullins, M.A., Nye, G.J., Craig, J.A., and Koziel, M.G. 1996. Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proc. Natl. Acad. Sci. USA* **93**:5389–5394.
44. Yu, C.G., Mullins, M.A., Warren, G.W., Koziel, M.G., and Estruch, J.J. 1997. The *Bacillus thuringiensis* vegetative insecticidal protein Vip3A lyses midgut epithelium cells of susceptible insects. *Appl. Environ. Microbiol.* **63** (in press).
45. Luttrell, R.G., Fitt, G.P., Ramalho, F.S., and Sugonyaev, E.S. 1994. Cotton pest management: a worldwide perspective. *Annu. Rev. Entomol.* **39**:517–526.
46. Perkins, J.H. 1980. Boll weevil eradication. *Science* **207**:1044–1050.
47. Purcell, J.P., Greenplate, J.T., Jennings, M.G., Ryerse, J.S., Pershing, J.C., Sims, S.R., Prinsen, M.J., Corbin, D.R., Tran, M., Sammons, R.D., and Stonard, R.J. 1993. Cholesterol oxidase: a potent insecticidal protein active against boll weevil larvae. *Biochem. Biophys. Res. Commun.* **196**:1406–1413.
48. Corbin, D.R., Greenplate, J.T., Jennings, M.G., Purcell, J.P., and Sammons, R.D. 1995. Method of controlling insects. *World Intellectual Property Organization* WO 95/01098.
49. Linder, R. and Bernheimer, A.W. 1984. Enzymatic oxidation of membrane cholesterol in relation to lysis of sheep erythrocytes by corynebacterial enzymes. *Archs. Biochem. Biophys.* **213**:395–404.
50. Greenplate, J.T., Duck, N.B., Pershing, J.C., and Purcell, J.P. 1995. Cholesterol oxidase: An oostatic and larvicidal agent active against the cotton boll weevil, *Anthonomus grandis*. *Entomol. Exper. Et Aplicata* **74**:253–258.
51. Corbin, D.R., Greenplate, J.T., Wong, E.Y., and Purcell, J.P. 1994. Cloning of an insecticidal cholesterol oxidase gene and its expression in bacteria and in plant protoplasts. *Appl. Environ. Microbiol.* **60**:4239–4244.
52. Cho, H.-J., Choi, K.-P., Yamashita, M., Morikawa, H., and Murooka, Y. 1995. Introduction and expression of the *Streptomyces* cholesterol oxidase gene (*choA*), a potent insecticidal protein active against boll weevil larvae, into tobacco cells. *Appl. Microbiol. Biotechnol.* **44**:133–138.

Critical Reviews in Plant Sciences, 14(2):149–178 (1995)

In Vitro Culture of Wheat and Genetic Transformation — Retrospect and Prospect

N. Maheshwari,[1] K. Rajyalakshmi,[2] K. Baweja,[1] S. K. Dhir,[3] C. N. Chowdhry, and S. C. Maheshwari [1,2]

[1]Department of Plant Molecular Biology, South Campus and [2]Department of Botany, University of Delhi; [3]Monsanto, 700 Chesterfield Village Parkway, St. Louis, MO 63198, U.S.A.

Referee: Dr. K. J. Scott, Department of Biochemistry, University of Queensland, John Hines Building, Brisbane Qld 4072, Australia

ABSTRACT: This review focuses on the improvement of wheat (*Triticum aestivum* L.) via tissue and cell culture and its use in gene transfer techniques. Success of the latter critically depends on the ability to regenerate plants from cells or tissues cultured *in vitro*. Hence, we have devoted attention to the attempts made so far in obtaining regenerants from diverse explants. Although it is known that immature embryos are the best source for initiating morphogenic cultures, basic information related to the process of differentiation can also be gained by studying less responding tissues. The opportunity provided by anther and microspore culture in wheat improvement and the progress made is also presented.

To enhance tissue culture responses, identification of chromosomes, gene loci, and genes is of cardinal importance. We have also surveyed the progress made in this regard by conventional but incisive plant-breeding techniques. Gene rearrangements in tissue culture leading to the appearance of somaclonal or gametoclonal variation are of interest in selection of useful cell lines. The last part of the review is devoted to work done on transient gene expression and transformation with emphasis on recent developments.

KEY WORDS: wheat anther and pollen culture, suspension and protoplast culture, wheat somaclonal variation, genetic basis of tissue culture response.

I. INTRODUCTION

As the principal crop of the world, wheat (*Triticum aestivum* L.) has received the constant attention of scientists. Its total annual production is 565 million metric tons compared with rice, which is 527 million metric tons. For several decades, plant breeders improved it assiduously, but their efforts were reaching a plateau, especially with respect to yield. Fortunately, the field of recombinant DNA technology has potentially opened up new avenues to modify crops according to specific needs. Central to the new technology for plant improvement are *in vitro* culture techniques. Although these have existed for many years, they are now finding an essential role in genetic engineering because this technology requires the ability to regenerate plants from cells or tissues cultured *in vitro*.

In this review, we examine the achievements and problems in regenerating plants from various explants, somatic and gametic cells (especially microspores), and protoplasts of wheat. Haploid techniques are central to improvement programs of many crops. In wheat, there is intense activity in this area; hence, we have covered these developments in detail. Because the genetic makeup is an important component in tissue culture response, we also propose to review what has been learned over the years with respect to this aspect. Also included is a discussion of somaclonal and gametoclonal variation supposedly of value in finding useful traits in tissue culture. The last part of our review is devoted to the progress made in obtaining transgenic plants. The earlier gene transfer experiments depended on protoplasts, the isolation and culture of which is very demanding. However, newer methods involving faster and more convenient ways of gene delivery into plant cells and tissues are replacing or supplanting the protoplast approach.

For a detailed analysis of some of the above-mentioned aspects, one may also consult the compilation of excellent articles from various experts in a volume on wheat edited by Bajaj.[11] Although less recent, *Wheat and Wheat Improvement*, edited by Heyne,[113] is also a valuable compendium. Reviews on more limited aspects other than those in Bajaj are by

Abbreviations: ABA Abscisic acid; ACC 1-Amino-cyclopropane-1-carboxylic acid; CCC Chlormequat chloride; CH Casein hydrolysate; CW Coconut water; GUS B-glucuronidase; KN Kinetin; PAA Phenylacetic acid; 2,4-D 2,4-Dichlorophenoxyacetic acid; 2,4,5-T 2,4,5-Trichlorophenoxyacetic acid; Picloram 4-Amino-3,5,6-trichloropicolinic acid.

0735-2689/95/$.50

Picard et al.[230] on haploids, by Morrish and Fromm[197] on cereal transformation methods, and by Baum et al.[20] in *Wide Crosses in Cereals*. However, there is no recent review that deals with the biotechnology of wheat and covers the broad field as we have attempted here.

II. TISSUE CULTURE AND REGENERATION FROM VARIOUS EXPLANTS OF SOMATIC TISSUES

In the past, most of the important advances relating to tissue culture were made on dicotyledonous species. In all the exciting developments starting from the development of culture media, the use of synthetic hormones, deciphering the importance of auxin cytokinin ratios in differentiation, culture of ovaries and ovules, totipotency, single-cell culture, the discovery of somatic embryogenesis, discovery of haploids, use of protoplasts and somatic cell techniques, suspension cultures, etc. monocots figured nowhere and for many years they were given the epithet "recalcitrant" because they were not amenable to the strategies that worked well with dicots. Gradually, however, the scene changed and reports on cereal tissue culture started appearing in the late sixties (see References 159 and 271). In this section, we give a historical background of developments pertaining to use of various explants. Although immature embryos have been the explant of choice, for regeneration, establishing suspension cultures, isolating protoplasts, or obtaining transformation, it is the early work with other explants that paved the way for the modern work and therefore must be chronicled.

A. Caryopses (Seeds)

Although immature embryos are excellent material for the establishment of regeneration systems, their nonavailability throughout the year has prompted the use of seeds as a reasonably satisfactory alternate. When cultured on a medium with different phytohormones, seeds callused from scutellum and embryonal axis within a few weeks.[50,69,85,114] Only rarely did callus arise exclusively from roots.[22,202] The cultured callus was difficult to regenerate. In the earlier studies, only root regeneration and sporadic development of shoots was observed.[50] Later, enhancement in shoot regeneration in IAA and zeatin-supplemented media was reported.[85] According to Eapen and Rao[69] although both root and shoot apices form, complete plantlets were not regenerated and shoot buds probably arose from preexisting embryonal axillary shoot apices.

With high levels (20 mg/l) of 2,4-D, embryo cultures from seed showed three types of calli: sm embryogenic, nodular embryogenic, and embryogenic.[114] These were manually separable showed marked differences in their regenerat tential. As expected, the embryogenic, bu nonembryogenic, calli regenerated plants. The r eration potential of calli during long-term mainte declined with age.[114]

B. Leaf and Other Seedling Parts

1. Leaf

Leaf tissues cultured *in vitro* offer a speci vantage to investigate the molecular basis of dif tiation because they lack meristems except at the Therefore, they are useful for obtaining totally ferentiated callus. This callus can be used for ir gating factors necessary for regeneration of plants.[299]

In the studies prior to 1980, even callusing not be obtained from leaf segments.[22,206] Duri early 1980s, successful regeneration of plants leaf explants became possible.[7,297,299] There wa ference in response depending on whether t segments were taken from field-grown or *in vitro*- seedlings. A larger number of segments of th leaves callused compared with the younge soil grown plants,[299] while with *in vitro*-gro lings there was no difference in response due to a In fact, seedlings of all ages between 5 to 30 d an identical response.[7] Only the basal 10- to region responds in terms of callusing and reg tion.[7,287,299] Attempts have also been made to callus from upper segments of the leaf. With a acid in the medium, callusing could be exten 40 mm from the leaf base.[7] Cysteine an vinylpyrrolidone were ineffective. Increasing th concentration to 30 mg/l also enhanced callus tion in semidifferentiated cells but mature se did not respond.[287]

In earlier work on leaf explants, regenera only by organogenesis.[7,299] Wernicke and Milk obtained callus with structures suggestive of embryogenesis. Only recently has regener somatic embryogenesis been reported, from 2, duced callus obtained from basal segments of ond and third leaves on hormone-free mediu

To improve the response of leaf explant ture, Grossmann et al.[87] studied the influence classes of compounds, chlormequat chlorid retardants (CCC, tetcyclacis, LAB150978) a

ene-forming compounds (ethephon and 1-amino-cyclopropane-1-carboxylic acid). Donor seedlings treated with both classes of these compounds showed reduced shoot growth. With growth retardants, leaf development was delayed but the competence of the second leaf to form callus was enhanced. Ethephon treatment of either the donor seedlings or *in vitro* pretreatment of the explants substantially increased the propensity of leaf bases to produce callus.

A comparison of total protein levels and polypeptide profiles from morphogenic and nonmorphogenic leaf base calli showed significant differences in the soluble protein content on a fresh weight basis and quality as well as quantity of specific proteins.[235] Similarly, differences in isozyme profiles of several enzymes in embryogenic and nonembryogenic calli derived from immature embryos have also been reported as a selective expression of several isozymes during the course of somatic embryogenesis.[15]

a. Rapid Induction of Somatic Embryogenesis in Leaf Bases by 2,4-D

Conventionally, cereal tissues are maintained on 2,4-D medium for days or weeks and then transferred to medium without 2,4-D for regeneration. Therefore, it was unexpected that exposure to 2,4-D can be extremely brief, in order to elicit a response. In experiments with cultivar HD2329, Mahalakshmi et al.[169] showed that treatment of leaf bases with 2,4-D (10^{-5} *M*) for only 1 d is enough to trigger the appearance of structures that give rise to normal seedlings within 1 month after transfer to MS basal medium. Interestingly, at higher concentrations (10^{-4} *M*), even a 1-min application is inductive. This may be one of the most rapid hormone responses observed in tissue culture and is, essentially, genotype independent. The induction process was found to be calcium/calmodulin mediated.[169] This was indicated by the following observations: EDTA inhibits the response and the calcium ionophore A23187 enhances it. Calcium channel blockers, calcium antagonists, and calmodulin antagonists or inhibitors, as well as inhibitors of phosphoinositol metabolism, all inhibit the induction of these structures. Work is in progress to determine if the structures develop via somatic embryogenesis.

2. Mesocotyl, Nodes, and Internodes

All parts of young seedlings — roots, root tips, mesocotyl segments, stem axes, nodal and internodal regions of stem, apical meristems, and leaves — have been used as explants to initiate callus. In earlier reports there was no regeneration of plants even though callus was maintained for as long as 4 years.[266] Only roots were obtained even after several possible variables were modified: for example, hormone concentration, incorporation of substances such as corn steep liquor, lactalbumin hydrolysate, diphenylurea or CM, and altering pH, O_2 tension, or subjecting the growing callus to gentle centrifugation.[176,177,179] Interestingly, continued attachment of mesocotyl segments to the callus ensured regeneration of plants.[178]

Efficient plantlet regeneration has been obtained from mesocotyl. As much as 30 to 40% plant regeneration could be obtained; however, it was genotype dependent.[296] The size of the initial inoculum influenced ability to regenerate; however, a very small inoculum (55 to 70 mg) did not go beyond the callus stage.[67] Recently, a new method of culturing mesocotyls of mature wheat embryos was reported from which plant regeneration was obtained via organogenesis. The "germ tissues", coleoptile, and coleorhiza were surgically removed and callus was induced from mesocotyl still attached to the endosperm on medium containing 2,4-D.[18]

Callus that could be maintained for about 3 years by frequent subculture was obtained from nodal and internodal regions or stems of seedlings.[206] A few shoot buds also appeared. This was in contrast to rapid regeneration of plants obtained if coleoptilar nodes were cultured.[187] The plants were successfully transferred to soil and grown to maturity.

3. Apical Meristem

Apical meristems are superior explants compared to other seedling parts with regard to both yield of callus and plant regeneration.[120] Regeneration via somatic embryogenesis was achieved for the first time from shoot meristem cultures although a strong genotypic effect was noted.[288] The authors showed an enhanced effect of 2,4-D on morphogenesis. Callus from excised shoot meristems cultured on medium without auxin developed into shoots in a normal fashion but no roots were formed. Recently, regeneration of apical explants up to the formation of fertile plants was shown by Simmonds et al.[261]

4. Root

The first attempt to culture wheat roots was made by Burström[28] from a physiological viewpoint. Later, many workers obtained prolific callus from root seg-

ments but there was no plant regeneration irrespective of the type of culture medium, hormones, and other substituents.[41,64,120,177,179,205,206,266]

Plantlet regeneration from root cultures was first reported by Shimada et al.[254] and Adachi and Katayama.[1] Inclusion of asparagine enhanced regeneration of plantlets[22] but, as the authors state, these results were not reproducible.

More consistent regeneration from root-derived cultures was obtained by Nabors et al.[202] They selected embryogenic (isodiametric cells up to 31 μm diameter) and nonembryogenic (long, tubular cells up to 52 μm in width and 355 μm in length) calli visually; the embryogenic calli showed 33 times more regeneration than the nonembryogenic. Optimization of hormonal additives led to the production of about 1000 plantlets from an initial 26 seeds after six passages in culture.

C. Inflorescence

Calli induced from immature inflorescence explants produce comparatively more morphogenic cultures than those derived from other explants because they have a number of suppressed meristematic regions that proliferate on contact with the nutrient medium.

In the earlier work, rachises free of glumes were cultured. Regeneration was limited to obtaining shoots, 7.1%[64] or 35%.[85] Poor callusing response was obtained from young inflorescences and there were no signs of differentiation.[41] Later, Ozias-Akins and Vasil[216] reported compact callus and complete plant regeneration, although with low frequency (8%). Very improved regeneration frequencies were reported when immature inflorescences were cultured on MS medium with CM[166] or with CM and CH.[234] Depending on the genotype, the response in terms of plant regeneration varied as much as from 33 to 100%[166] or 0 to 66%.[234] Regeneration was via somatic embryogenesis.

D. Mature Embryo

Because of easy availability, mature embryos were used as explants by earlier workers. In the first report, Chin and Scott[41] obtained 7% regeneration on a hormone-free medium. The regenerants were raised from callus obtained on 2,4-D medium with parachlorophenoxyacetic acid. A combination of 1 mg/l NAA, 5 mg/l KN led to production of shoots in more than 20% of the cultures. O'Hara and Street[206] compared various genotypes and found varietal differences in the induction and extent of callusing. Other than genotype, factors that affected callus proliferation significantly were hormones (IAA, 2,4-D, 2,4,5-T, picloram, and KN) and to some extent environmental fac tors such as light, with low light intensities favorin root formation.[141] Ozias-Akins and Vasil[217] have sh that callus originates from parenchyma cells in procambial zone of mature embryos axes.

A comparison of the regenerative potential o embryonal axes (embryos without scutellum) and plu mules (coleoptile and scutellar node) has shown th while calli from plumules were regenerative, th radicles produced calli with only rooting ability.[74]

In a recent study[192] it was noted that if hig molecular-weight osmotica such as PEG are used, ag does not solidify. This can be circumvented by usin filter bridges. These proved superior to growing ex plants directly on agar because the amount of embry genic and the number of regenerants per gram of callu were almost twice as much as from cultured matur embryos.

E. Immature Embryo

Immature embryos are far more responsive th mature embryos in culture. Complete plantlets wer first regenerated from them by Shimada.[256] In a lat and more intensive study, Shimada and Yamada optimized various factors, such as age and size of th embryo and auxin concentration for callus inductio Younger embryos, about 14 d old, could be induc form callus with green spots (53.2%). These late duced shoots when transferred to 2,4-D-free mediu Embryos older than 27 d did not develop green sp or shoots. Gosch-Wackerle et al.[85] also used 14- to 1 d-old immature embryos, and shoot regeneration wa achieved using a combination of hormones (1 mg IAA and 1 mg/l zeatin). Sears and Deckard[252] a obtained plant regeneration from immature embry but their strategy involved a stepwise decrease in th 2,4-D level, finally transferring to 2,4-D minus m dium.

Although the term *somatic embryogenesis* per s was not used, such embryos were observed in imm ture embryo cultures by Ahloowalia[5] who repor normal chromosome numbers in the regenerated plant Also in 1982, the first unequivocal demonstration regeneration via somatic embryogenesis with suppo ing evidence from scanning electron microscopy wa provided by Ozias-Akins and Vasil.[216] Normal somati embryos formed in high frequency were reported the following year.[217]

Because cultures from immature embryos hav been the subject of several studies, we give belo detailed account of factors that influence their respo *in vitro*.

1. Growth Regulators

The auxin 2,4-D (1 to 2 mg/l) is routinely used for callus induction. Regeneration via somatic embryogenesis occurs upon its withdrawal from the medium. Initiation or maintenance of callus cultures on 2,4-D and a low concentration of KN was found to be better for regeneration after formation of embryogenic callus.[114,220] Substitution of 2,4-D by dicamba (3,6-dichloro-*o*-anisic acid) led to the initiation of larger calli and a large number of shoots, particularly from cultures maintained in the dark.[119,220] Surprisingly, however, the substitution of 2,4-D by dicamba in one of the media had no significant effect[32] if the number of embryoids induced per explant was scored. This may be due to rapid metabolism of dicamba in wheat tissues resulting in early germination of embryoids. Hunsinger and Schauz[119] found dicamba to be of great value both in induction of somatic embryogenesis as well as in reducing the time required for regeneration even in light-grown cultures. However, compared with other graminaceous species, they found low levels (1 mg/l) to be effective even when used without 2,4-D.

While 2,4-D induces somatic embryogenesis, other nutritional factors as well as hormones often cause precocious germination of embryos. Various protocols have been devised to suppress precocious germination and thereby enhance the frequency of somatic embryogenesis. Use of high 2,4-D concentration was recommended by Bapat et al.[14] Dicamba also effectively suppressed it as shown by Papenfuss and Carman[220] and Carman et al.[33] These workers further found that this response could be improved by incubating cultures on medium with both L-tryptophan and 2,4-D. Dicamba also had a positive effect on callus growth and along with KN markedly enhanced the shoot-forming ability.[32] In general, dicamba promoted formation of a significantly larger number of embryoids than did 2,4-D.[33] Sorbitol (0.3 *M*) also suppressed precocious germination.[245]

Interestingly, low O_2 levels (9%) also decreased precocious germination of immature embryos as did ABA (1.9 µmol/l) in the callus induction medium.[31] ABA-induced inhibition of embryoid germination could be relieved by a brief period of desiccation.[31] While precocious germination of late-stage embryos (21 to 25 d after anthesis) could be suppressed by ABA and its analogues such as methyl abscisate and methyl epoxybetaionylideneacetate,[233] these elicited a negative response in early-stage embryos (10 to 14 d after anthesis). The difference is probably related to the high endogenous levels of ABA in the early stages. Increased osmotic potential and low concentrations of ABA promote embryogenesis.[27] Using an ABA pretreatment, immature embryos cultured *in vitro* for 10 d can be cryopreserved without loss of viability. Polar diacyl lipids induced by ABA treatment may be one of the factors contributing to low temperature tolerance.[129]

2. Other Components of Medium

To maximize efficiency of regeneration, investigators have manipulated not only the basal medium itself, but have also included complex substances such as coconut water or amino acids (alone or in various combinations). A double-strength MS medium with normal or double concentration of vitamins together with casein hydrolysate and other adjuvants supported a 30% increase in the induction of embryogenic callus from which somatic embryos with typical scutella, coleoptiles, shoots, epiblasts, roots, and coleorhizae were differentiated.[218] He et al.[103] confirmed enhancement in the frequency of white callus upon increasing the concentration of macroelements from half-strength to full or double strength. Coconut water supplemented in the medium had a promotive effect on callus induction, embryoid formation, and plant regeneration.[166,183] Addition of tryptophan significantly increased embryogenic callus formation.[33,202]

Interestingly, two antibiotics, cefotaxime and carbenicillin, which are used to eliminate *Agrobacterium* in coculativation experiments, were reported to stimulate callus growth and regeneration.[181] Cefotaxime also suppressed precocious germination of embryoids. However, if autoclaved, the two antibiotics were found to be inhibitory. These results were substantiated in durum wheat where the regenerative capability of a low-responding genotype was vastly improved by inclusion of cefotaxime in the medium.[24]

A study of various carbohydrates showed that 2% sucrose supported optimal callus growth and high number of shoots.[82] It was also shown by Galiba and Yamada[83] that 40 m*M* NaCl or KCl could significantly increase the frequency of somatic embryos.

Shoot regeneration was remarkably enhanced by inclusion of $AgNo_3$ in the culture medium.[231] This effect may be due to the inhibition of ethylene action because ethylene is known to suppress shoot regeneration.

3. Age, Position of Embryo, and Ontogeny of Embryogenic Structures

The effect of embryo orientation on the medium surface has been found to influence the amount of nodular embryogenic callus. Various studies showed that there was more proliferation of embryogenic cal-

lus if the scutellum faced away from the medium and the frequency of plant regeneration was increased.[69,104,114] It is now generally known that somatic embryogenesis is influenced by embryo age; younger embryos produce comparatively more somatic embryos (see References 103 and 119).

Anatomical observations of cultured immature embryos showed that somatic embryos arise from the scutellar tissue[216] and all three tissue systems of the scutellum, namely, dermal, ground, and vascular, proliferate and the embryogenic development is similar to zygotic embryogeny.[168] In a detailed study, He et al.[103] traced the histological development of both scutellar- and epiblast-derived callus and found both somatic and organogenetic modes of embryogenesis. The organogenetic mode could be documented by the vascular connection to the basal callus. Not all regenerative outgrowths were of single-cell origin. Many embryoids were broad based and likely had a multiple-cell origin. Ryschka et al.[245] also investigated the anatomy of proliferating scutellar tissue. Single cells of the abaxial scutellar epidermis and the subepidermal layer formed embryogenic structures similar to the ones seen during the course of zygotic embryos development. The procambium also gave rise to two- to four-celled proembryos.

From the few histological investigations one can conclude that somatic embryos originate mainly from single cells, like their zygotic counterparts but a multiple-cell origin of embryogenic structures cannot be ruled out. However, this is not in line with the established principles of embryo development.

III. CELL SUSPENSION AND PROTOPLAST CULTURE

A. Suspension Culture

Cell suspension cultures were developed in the 1960s by Trione et al.[266] and Gamborg and Eveleigh[84] from root cultures. In Cocking's laboratory[7] suspension cultures were raised from leaf base callus and even a few plantlets were regenerated and transferred to soil. Subsequently, cell suspension cultures have been initiated from callus cultures derived from immature inflorescence,[147,165] immature embryos,[165,237,262,274,281] and anther.[94]

With improved protocols Vasil and co-workers[237] have been able to obtain high-frequency regeneration from cell suspensions derived from aged calli (5 to 8 months old) that are nodular and compact. The calli themselves were obtained from immature embryos. Vasil et al.[274] have been able to identify and characterize highly embryogenic cultures (types C and D) separable from the less embryogenic A and B. Types C and D yield prolific suspension cells.

A strategy adopted by Wang and Nguyen[281] to obtain high-frequency regeneration from suspension cells derived from embryo callus was to systematically discard root-forming clumps and select shoot-competent cells. Such cells, maintained for 2.5 years, when plated showed regeneration frequencies as high as 80%. In a more recent report from the same laboratory, Chang et al.[35] were also successful in establishing regenerable suspension cultures from long-term, shoot-competent cultures. Although a high yield of protoplasts was obtained from the suspension cells, the actual number of plants regenerated was very small.

Yang et al.[294] established fine embryogenic cultures derived from immature embryo callus, which maintained their regenerative potential for 20 months. This was achieved by continuous selection of embryogenic clumps at each subculture. On differentiation medium, these developed into embryoids and green plants with the normal chromosome complement.

Other than for obtaining regenerable protoplasts, suspension cultures are proving useful for selecting special cell lines. From immature embryo-derived suspensions, a thermotolerant cell line that could be grown at 48°C was selected.[278] The cell line maintained synthesis of normal proteins and showed a quantitative increase in high-molecular-weight (70 to 90 kDa) and two unique low-molecular-weight (16 to 17 kDa) heat shock proteins. Although only sterile plants could be regenerated and the regenerative ability was lost with passage of time, this kind of investigation is promising for developing wheat plants suited for growing under diverse ecological situations.

Interestingly, cell suspensions can be cryopreserved and the ability to form callus and regenerate plants can be retained after thawing.[37] Cold hardening, before cryopreservation, and prefreezing in DMSO and sorbitol increased the efficiency of survival of cells after thawing. Although some calli lost the ability to regenerate, a few produced normal plants that were subsequently transferred to soil.

Despite recent successes, the establishment and maintenance of suspension cultures is difficult. There are reports of chromosomes loss[127,262] and plants regenerated from suspension cells show polyteny, giant nuclei, genome rearrangements, minute chromosomes, and granulated chromatin.[253]

B. Protoplast Culture

Protoplast culture in monocots has proved to be far more challenging than in dicots — and mesophyll

tissue used so successfully in dicots is totally unsuitable here. In wheat mesophyll, protoplasts do not go beyond a few cell divisions.[13,57,65,75,151,159,170,207,247]

Failure of mesophyll cells to respond has led to search for alternate tissues. Based on their success with *Pennisetum*, Vasil et al.[274] advocated the use of embryogenic cultures for regeneration of protoplasts. Indeed, it is now realized that the starting material itself for obtaining protoplasts must have high regenerative capacity. If this is ensured the medium can be relatively simple and inexpensive; agarose or nurse cells or Ficoll need not be incorporated.[105] Initially, even with embryogenic tissues success was limited,[100,165] but in recent years there has been a spate of reports describing varying degrees of success in regenerating complete plants from protoplasts obtained from suspension cultures of various explants: anthers,[94] mature embryos,[278] young inflorescence,[152,239] and immature embryos.[4,35,90,91,105,153,232,274] Vasil et al.[274] showed that from aged embryogenic calli one can obtain embryogenic suspensions more efficiently and establish regenerable protoplast cultures. Chang et al.[35] also mention obtaining protoplasts from a secondary suspension culture initiated from a long-term regenerable callus culture of the cultivar Mustang. In work with a Finnish winter wheat variety "Aura", the investigators emphasize the need to employ a particular embryogenic callus type for obtaining fine suspension cultures. The protoplasts isolated from these regenerate into fertile wheat plants.[221]

From protoplasts the plants can be obtained via somatic embryogenesis and are successfully transferred to soil.[274] The frequency of regeneration is still low and appearance of fertile plants is somewhat rare. Qiao et al.[232] raised a few plants to the heading stage, but whether they were fertile was still to be determined. In another report,[4] out of 500 plants transplanted to a greenhouse, only one was fertile. Li et al.[153] also transferred regenerants to soil, but no plant grew to maturity.

In summary, success in wheat tissue culture has come from a number of factors. First, and most important, has been selection of explant. There is increasing reliance on immature tissues for callus initiation and regeneration. Immature inflorescence, immature leaf bases, mesocotyls, and immature embryos have proven definitely superior to fully differentiated mature tissues, which give extremely low frequencies of callus induction and regeneration. Second, medium composition is also important. A double-strength MS with casein hydrolysate or coconut water is required. Third, any treatment that suppresses precocious development enhances embryogenesis. Fourth, for organogenesis, especially development of shoots, it is important to suppress ethylene action. Also important are factors such as mass of the inoculum for initiating callus growth and positioning of the explant (embryos especially) on the surface of the nutrient medium at the time of culture. Although we have left out of the discussion any aspect of somatic embryogenesis, it is clear that it is triggered by stress and then withdrawal from that stress, for example, hormonal, desiccation, heat, cold, osmotic, etc. Relieving tissues of any of these leads to a burst of somatic embryogenesis. Finally, for isolation of regenerable protoplasts, fine suspension cells are required and must be obtained from a particular embryogenic callus. Paradoxically, young callus cultures are unsuitable for obtaining protoplasts with high division and plating frequencies. Long-term subculturing to obtain fine cell suspensions is a prerequisite for satisfactory yields of regenerable protoplasts.[262]

From what was an impasse just a couple of years ago, now several laboratories are actively engaged in establishing successful protocols for raising plants from protoplasts. From this progress it appears that it will soon be possible to establish techniques such that protoplast culture will become routine with standard protocols in various laboratories around the world.

IV. GENERAL CONSIDERATIONS

A. Somaclonal and Gametoclonal Variation

Since Larkin and Scowcroft[136] first drew attention to somaclonal variation induced by tissue culture, a great deal of literature on this subject has accumulated on various plant species. Nonetheless, their usefulness for improvement of crops has remained controversial. Depending on whether this phenomenon is useful, it is important to know how to control its occurrence.

Interest in somaclonal variation has stayed alive because of its value in permitting change of one or more characters in an otherwise outstanding cultivar without affecting the rest of the genome. Further, no additional expense is involved in picking up a desirable somaclone. With these considerations its potential to contribute suitable genetic variation for wheat improvement has been explored over the years. As a result, a large number of morphological and biochemical variants have been characterized. These include plant height, presence or absence of awns, storage proteins, amylases, and tolerance to aluminum (see Reference 34).

Plants with early or late flowering have been reported among somaclones of wheat.[137] Lazar et al.[145]

selected somaclonal lines with greater freezing tolerance. Of even greater importance are agronomic traits that enhance yield potential. Several investigations have focused on traits such as kernel weight, heading date, and grain number per spike, and heritable somaclonal variants have been selected with regard to these.[6,34,137,167,244] Variation in spike length and number of grains per spike were reported in somaclone lines compared with the nontissue culture-raised parents.[191,192] However, the assessment of yield of somaclones has so far been difficult because testing in conventional yield plots is hindered by the small number of selfed progenies of regenerated plants.

If tissue culture is regarded as a mutagenic procedure,[224] all forms of mutational events are expected to occur. Changes in ploidy level, chromosome breakage and rearrangement, gene amplification, single-gene mutations, variation in quantitative traits, and activation of transposable elements have all been reported. The work of Karp et al.[127] showed gross changes, usually a decline in chromosome number, in wheat cell suspension cultures and protoplasts. These led to cessation of cell division and growth. Chromosome instability is dependent on ploidy level; diploids are the most stable, while hexaploids are least stable.[290] However, less drastic changes may be sustained and perpetuated and may in fact be useful. Indeed, in wheat, genetic changes resembling single-gene mutations have been detected, for example, variants in height (see Reference 137). A grass dwarf mutation induced by wheat callus culture is also a single partially dominant gene characterized as a new allele at one of the two dwarfing loci, D_1 and D_4.[92]

Detection of somaclonal variability in protein level has also been reported.[48] Gliadins are considered to be excellent biochemical markers for assessing genetic constitution of wheat plants. Significant heritable variation in gliadin patterns was noted in regenerated plants.[48,137] Changes in isozyme patterns represent a widely employed index of variation. Compared with other characters, isozymes can be screened easily and are not subject to environmental variation. Ryan and Scowcroft[243] analyzed B-amylase isozymes and found a variant heterozygous for a pattern that was different from the control and showed five additional isozymes.

Several investigators have attempted to determine changes in nuclear and organellar genomes associated with somaclonal variation. Breiman et al.[25,26] analyzed the spacer regions or rRNA genes at *Nor* loci (nucleolus organizing region) in progenies of regenerated plants. Their method depended on digesting DNA with *Taq1* endonuclease and probing the Southern blots with a plasmid (pTA 71) designed to reveal *Nor* region. Reduction in the number of rRNA spacers was noted in the *in vitro* grown lines compared to the controls. For comparison, genes for two grain proteins — gliadin and glutenin — were also investigated. These showed low or no somaclonal variation.

For a number of years, Henry, De Buyser, and colleagues in France investigated mitochondrial genome variability in tissue cultures of wheat[95,241,242] in order to assess its contribution to somaclonal variation observed in tissue cultures. They concluded that the longer the duration of the *in vitro* steps preceding regeneration, the more the variability in mitochondrial DNA. Embryogenic and nonembryogenic cultures differed in the organization of the mitochondrial genome. Further, plants regenerated from short-term cultures showed mitochondrial genomes more like the parental cultivar, while regenerants from long-term cultures showed little resemblance to the parental cultivar and instead resembled the nonembryogenic cultures.[96] Regeneration capacity was correlated with a particular genome organization of mitochondria. The loss of certain fragments of the hypervariable (noncoding) region of mitochondria was associated with the loss of regenerative capacity. The genome organization was genotype dependent, and in one variety where deliberate comparisons of organs and tissues was made, the mitochondrial genome reorganization was clearly organ or tissue specific. The conclusions were based largely on a study of a hypervariable region of the mitochondrial genome.[195,196]

This brings us to the question of how and when in the tissue culture process the variations are induced and whether it is possible to control their occurrence and frequency. In a recent study partitioning total somaclonal variation among three potential sources — single-embryo-derived calli, regenerants from these, and spike-derived lines from a common regenerant — it was concluded that somaclonal variations originated during the callus phase, while the regenerated embryos did not consistently provide a significant source of variation.[34] By manipulation of medium, somaclonal variation can be reduced as is indicated by evaluation of the effects of three different media on genetic stability of calli.[97] Growth regulators like 2,4-D induce sister chromatid exchanges in cultured wheat cells, while NAA does not.[201]

It has been suggested that a molecular basis of somaclonal variation may be methylation of DNA.[225] In maize, the Ac element showed progressive demethylation with time spent in culture.[60] Again in maize tissue cultures, transposition of the mutator

element occurs when it is unmethylated and genetically active, while the methylated or genetically inactive Mu does not transpose. Transposition of active Mu elements brings about additional genetic and cytological rearrangements.[60] In carrot tissue cultures, auxins dramatically increase methylation of DNA.[160]

If stable and sexually transmitted, somaclonal variants could be useful. However, examples of instability, although not reported so far in wheat, are a possibility. In rice, a dwarf regenerant was lost after out-crossing.[208] In sugar cane, eye spot disease resistance obtained via somaclonal variation was lost after 10 years of asexual reproduction.[173]

In conclusion, somaclonal variation is not looked upon as favorably as it was some time ago despite its possible advantages. Indeed, many workers[34,244] recommend selection of genotypes that minimize or eliminate somaclonal variation. In an investigation of grass dwarf mutation that leads to reduced yields, Guenzi et al.[92] also recommend working with genotypes that do not yield variants, at least of this kind. Somaclonal variation is also not desirable if the goal is to regenerate plants from transformation products or from *in vitro* selection procedures. However, the real challenge lies in gaining an understanding of the underlying processes that cause somaclonal variation so that one can exercise more control over it.

B. Genotype Effect

It is now generally accepted that various *in vitro* responses are genotype dependent. These include callus induction, somatic embryogenesis, the occurrence of somaclonal variation, and regeneration. Strong genotypic effects in wheat were observed in callus induction from immature embryos and regeneration.[5,32,182,183,220,252,281] In a detailed examination of 25 cultivars, embryoid formation and shoot regeneration were found to vary from as low as 12% to as high as 96%.[166] In another survey of 30 cultivars, a genotypic effect was recorded with respect to callus induction and growth in response to different hormones and light intensity, as well as shoot regeneration from immature and mature embryos.[141] Percent callusing and the ability to regenerate varied greatly according to genotype of 20 cultivars studied.[234] Production of plants from embryogenic and nonembryogenic calli[114] and callus induction from the epiblast[102,104] were clearly genotype dependent. Even the frequency of precocious germination may be genotype dependent. In the selection of "Bobwhite" for achieving rapid transformation,[286] the importance of genotype was evident.

C. Genetic Basis of Tissue Culture Response

Much effort has been expended in relating the various tissue culture responses to genes or to specific chromosomes and regions within chromosomes. Through conventional breeding and selection strategies, it is now possible to design genotypes with enhanced tissue culture response characteristics. It is not entirely clear whether the chromosomes governing the tissue culture response of anthers and somatic tissues are the same or different. According to Agache,[2] genetic control of the two is quite separable and is not determined by the same genes. It should be possible then to select genotypes that combine alleles for good response to both systems. Wheat chromosomes shown to be important in the tissue culture response of somatic tissues are 4B, 3A, 2D, and 2BL. Early work of Shimada and Makino[255] and Baroncelli et al.[17] with ditelosomic lines established that deletions of the whole arm of chromosome 4B were deleterious for callus initiation and growth.

Work of Mathias et al.[182] is supportive of the role of chromosome 4B in the tissue culture response. The cultivars Chinese Spring (CS) and Cappelle-Desprez (Cap) have nearly identical responses in callus cultures, but the substitution of chromosome 4B of Cap into CS resulted in the enhancement of the growth rate and regeneration potential of immature embryo-derived calli. Normally, regeneration is suppressed by high 2,4-D concentrations, but in the substitution lines normal shoot regeneration occurred. This work showed that the 4B substitution changed the sensitivity to 2,4-D concentration. Extending this study, Higgins and Mathias[115] found that the deletion of the short arm of 4B, as in the CS ditelocentric lines, resulted in marked reduction of shoot regeneration. The variation among different substitution lines was not altered by light or dark conditions, which normally affect shoot formation. These results and also the substitution (CS for Cap 4B) strongly suggest that the gene(s) on chromosome 4B are indeed responsible for the tissue culture response. Similar results were obtained by Felsenburg et al.[78] employing aneuploid lines of CS.

In a subsequent study with CS designed to identify the specific chromosomal arm responsible for tissue culture responses, Kaleikau et al.[125] identified 2BL as the major regulatory segment controlling the response. Crosses involving a monosomic series showed high-frequency regeneration (93.6%) and high callus growth rate (1.87 g in 90 d) if 2D chromosomes were present.[124] An interesting report[180] linked a known genetic locus with modified *in vitro* response in imma-

ture embryos. Chromosomes 4A and 4D carried the reduced height (*Rht/GA*i) loci and the D3/dwarf clump locus is on chromosome 4B. The dwarfing genes (*Rht* 1, *Rht* 2, and *Rht* 3) alter gibberellic acid and auxin metabolism and have significant effects on the growth and morphogenesis of calli. The *Rht* effect in culture supports the earlier work on 4B chromosomes mentioned above[115] and emphasizes the strong interaction between *Rht* genes, 2,4-D concentration in the medium, and the genotype of the plant.

More recently, Langridge et al.[134] found that the substitution of the short arm of rye chromosome 1 (1RS) into wheat results in stimulation of callus growth and embryogenesis from immature embryos cultures. Two distinct regions on 1RS could be identified that are involved in this response. The "enhancement of embryogenesis" is located between the centromere and the Sec1 locus while the "stimulation of callus growth" is situated close to Sec1 locus.

That there is a genetic basis for tissue culture responses is established beyond doubt. However, there is no consensus on the precise chromosomes and genes involved in the various responses. These questions can be further resolved only by more detailed gene mapping studies.

D. Long-Term Maintenance of Regeneration Potential

A concern, of general interest, is the long-term retention of regeneration potential of tissues in culture. That the regeneration frequency declines with increased culture period has been known for a long time.[202] Yet, only a few studies are addressed specifically to this problem. Rajyalakshmi et al.[234] could obtain complete plant regeneration even after keeping the cultures for 18 months. In 12-month-old cultures, the morphogenic potential gradually diminished, but subculturing at short intervals was helpful in restoring it. In a recent publication, Vasil et al. reported successful maintenance of inflorescence-derived "aged callus" for a period of 19 months.[236]

V. ANTHER AND MICROSPORE CULTURE

A. History and Significance of Study of Factors Controlling Androgenesis

It is well known that production of haploids by anther culture, discovered by Guha and Maheshwari,[88,89] not only dramatically reduces the time frame for generating pure lines but also offers several other advantages. These include screening of smaller populations for selection of desired genotypes, uncovering of recessive alleles, and efficient selection of mutants after mutagenic treatment. In recent years, with rapid developments in this field and application of the technique to many crops, such as *Brassica* among dicots and rice among monocots, there has been much interest in the use of microspore-derived haploid lines also in wheat. Indeed, considerable progress has been made in obtaining haploid lines in this crop and new varieties have been released by Chinese and French workers. These are discussed below.

Early progress in obtaining calli from anther cultures came from Fujii in Japan[81] and Gadgil and coworkers in India.[93] Haploid plants were obtained soon after in 1972 and 1973 in China by the groups of Ouyang and Chu[8,45,212,277] and in France by Picard and De Buyser.[225] By the end of the 1970s, work on haploids was also being conducted in Canada,[49] Japan,[255] and the U.S.A.[249] In the 1980s, the first field trials of anther culture-derived wheat plants were made by the Chinese and French workers.[117,109,230] The Chinese were the first to release varieties in the field. At least 20 varieties have been developed. Jinghua-1 (a winter wheat) is an outstanding example; it is reported to have outyielded existing varieties in every location tested in China.[117] A new cultivar, Florin (also a winter wheat), was developed and released in France. It has proven to be superior to the existing French varieties.[109,230] Considerable work has also been done by Schmid and coworkers[250] in Switzerland with *T. speltoides*, the "spelt" wheat, which is valuable for southern Europe. Advanced programs are also underway on German,[77,80] Hungarian,[16,21] American,[133,146] Canadian,[128] and Australian wheats.[163]

The rapid advances that have taken place with respect to anther culture of wheat now parallel the remarkable successes already achieved in barley and *Brassica*. This has come as a result of continuous improvement in culture conditions and painstaking studies on the effect of a variety of factors on anther response. These can be classified broadly as follows: (1) state of the donor plant, (2) pretreatments prior to anther culture, (3) various physical and chemical factors during culture and induction of androgenesis, (4) culture conditions for regeneration, and, finally (5) genotype effects. These are discussed in the following pages in the order indicated.

B. Anther Donor Plant and Stage of Anther

Health of the anther donor plant should be the first concern of any investigator for anther culture response.

It has been found that the response of spring-sown donor plants is twice as good as that of autumn-sown plants with respect to callus induction and regeneration, even in winter varieties.[122] Comparisons have also been made of the suitability of plants grown in greenhouses or growth chambers with those grown in the field. It is desirable to have experimental material all year round and it may be expected that conditions in modern growth chambers can be controlled sufficiently to give a response as good as that in the field. However, in actual practice this may not be true. Chinese workers have relied extensively on field-grown material and greenhouse-grown plants were found to be definitely inferior.[213,222] Similarly, in a recent study in the U.S.A., Lu et al.[161] found that anthers from field-grown plants callused more than those from a greenhouse or even controlled environmental chambers. However, according to Bjornstad et al.,[23] who conducted their study in Norway with a Scandinavian spring wheat, field-grown plants responded more poorly than those grown in growth chambers. Lazar et al.,[144,146] who have done extensive studies in the U.S.A., also found growth chambers to be suitable for growing donor plants.

Anther response is also dependent on the stage of the microspores at the time of excision of the spikes or the anthers. Detailed studies have shown that the mid- or late-uninucleate stages are the best for anther response.[101]

C. Pretreatments

A well-known practice in anther culture of dicots is cold pretreatment. Even among monocots, for example, barley, the accepted pretreatment is storage of anthers or preferably spikes at 4°C for about 28 d in darkness at high relative humidity.[163] In wheat, Picard and co-workers[228] subjected experimental material to cold treatment, typically at 3°C to 5°C for 1 week, although the duration of treatment varied from a minimum of 2 d to 3 to 4 weeks. Further, as in barley, such treatment is given more conveniently to spikes and sometimes to entire tillers. Several other laboratories, for example, in China[117,215,284] and in the U.S.A.,[146,249] have also routinely employed cold treatment for enhancing the response of anther cultures. Eapen and Rao[70] and Chu and Hill[43] also utilized cold treatment for improved haploid production. Among various reports, the work of Datta and Wenzel[53] deserves special mention. Their data showed that cold treatment for 6 to 8 d was highly beneficial and stimulated divisions from isolated microspores approximately threefold and embryoid formation 4- to 12-fold.

Entire donor plants can also be subjected to low-temperature treatment. Maintaining donor plants at 10°C during the day and at 5°C during the night markedly enhanced anther culture response and production of green plants in all Canadian winter wheat cultivars tested.[260]

Thus, the work in a large number of laboratories shows that cold treatment improves anther culture response. However, in some studies pretreatment was found to be unnecessary or even inhibitory.[150,154,156,162,175] The reason for this difference in results is not known. One possibility could be differences in technique. Some workers emphasize that the material should be kept humid at all times and that desiccation must be prevented. Another reason could be that the response is genotype dependent. Differences could also arise on account of donor plant environment. Indeed, Henry and De Buyser[109] found that cold treatment is not very effective for anthers from field-grown plants.

In addition to cold treatment, subjecting anthers to high osmotic pressure and to γ-rays has also been recommended. Wang et al.[279] found the treatment of excised wheat anthers prior to inoculation with 0.8 *M* sucrose, or a 0.8 *M* mixture of KCl + $MgSO_4$ for 1 h significantly increased the response. Ling et al.[156] recently reported that ^{60}Co γ-rays markedly stimulated embryoid formation. At 3 Gy, the percentage increased form 18.2 to 66 in cultivar Grebe and in the cultivar Kite it increased from 0 to 4.4 at 1 Gy. Chinese workers have reported similar effects.[276,295]

D. Medium Composition

1. Major Salts

The first culture media employed for wheat anther culture leading to production of haploid plants were MS[8,45,212] and Miller's.[225] MS was employed by Chinese workers and the other medium by the French. However, there have since been remarkable improvements in culture medium. These include, for example, modification in concentration of various salts. Yang[293] reduced four of the major salts of MS medium to half the original strength and increased the phosphate concentration. This was termed Maiji-I medium. It gave superior response over MS medium. The same year (1975), Chu et al.[46] developed a further improved synthetic medium, the N_6 medium, for rice. This had a reduced nitrogen level, ammonium nitrate was totally deleted, although KNO_3 was increased by about 1.5-fold, but provided a higher level of phosphorus than MS medium. These media also contained a higher level of sucrose (6%). Subsequently, the N_6 medium

has been employed widely; its efficacy can be judged from the fact that it increased the number of pollen calli 11-fold over that in MS medium.[46] However, other media have also come into existence, such as the C_1 (also synthetic) described by Chen and Wang,[36] or C17, further improved by Wang and Chen,[280] and the D medium, by Zhang and Meng.[300] Currently, the modified C17 medium, called MC17, is finding favor with several workers, for example, Luckett et al.[162] Ouyang and co-workers have also designed another series, W_4, W_5, etc. (see Reference 214). In all these media the level of nitrogen, whether in the form of ammonium or nitrate, has been further decreased. In general, the pollen induction frequency reported is significantly higher but the percentage of green plant is lower than in N_6 medium.

The need for more intensive studies on the role of various ions is shown also by the results of Feng and Ouyang[79] on optimal level of KNO_3. Employing W_5 medium, they found that although KNO_3 at levels beyond 20 m*M* decreased plantlet regeneration, the ratio of green plants nevertheless increased up to about 35 m*M*.

2. Hormones

Hormones are required in all the media for the induction of microspore embryos. Usually, two media are used, one for inducing androgenesis and the other for regenerating plants. For the former, inclusion of 2 mg/l 2,4-D has been a standard practice, while in the regeneration medium the auxin concentration is either lowered or omitted. Generally, in the regeneration medium, a weaker auxin such as IAA or NAA has been used to suppress callusing and to promote organogenesis. However, to save time and eliminate the transfer of calli, Pan and Gao[219] and Liang et al.[155] described single-step procedures. In these procedures, 2,4-D is replaced by either IAA or NAA in the initial medium. However, this sometimes results in lower rates of regeneration and fewer plantlets. Therefore, most workers have continued to employ the two-step procedure (as mentioned later, the optimal sugar concentration also varies in induction and regeneration media). For further development of the plantlets, no hormones in the medium are necessary and some workers, for example, Chu and Hill,[43] have even employed a third medium, lacking any hormone but containing some sugar, before transferring plants to soil, for facilitating rooting and stabilization in the new environment.

Some workers have found a new hormone, PAA, distinctly superior to 2,4-D in their experiments.[163,305] Ziauddin et al.[305] reported that nearly 150 green plants per 100 anthers were obtained with PAA when compared with only 30 to 60 green plants per 100 anthers in 2,4-D medium. There is evidence that PAA may be a natural hormone.[163,305] Also, unlike IAA, which is unstable in light and is destroyed in presence of basal salts, PAA is stable in their presence, withstands autoclaving, and is not degraded during storage.

3. Potato Extract

Another major contribution of Chinese workers in anther culture of cereals was the discovery of the high efficacy of potato extracts. This led to the development of the potato media (P-1, P-2, etc.). Originally, the potato medium with only 20% strength of the extract of fresh potatoes was described by Research Group 301 in 1976;[9] the major and minor salts were all supplied by the extract. Excellent results were obtained with tobacco, then rice,[38] and later also with wheat as reviewed by Ouyang.[214] To reduce variability due to different potato batches, the strength of the extract was later lowered to 10% (Potato-2 medium) and salts of a synthetic medium, namely, W_4, were incorporated at half-strength[47] (see also Reference 214). Although many workers prefer to use a purely synthetic medium such as N_6, results of most studies showed that the yield of calli was always at least twice as high on potato than on synthetic media. Use of potato extracts has thus resulted in a major breakthrough in anther culture of wheat.[186] Continued improvements have resulted in new media, such as P_3, P_4, etc.[12,302] Of interest is the recent work of Marburger et al.[172] who showed that peeled potatoes produce superior results. The authors point out that acidic phenols, glycoalkaloids, chlorogenic acid, and ABA are present in unpeeled potatoes, which may be inhibitory for callusing of pollen.

4. Sugars

Improvement in response has also resulted from altering the concentration and type of sugars employed in the media. The inclusion of 6% sucrose in N_6 and potato media (instead of the 2 to 3% normally in MS medium) produced significantly improved anther culture response. Currently, 9% sucrose is recommended in induction media (see References 12 and 212). Yet it is to be noted that optimal concentration varies not only with the genotype but also whether the medium in question is for induction or regeneration[12] because in regeneration media, sugar concentration has to be

low (2 or 3%). However, another very significant point that has emerged from recent studies is that other sugars are much better provided alone or in combination with sucrose, for example, glucose (Reference 130, cited in Reference 149; Reference 44) and more particularly maltose.[44,76,210,302] Zhou et al.[302] employed 13.5% maltose and obtained a significant increase in the number of green plant regenerants over albinos. Other workers have reported increases in overall embryo yields of cultured anthers by as much as three- to fourfold by use of maltose instead of sucrose (e.g., Reference 139). In a recent investigation on isolated pollen culture Mejza et al.[188] found that inclusion of maltose was critical for success.

It is not known with certainty why higher concentrations of sugars increase the response. The availability of a larger quantity of a carbon source need not be the sole or even the principal factor. In fact, it appears from a number of studies, for example, by Ball et al.,[12] that the osmotic stress is critical for not only callus induction, but also for plant regeneration and for obtaining higher numbers of green plant regenerants, and possibly this is what may cause the higher response. Of particular interest is the recent work of Zhou et al.[302,303] They studied the relationship between anther response and osmotic pressure by including 200 g/l Ficoll (in liquid medium). This resulted in a much higher ratio of green plants over albinos. The overall yield of plants was reduced, but there was a higher number of green regenerants.

E. Gelling Agents and Liquid vs. Solid Media

Although agar is a common constituent of tissue culture media, it has been known to have certain components inhibitory for cell culture. Henry et al.[107] found that agarose gave much better response than did bacto-agar and many workers have subsequently preferred it to agar. A major disadvantage, however, is that like Ficoll, it is much more expensive. Recently, a study was conducted to determine whether other agents that increase viscosity and bring about gelling, such as starch, could be employed.[259] It was found that corn and wheat starch could not only replace bacto-agar, but wheat starch supported high frequencies of embryo regeneration. In fact, three to four times as many embryos were produced.

Comparative studies have shown that liquid culture may be superior to solid medium. The first study was reported by Henry and De Buyser[106] who used the term *float culture*. Despite this report, liquid culture has received limited use because of the need for repeated transfer of explants, which is tedious, and the increased possibility of contamination. Yet, in many other laboratories liquid cultures have been used and whenever comparisons have been made it has generally been shown that they are superior to solid cultures.[43,76,123,128,133,144,149,284,306] Chu and Hill[43] reported that induction frequency of embryoids increased twofold in liquid media and Konzak et al.[132] recorded an almost 20-fold increase in callusing in liquid medium.

For liquid cultures, it is necessary that a very limited amount of medium be used in the culture vessel so that the liquid layer is shallow. Alternatively, Ficoll (40%) is used to increase viscosity of the medium and enable anthers to float. A particularly interesting technology development is the membrane raft[162] that produced a high frequency of green plant regeneration. A two-step procedure has been used and membrane rafts, which are now commercially available, greatly simplify transfer of explants from one medium to another.

F. Miscellaneous — Temperature/Light during Incubation and Pretreatment by Gametocidal Agents

Activated charcoal has been reported to increase embryoid production in agar-gelled medium, although plant regeneration frequency was decreased.[123] Other workers, however, have failed to find any significant effects with charcoal.[146,154]

It has also been found by several workers that a higher temperature for culture may be of great advantage. Normally, tissue cultures are maintained at 25°C, but 28 to 30°C may be optimal.[163] Even higher temperatures such as 30 to 35°C may be employed, but for shorter periods, such as 5 d.[118]

Light/dark regimes also play an important role in the anther culture response. One of the more detailed studies is that of Bjornstad et al.[23] on two Scandinavian cultivars. They found that darkness or low light intensity promoted induction, but high light intensity (75 μM Einsteins m^{-2}-s^{-1}) strongly inhibited it. Light can also have an independent effect on frequency of green plants. In an investigation with German spring cultivars, Ziegler et al.[306] found that regeneration of green plants was improved by differentiation in the dark (25% of the plants were green), while under light only 2.9% of the plants were green.

Experiments have also been conducted with chemicals that induce male sterility, the so-called gametocides or chemical hybridization agents[228,251] and male sterile mutant lines.[110] Schmid and Keller[251] conducted their experiment with *T. speltoides*. However, Picard et al.[228,230] describe results with *T. aestivum* in which the

induction response was greatly stimulated (about 20 times over the control in the Fielder Chris hybrid) by spraying donor plants just before or at the time of meiosis with a chemical hybridizing agent Fenridazone-potassium (RH0007, manufactured by Rhom and Haas, U.S.A.). It is not known how precisely these chemicals work. However, pollen embryogenesis has been considered to be a consequence of an interference of the normal pathway, that is, toward development of a pollen tube and release of sperm cells, and it has been hypothesized that any substance that disturbs the sex balance, or induces sterility or feminization, leads to androgenesis.[110] It has also been observed that the androgenetic response is much greater in cytoplasmically induced male sterile lines. Certain types of cytoplasms (e.g., *T. timopheevi*) seem especially conducive in this regard.[110] To conclude, more research on these lines is needed.

G. Regeneration from Isolated Microspores

This section is devoted to the new emerging technique of isolated microspore culture. This may gradually replace anther culture for raising haploid plants and offers distinct advantages in the selection of valuable androgenetic lines (see References 54 and 163).

Two techniques have been employed for isolating microspores. In the early studies, pollen grains were shed from anthers in liquid medium. Several laboratories have adopted the procedure of serial culture in fresh medium. This allows the senescing anthers to be left behind so that substances leached from them do not adversely affect developing embryoids. Although these manipulations entail some effort and also consume time, recent innovations like the membrane raft technology can simplify manipulations during culture and ensure higher success rates in regeneration from microspores. The other technique, which has been used less frequently but may become a method of choice, is mechanical isolation. This involves grinding the anthers and separating pollen from anther debris by sieving or centrifuging. With gradient centrifugation, pollen of different sizes or stages can be collected.

The culture of microspores was first reported by Henry and De Buyser[106] and Wei.[284] In the former work, anthers were precultured on solid potato medium for 12 to 14 d. It was possible to show a limited extent of embryo development (1 to 2 mm) if a pollen suspension was made and then transferred to R3 regeneration medium containing potato extract, glutamine, and zeatin (modified after Chuang et al.[47]) Also, in the Chinese work,[284] pollen was allowed to be shed but anthers were cultured in liquid medium. When collected by centrifugation and resuspended in fresh medium, the pollen grains grew into calli. The induction frequency was about sixfold greater than in agar medium anther culture on a per anther basis. However, pollen calli had less regeneration ability than calli derived from anther cultures.

Datta and Wenzel[53] obtained embryoids directly from microspores without a callus phase. As in other work, anthers cultured in liquid medium shed microspores within 4 to 6 d. The microspores were then cultured in N_6 medium and these formed multicellular embryos after 2 weeks; plantlets formed later. Addition of potato extract was not necessary but the inclusion of Ficoll (10%) along with a high concentration of 2,4-D (22.5 μM) was critical for induction of embryogenesis.

More recently, Mejza et al.[188] succeeded in culturing isolated microspores from cultivar Pavon, without prior anther culture, and could regenerate fertile plants. In contrast to Datta and Wenzel,[53] who had only limited success with microspores isolated mechanically, these workers obtained excellent results from microspores isolated by microblending. However, incorporation of maltose (this sugar replaced sucrose also for gradient centrifugation) and coculture with live wheat or barley ovaries was critical. Immersion of cut ends of spikes in water for 2 d prior to culture of microspores was also necessary to reduce the variability of response. The contribution is significant because with selected bands on density gradients as many as 7% of the microspores formed embryos. It is expected that this technique will be further perfected and may ultimately replace anther culture.

H. Genetic Basis of Anther Culture Response

Earlier we discussed the genetic basis of tissue culture response of somatic cells. Many investigators have likewise been interested in determining the genetic basis of anther culture response.

There are numerous reports of a strong genotype dependence of anther culture response. For example, in Hungary, Heszky and Mesch[112] obtained a response from only 11 of 66 wheat cultivars. In England, Bajaj[10] obtained a response from 10 of 21 cultivars and in a U.S. study, Schaeffer et al.[249] obtained callus from 8 of 17 cultivars tested and only 3 formed plantlets.

These investigations clearly indicate that genotype is important in mediating androgenetic responses. Additional evidence comes from the work of Picard and De Buyser.[227] These workers showed an increase in androgenetic capacity from doubled haploids com-

pared with the parent cultivars. In fact, these results led to a greatly heightened interest in investigations of this type. If a genetic basis can be firmly established, perhaps the genes controlling this trait can be transferred to nonresponding varieties that possess other desirable agronomic traits.

Apart from the pioneering work of Picard, De Buyser, and their co-workers in France, such work has also been conducted by Schaeffer and co-workers in the U.S.A. and Zhang and Li in China. In recent years, important contributions have also come from the laboratories of Barnabás in Hungary, Konzak in the U.S.A., and Muller in Germany.

In principle, anther culture response can involve both nuclear and cytoplasmic genes. Firm evidence for the involvement of nuclear genes came in 1982 from Bullock et al.[29] who studied the F_1 and their reciprocal crosses from a few important American wheat varieties. Because no significant differences were found in the reciprocal crosses, they excluded a cytoplasmic component. More detailed studies were undertaken later by the same group.[142-144] In a series of contributions, they showed that anther culture response was the end result of three independently governed traits concerned with (1) callus induction frequency, (2) regeneration frequency, and (3) frequency of green vs. albino plants. That the traits are inherited independently was evident from observations that the highest callusing frequency was exhibited by one cultivar, high regeneration frequency by another, and a high ratio of green to albino plants in yet another cultivar. These conclusions were supported by later studies of De Buyser and co-workers in France[2,3,108] as well as by Sági and Barnabás[246] in Hungary, Ekiz and Konzak[71-73] in the U.S.A., and Andersen and co-workers in Denmark.[267]

Beginning with the early work of Shimada and Makino,[255] monosomics, special substitution and translocation lines, disomics, nullisomics, and ditelocentric chromosomes have been used to determine not only which chromosome(s) may be involved in the AA BB DD genomes in the various responses, but also where such genes may be located, that is, whether they are on the short or long arms.[3,55,246,263,301] Results of these studies reveal that the genes are located mainly on chromosome 1 of the B genome, chromosome 2 of the A and D genomes, chromosome 3 of the A genome, chromosome 5 of the A and B genomes, and chromosome 7 of all three genomes. In most cases, both chromosomal arms promote production of plants from microspores but 1BS and 1BL reduce it.[55] 1BL also increases albino frequency.

A particularly interesting case concerns the 1BL/1RS translocation in cultivars such as Aurora, which contain a portion of the rye genome. In this case, traits such as the frequency of embryoid formation, plant regeneration, and formation of green plants are all greatly stimulated.[108] Also, the 1B chromosome, by itself, stimulates the microspore response. These observations on the significance of translocation of genetic material are also supported by recent studies of Muller et al.,[198,199] Lashermes et al.,[138] and Devaux,[61] who have used a number of cultivars originating from both Russia and the U.S.A.

It is expected that in the next few years rapid advances will be made in identifying various genes both in wheat and rye through chromosome walking and other modern recombinant DNA methodologies. When this is accomplished, investigators may be able to obtain directed transfer of anther culture ability.

Finally, a brief reference may be made again to the maternal effects. Although in the early work no such effects were found and anther culture response was supposed to be exclusively transferable by nuclear genes, recently Barnabás and co-workers[246] and Ekiz and Konzak[71-73] reported not only differences in reciprocal crosses but that for *all* components of anther culture response, cytoplasmic genes may interact with nuclear genes. At present, it is not clear where these genes are located, for example, in chloroplasts or mitochondria. To determine this, further work is required.

I. Conclusions

To summarize the research on anther culture, there is no doubt that great progress has been made in the 2 decades following the first reports of development of haploid wheat plants *in vitro*. In addition to the release of superior varieties in France and China produced by anther culture, much research has been directed toward improving anther culture response in various wheat varieties, including those that are recalcitrant. This research has followed two approaches: one physiological and the other biochemical. In the physiological approach, extensive studies have been done on the optimal environment of donor plants, on medium composition, on the significance of sugars such as maltose or agents that subject the explant to osmotic or other stress, and use of new hormones like PAA. Considerable advances have also been made on microspore culture. With the genetic approach, that is, using aneuploids and substitution and translocation lines, regions in the genome have been identified that control not only overall anther culture response, but also characteristics such as regeneration and frequency of albinism. When these genes are identified, even recalcitrant species may be made responsive to anther culture.

Significant developments, new in the field, include molecular genetic approaches of analysis of pollen development exemplified by the work of Reynolds and Kitto.[240] By construction of cDNA libraries of normal pollen and pollen undergoing embryogenesis, followed by differential screening, genes are being identified that may regulate pollen development. Eventually, such work may lead to rapid strides in our understanding of the control and manipulation of the pollen development pathway.

VI. STUDIES ON GENETIC ENGINEERING

A. Transient Gene Expression

In wheat, early gene transfer studies were limited to transient gene expression or transgenic cell lines because of the lack of suitable regenerative systems. This limitation has been overcome; not only is the regeneration of complete plants possible from cell suspension-derived protoplasts,[274] but transformation of intact cells has yielded fertile transgenic wheat plants.[273,275,286] Nevertheless, transient assays are still necessary to analyze efficiency of gene delivery systems and as a rapid method to study gene expression. The level of gene expression in a transient assay is free of position effects and is higher than in stable transformations. This may be because the number of copies of the gene in question that do get integrated in the genome is restricted.

Diverse methods have been used to deliver foreign DNA into protoplasts. The earlier methods used polyethylene glycol.[148,158,289] Later, electroporation[98,147,211] and a combination of the two were tried.[203] More recently, other novel methods have been employed that can make use of intact cells.[189] Whatever the method, the choice of markers and reporters or promoters and introns greatly influences the final outcome and these are briefly considered here.

1. Marker and Reporter Genes

In order to see the expression of foreign genes the selectable markers commonly employed are *nptII*, which confers resistance to kanamycin, or *hpt*, which confers resistance to hygromycin.[157] A reporter gene of choice has been *cat*, which codes for the enzyme chloramphenicol acetyltransferase. More recently, the *bar* gene, which codes for the enzyme phosphinothricin acetyltransferase (PAT), has found favor. This gene detoxifies phosphinothricin and bialaphos and allows better regeneration than other markers (see Reference 197).

Because of the importance of antibiotics in the selection of transgenic lines, Simmonds and Grainger[262] conducted a special study to assess the phytotoxic response of a range of antibiotics. Toxicity was evaluated with regard to cell division and plating frequencies in agarose-embedded protoplast cultures initiated at low plating densities (5.0×10^4 protoplasts per milliliter). The aminoglycoside antibiotics hygromycin, G418, bekanamycin, and kanamycin were extremely toxic to wheat cells (50% inhibition at 15 μg/ml). This is in contrast to an earlier finding of Hauptmann et al.[99] who reported resistance to very high levels (800 μg/ml) of kanamycin and hygromycin. The discrepancy may be due to the high cell density cultures used by Hauptmann et al.,[99] which may have led to "escapes". Simmonds and Grainger[262] emphasize the need for relatively low plating densities for measuring the response reproducibly. They further reported resistance of suspension cells to high levels of cefotaxime (100 μg/ml), vancomycin (3000 μg/ml), and spectinomycin (3000 μg/ml). These concentrations are much above the antibacterial activities. Indeed, vancomycin (1000 to 2000 μg/ml) enhanced cell division and plating efficiencies twofold, indicating its usefulness in establishing protoplast cultures.

Among new markers, anthocyanin is very promising because it obviates the need for artificial substrates; it can be visualized easily and allows cell-autonomous expression. Unlike GUS, which can give false positives because it shows expression even in damaged cells, anthocyanin accumulates in intact vacuoles and depends on coordinate expression of many genes. Expression of anthocyanin genes in transformed tissue can provide a system for investigating the regulation of gene expression, which in many ways is difficult with GUS. In addition, anthocyanin markers permit visualization of transgenic tissue from the beginning and throughout development without sacrificing the transformed tissues. First used as a novel visible marker for maize transformation (see Reference 164), apparently anthocyanin biosynthesis in various cereals, including wheat, can also be activated by introduction of regulatory genes B or C1 from maize.[291] In preliminary experiments with particle bombardment, regulatory genes introduced under the control of the constitutive promoter CaMV 35S led to the appearance of pigmented cells in 1 to 2 d in wheat.[291] The Monsanto group has shown stable expression of anthocyanin in calli and suspension cells with three selectable markers, glyphosate, bialaphos, and methotrexate, using microprojectile bombardment.[62] The same group

also produced transformed wheat shoots expressing anthocyanin genes by bombardment of young embryogenic callus with plasmid constructs carrying selectable marker and anthocyanin genes. Under selection, embryogenic callus carrying selectable marker and anthocyanin gene produced shoots. Phenotypes of the leaves expressing anthocyanin varied from solid red to striped.[63]

Klötti et al.[131] also made use of the anthocyanin marker to optimize conditions for gene delivery by electroporation in zygotic embryos. Anthocyanin accumulation could be detected in vacuoles within hours of electroporating the regulatory genes of anthocyanin biosynthesis. The scutellar cells accumulated more anthocyanin than other cells of the embryo, which, according to the authors, is due to tissue specificity of electroporation.

2. Promoters

Efficient expression of a gene requires selection of an appropriate promoter as well as polyadenylation sequences. It is evident that in wheat, heterologous promoters such as those of shrunken gene of maize,[289] copia LTR from *Drosophila*,[211] CaMV 35S,[98,147,211] and nopaline synthase 2′ gene[98] are able to express the introduced genes. However, because the level of expression of introduced genes is quite low in cereals [illegible] and four repeats of the *ocs* enhancer elements from the octopine synthase gene. These are placed in front of a modified *Adh1* promoter with the *Adh1* intron 1. The promoter showed greatly improved GUS expression in five different monocot species, including *T. monococcum*.

The need for exploring new promoters was prompted because those generally in use, such as CaMV 35S, were not active in all cell types of monocots. Interestingly, a duplicated CaMV 35S (enhanced 35S) is an effective promoter.[298] Used alone or with TMV leader sequence, it promotes higher levels of GUS transient activity than the *Adh1* promoter and intron.[298]

Based on the premise that actin is a universal component of the plant cell cytoskeleton, McElroy et al. (see Reference 185) used the 5′ region of the rice actin 1 gene (*Act1*) in a chimeric construct with GUS and found high levels of GUS expression in transient assays of transformed rice and maize cells. Subsequently, the actin promoter was also tested in wheat. In one study, the strength of various promoters was compared.[223] The efficiency of GUS expression from four different promoters was monitored in the calli of wheat scutella: the 35S CaMV promoter, with the addition of the Ω translation enhancer element, the full-length rice actin promoter, *pAct1*-D, and a truncated rice actin promoter *pAct1*-F. The 35S CaMV promoter with or without the enhancer element was barely active, while the actin promoters were very efficient in wheat scutellar tissue. Another new promoter that appears to be effective in transient assays is the ubiquitin promoter plus its first intron. Weeks et al.[285,286] reported that the ubiquitin promoter conferred high levels of expression of the marker genes *bar* and *gus*. In a detailed study, Taylor et al.[264] showed enhanced GUS expression with a ubiquitin-based plasmid construct, PAHC25, in a number of cereals and grasses, including wheat. The study was conducted with both suspension cells and immature embryos subjected to microprojectile bombardment.

Less attention has been paid to the elements at the 3′ end. Various polyadenylation signal sequences from genes such as *ocs*,[289] *rbcS* and *SV40*,[211] and *nos*[147] have been shown to be functional in wheat.

3. Introns

[illegible] expression of genes severalfold. The role played by introns in eukaryotic gene expression is poorly defined, but perhaps they affect the stability of the transcripts that are not degraded due to their presence. In plants, Callis et al.[30] were the first to show that insertion of a maize intron in the 5′ noncoding region of the chimeric gene resulted in 50- to 100-fold stimulation of expression. The mechanism whereby increased expression is obtained could be acting at the splicing level. A chimeric gene construct containing the maize intron sequences (*Adh1* intron 1) markedly enhanced transient expression of CAT in aleurone protoplasts isolated from wheat caryopses.[148] Even more remarkably, a 30- to 185-fold greater CAT activity was seen by using *Adh* intron 6 placed between the promoter of CaMV 35S and *cat* sequences.[203] Similarly, as mentioned before, the maize ubiquitin promoter plus its first intron was used to express marker genes for *bar* and *gus*.[285]

4. Use of Particle Bombardment and Air Guns

Success in genetically manipulating monocots has been difficult to achieve. This is not because of a lack of adequate delivery methods,[19] but until recently serious problems were encountered in regeneration of plant protoplasts. Alternate approaches that allow use of intact cells or multicellular tissues and organs have thus been devised.

Wang et al.[283] showed transient expression of *gus* and *cat* genes by bombarding DNA into aggregates of suspension cells of a *T. monococcum* cell line. Prior to bombardment, the cell aggregates were evenly spread onto filter paper discs in a Petri dish. Unfortunately, the frequency of gene transfer was not possible to calculate because of the uneven distribution of particles and lack of monolayers of plated cells.

Intact germinating wheat embryos were bombarded with tungsten particles carrying the *gus* gene. Appearance of blue patches was indicative of positive GUS reaction.[157] Transient expression of two marker genes, *cat* and *gus*, was also obtained by microprojectile bombardment of immature zygotic embryos.[40] GUS activity was confirmed both histochemically and fluorometrically, while CAT activity was detected by an ELISA technique using CAT-specific antibodies. Interestingly, the maximum expression of both genes was achieved if the *Adh1* intron 1 was present between the promoter and the coding region of the gene.

Recently, Daniell et al.[52] showed the expression of a foreign gene (*gus*) and the ability of a dicot chloroplast promoter (*psbA*) to function in a monocot chloroplast by biolistic delivery into wheat chloroplasts. They used the nuclear GUS vector pBI121 and a chloroplast expression vector with the *psbA* promoter to show the expression of the reporter gene in green as well as albino leaves and in immature embryo-derived calli. The expression of GUS in albino leaves bombarded with a chloroplast expression vector suggests the presence of functional protein synthesizing machinery even in albino plastids.

Pollen embryos derived from cultured anthers bombarded with plasmid DNA (pBI221)-coated gold particles showed transient expression of *gus*.[258] Normal green plants were regenerated from the pollen embryoids after 1 month at high frequency.[258] Embryogenic wheat callus subjected to microprojectile bombardment with the biolistic technique showed transient expression of two marker genes, *bar* and *gus*.[285]

At present, the biolistic method suffers from low efficiency of DNA delivery as well as reduced regeneration capability of the bombarded tissue. Simple improvements in protocols can, however, lead to severalfold higher expression levels. Perl et al.[223] recommend maintaining scutellar calli in 0.25 *M* mannitol before and after bombardment, the inclusion of silver thiosulfate and calcium nitrate, and the exclusion of $CaCl_2$ and spermidine from the microprojectile mixture. Further, the introduction of a liquid culture phase led to higher efficiency of regeneration of bombarded material.

The problem of tissue damage and accurate targeting of genes was overcome by Potrykus and coworkers by using an innovative modification of the particle bombardment technique.[248] They constructed a novel microprojectile accelerating system that accurately microtargets particles to restricted sites in tissues with minimum tissue damage. Unlike the biolistic method, the DNA is not bound to the microprojectiles. Instead, a mist of micrometer-sized droplets containing both DNA free in solution and microprojectiles are accelerated by a gas stream. The particles make holes in the cell wall as well as the membrane and DNA moves in independent of the projectile. The efficiency of the system was tested on wheat scutellar tissue. A very high density of transiently expressed GUS activity was obtained in undamaged cells.

Oard et al.[204] designed a low-cost air gun, which avoids the use of chemical propellants that release gases and lead to target cell trauma. The device has been tested on various cereals including wheat. Leaf bases and shoot apices of wheat bombarded with plasmid constructs containing the *gus* gene showed low levels of expression in leaf base tissue, while the apical dome showed no expression at all although the tungsten particles penetrated down to the fourth cell layer. Further refinements in the selection of promoters, substrate penetration, etc. should produce better results. An important limitation of the particle gun delivery system is the restriction of gene transfer to the first two or three surface layers of the target organ. A recently developed electric discharge gun (Agracetus Corporation) apparently assures deeper penetration into many cell layers and with much less tissue damage.[42] However, this gun has not yet been used for wheat transformation.

5. Uptake of DNA by Dry Embryos

The simplest delivery method devised so far is the uptake of DNA by dry embryos.[265] Transient expression of the *nptII* gene with various promoters was shown in *T. aestivum* and *T. monococcum*. The expression of *nptII* was shown to be eukaryotic rather than prokaryotic in character. This ruled out bacterial contamination. Further evidence that the expression was

due to plant cells came from the use of a specific replicon, the wheat dwarf virus (WDV), in *T. monococcum* protoplasts. Dimers of the wild-type wheat dwarf virus genome (pWDVS2) were more highly expressed compared with supercoiled monomers (pWDVneo1) and were comparable to the dimers (pWDVneo2). Furthermore, DNA uptake took place only when the embryos were incubated first with DNA solution and then with the buffer; if the order was reversed there was no uptake. It is known from previous studies that the continuity of cell walls required at least 20% water content, while in dry embryos the water content is only 4 to 8%.[209] Hence, the possible mechanism of DNA uptake may be through imbibition, but only during the first few minutes. After that the reconstituted cell wall may disallow further uptake.

In addition to intact seeds used by Töpfer et al.[265] dissected seeds, with exposed apical meristems, have also been the targets for DNA delivery.[39] Using wheat dwarf virus DNA as a marker, these investigators compared the three currently used DNA delivery methods: agroinfection, direct DNA imbibition, and microprojectile bombardment. Agroinfection gave the highest levels of infection (79%) when the dissected seeds were soaked in the *Agrobacterium* inoculum. The level of infection was lowest (3%) with microprojectile bombardment. Imbibition of wheat dwarf virus DNA solution gave 16% infection. In the case of dissected seeds, no infection occurred if they were soaked for 10 min in water prior to imbibition by WdvDNA. This indicates that intact membranes prevent entry of DNA.

B. Cocultivation with Agrobacterium

It is thought that monocots are resistant to *Agrobacterium* as shown by lack of tumor formation. However, recently Deng et al.[59] in China showed formation of tumors with *Agrobacterium* injected onto inflorescence stems of wheat. The size of the tumors was larger if acetosyringone was mixed with the bacteria, although it was not absolutely required for induction. Scanning electron microscopic studies have also provided definitive evidence of attachment of *Agrobacterium* to wheat cells.[86] Recently, Mooney and Goodwin[193] showed the adherence of *A. tumefaciens* to the cells of immature wheat embryos by scanning electron microscopy. Although attachment was not wound dependent, wounding preferentially promoted adherence of the bacteria at the wound site. Like dicots, monocots also respond to acetosyringone. If the incubation medium of wheat seedlings was supplemented with acetosyringone, *vir* gene induction and further T-DNA circularization events were induced in *Agrobacterium*.[268,269] However, monocots apparently synthesize their own signal molecules, which are different from acetosyringone and are capable of inducing *vir* genes of *Agrobacterium*. Messens et al.[190] reported purification of a substance identified as ethyl ferulate from cultured suspension cells of *T. monococcum*. In test experiments, this compound gave significant levels of induction, although no transformation occurred. There may be a block in a step subsequent to *vir* induction that precludes transformation.

Partial enzymatic digestion of immature embryos prior to cocultivation with *A. tumefaciens* was used by Mooney et al.[194] for mobilizing DNA. However, frequencies of transfer were low (data not available) and Southern hybridization patterns indicated that T-DNA in one cell line had undergone extensive rearrangements. The authors suggest that this is indicative of a process of T-DNA transfer and integration in cereals different from that which occurs in dicots.

A method devised by Zaghmout and Trolinder[297] appears to be a shortcut in gene delivery because the tedious and time-consuming procedure of plasmid isolation is not required. A binary vector (pKIWI 105), still housed in *Agrobacterium*, was directly electroporated into embryogenic callus cells of wheat. The expression of GUS activity was 10.1 pmol 4-methylumbelliferone per minute per 10^5 cells in the nonelectroporated callus and 100.7 in the electroporated callus.

C. Agroinfection

Although there is an interaction between *Agrobacterium* and monocotyledons, there is still a problem in the detection of transferred DNA because of the low frequency of transfer. An alternative mode of transfer that promises to enhance the detectibility of inserted genes is C. Agroinfection. This is the insertion of viral DNA into a plant with the help of *Agrobacterium*. The virus that enters the plant cell replicates at a rapid rate and amplifies the transferred DNA. This results in practically all cells becoming infected. Introduced DNA, thus, can be easily detected even at the early stages by probing with the viral nucleic acid or with antisera of coat protein. Disease symptoms appear later, when the plant is older. Although no foreign DNA has so far been actually transferred by this method, there appears to be increased possibility of adoption such a procedure as indicated below.

Woolston et al.[292] showed that cloned wheat dwarf virus (WDV) DNA infected wheat seedlings efficiently through agroinfection. In this study, the clone of 2.75

kb representing the complete WDV genome was ligated between the left and right borders of the T-DNA. Further studies have been undertaken by these workers to find the most suitable inoculation procedure as well as explants for efficient agroinfection.[51] Wounding (stabbing) combined with vacuum infiltration gave a higher response than either procedure alone. Inoculation into the meristematic regions of 1- to 4-d-old *in vitro*-grown seedlings gave a response similar to that obtained in seedlings grown in a glasshouse. The addition of acetosyringone improved the response. Recently, efforts have been made to compare various *Agrobacterium* strains for efficient agroinfection.[174] The nopaline types are better than octopine types and *A. rhizogenes* is better than either.

From these studies has emerged the possibility of developing wheat dwarf viral genome as a plant gene expression vector.[184] The viral capsid protein was replaced by three marker genes: *nptII*, *cat*, and B*gal*. The recombinant genomes expressed the foreign genes in a transient replication and expression system of *T. monococcum* suspension culture-derived protoplasts. It was shown that the wheat dwarf virus genome could be adapted as a gene vector for wheat without the need for *Agrobacterium*.

D. Pollen as Vectors

A different approach made use of pollen as vectors for transmitting genes.[229] While germinating on the stigma, apparently pollen tubes can take up plasmid DNA added to the florets. The plasmids are special constructs with a wheat *ori* region of replication and the *nptII* gene fused to the *nos* promoter. Kanamycin resistance and expression of the *nptII* gene was found among the progeny of seeds derived from DNA treated flowers; however, no molecular evidence was given. Transfer of the maize zein gene has been reported in wheat plants by the same approach.[58] In this work, *nptII* also was used as a selectable marker. DNA of kanamycin-resistant plants was shown to contain both the *nptII* and zein gene specific hybridization signals and tissue-specific expression of α-zein was immunologically detected in the endosperm of selfed progeny.

Macroinjection into floral tillers of *T. monococcum* with *nptII* as the selectable marker was used to transfer the soybean legumin gene.[66] Legumin sequences were detected in DNA of some of the progeny. In a similar investigation, Hess et al.[111] pipetted a suspension of *A. tumefaciens* into spikelets to bring the bacteria and pollen in intimate contact. Apparently, pollen releases flavonol glycosides, which, like acetosyringone, induce expression of the *vir* region of the bacterial plasmid. Mobilization of the plasmid DNA was determined by Southern blot analysis and NPT II assays. Kanamycin resistance was detected in two successive generations and the transformants were normal and fertile. The transformation frequency was 1% or higher. The obvious advantage of this method lies in its simplicity; it eliminates the tedium of tissue culture and reduces associated problems such as somaclonal variation. Furthermore, chimeras may be avoided because the transformation may only involve the egg cell (see, however, Reference 135).

E. Transformation

The first attempt to transform wheat was with *T. monococcum* by Lörz et al.[158] They transferred the *nptII* gene with the *nos* promoter and *ocs* polyadenylation sequences into cell suspension-derived protoplasts. The transfer was facilitated by PEG. Protoplasts could be cultured only to the callus stage and the efficiency of transformation was about 1 in 500,000 if the plating efficiency was 25%. There was no comparable report on *T. aestivum* until 1991 when Vasil et al.[272] obtained stably transformed callus lines by microprojectile bombardment of plated suspension cells; they used DNA containing the *nptII*, *gus*, and *EPSP* synthase genes, either on the same or different plasmids. Later, Vasil et al.[273] reported the first fertile transgenic wheat plant. They employed microprojectile bombardment of type C callus, which is a long-term regenerable embryogenic callus. The selectable marker was phosphinothricin instead of kanamycin because wheat has been found to have natural resistance to kanamycin. The plasmid used, pBARGUS, had dual promoters: (1) the *Adh1* with *Adh1* intron 1 to drive the GUS reporter and (2) CaMV 35S with the *Adh1* intron 1 to drive the *bar* gene. Southern blots confirmed the presence of the *bar* gene in transgenic plants that had been selected by topical application of the herbicide Basta. The transformed plants were unable to self-fertilize, but by out-crossing with wild-type pollen or ova, fertile progeny was raised. The introduced *bar* gene showed Mendelian inheritance through two generations. This demonstration of fertile transformed wheat plants has opened the possibility of more effective application of recombinant DNA technology in wheat improvement programs.

Further improvements in procedure aimed at accelerating the pace and frequency of production of transgenic plants have also come from Weeks et al.[286] These workers made use of a highly embryogenic wheat cultivar, "Bobwhite". Calli were derived from immature embryos just as they began to proliferate 5 d after excision and bombarded with the *bar* gene.

Transgenic plants were recovered with high frequency (1 to 2 per 1000 embryos bombarded) and only 168 d elapsed between bombardment and anthesis of To plants. Vasil et al.[275] have discontinued use of type C callus, which occurs with low frequency and is difficult to identify and maintain. Their recent success has also been from direct bombardment of cultured immature embryos and 1- to 2-month-old embryogenic calli. Flowering transgenic plants were obtained in about 270 d after culture.

In spite of the current excitement of producing fertile transgenic plants by high-velocity particle bombardment, the use of protoplasts as transformation systems has not been totally discarded. Zhou et al.[304] recovered stably transformed wheat calli by electroporation-mediated direct gene transfer into haploid protoplasts that were obtained from suspension cells of anther-derived callus. The *bar* gene was introduced via special plasmid constructs pBARGUS and pBAS.

Zaghmout and Trolinder[298] also regard protoplasts as a good system to demonstrate transient gene expression. They prefer electroporation to microprojectile bombardment and consider the latter to be less efficient. Their efforts have been directed toward optimizing conditions for DNA uptake and expression and include the study of various promoters, comparison of protoplasts originating from slow-growing embryogenic and nonembryogenic callus, combining electropermeability with PEG, etc. Their data indicate that protoplasts derived from slow-growing embryogenic callus show a high level of transient activity. The inclusion of PEG during electroporation increases the response, and the promoter of choice is the enhanced 35S CaMV + TMV leader sequence.

VII. GENERAL CONCLUSIONS AND FUTURE OUTLOOK

Two important advances in the last few years have been the regeneration of plants from protoplasts, and the production of transgenic plants. Although one purpose of developing protoplast systems was transformation, they were not used in the production of the first transgenic wheat plants. These were obtained instead by particle bombardment of embryogenic callus. However, because protoplasts can now be made to regenerate, they can also be used for developing transformation technology. New advancements in technology, such as the helium gun developed by Du Pont,[121] the discharge gun developed by Agracetus,[42] and the novel method developed by Potrykus and co-workers (Sautter et al.[248]) to microtarget genes to a restricted site, promise to accelerate the pace of research. Another approach to target plasmid DNA is being developed by Simmonds et al.[261] By use of micromanipulators, DNA is injected into apical meristems such that it goes precisely into the hypodermal layer that gives rise to the floral meristem. Because these meristem explants can be regenerated to form fertile plants, the method has immense potential to obtain transgenic plants in a nonrandom manner.

In an atmosphere charged with environmental consciousness, the availability of the cre-lox system[116] to eliminate marker genes may aid in the transfer of transgenic plants from the laboratory to the field. Originally reported in bacteriophage P_1, cre-lox is a site-specific recombination system consisting of a site lox P (*lo*cus for crossing over) and a recombinase cre (*cau*ses *re*combination), a protein that carries out the recombination reaction. The lox P is 34 bp with two 13-bp inverted repeats separated by an 8-bp spacer. The two elements interact to affect recombination resulting in excision, inversion, or insertion of DNA, depending on location and orientation of the lox P sites. The system has been shown to function in plants and it can affect directed excision of a transgene from the genome, if it is no longer desirable to have the trait expressed. Antibiotic markers could thus be excised from the transgenic plants to make them environmentally benign. The lox P-cre system could potentially be applied also to direct site-specific integration.

With the burgeoning population and shrinking resources, there is an element of urgency in improving crop plants, especially the major cereals of which wheat is one of the most important. Many laboratories all over the world are engaged in improvement programs, and noticeable progress has been made in all aspects of wheat research during the last decade. In the coming years, attention must be given to isolate and characterize genes controlling traits of agronomic importance such as grain quality, disease resistance, and tolerance to stress. Parallel developments need to be made in design of expression vectors, preferably with monocot promoters and in identification of genetic elements in promoter and nonpromoter sequences that control gene expression. Indeed, a great deal of basic molecular biology has to be done before transformation technology can be fully exploited.

ACKNOWLEDGMENTS

The authors thank Professor I. K. Vasil and Dr. A. K. Tyagi for critical reading of the manuscript and helpful suggestions. It gives them pleasure also to thank Neelam Bedi, and Mahalakshmi for checking the successive drafts and Renu Dhawan and Satyendra

Patwal for expert technical assistance with the manuscript.

An addendum has been submitted for this article and will be run in the next issue.

REFERENCES

1. **Adachi, T. and Katayama, Y.,** Callus formation and shoot differentiation in wheat tissue culture, *Bull. Fac. Agric. Univ. Miyazaki*, 16, 77, 1969.
2. **Agache, S., De Buyser, J., Henry, Y., and Snape, J.,** Studies of the genetic relationship between anther culture and somatic tissue culture abilities in wheat, *Plant Breeding*, 100, 26, 1988.
3. **Agache, S., Bachelier, B., De Buyser, J., Henry, Y., and Snape, J.,** Genetic analysis of anther culture response in wheat using aneuploid, chromosome substitution and translocation lines, *Theor. Appl. Genet.*, 77, 7, 1989.
4. **Ahmed, K. Z. and Sagi, F.,** Culture of and fertile plant regeneration from regenerable embryogenic suspension cell-derived protoplasts of wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 12, 175, 1993.
5. **Ahloowalia, B. S.,** Plant regeneration from callus culture in wheat, *Crop Sci.*, 22, 405, 1982.
6. **Ahloowalia, B. S.,** Somaclones of wheat regenerated from primordial leaf callus, in *Genetic Manipulation in Plant Breeding*, Horn, W., Jensen, C. J., Odenbach, W., and Schieder, O., Eds., Walter de Gruyter, Berlin, 1986, 577.
7. **Ahuja, P. S., Pental, D., and Cocking, E. C.,** Plant regeneration from leaf base callus and cell suspensions of *Triticum aestivum*, *Z. Pflanzenzuecht.*, 89, 139, 1982.
8. **Anon.,** Research Group 301, A preliminary report on the induction of pollen plants from rice and wheat, *Genet. Commun. (Beijing)*, 1, 1, 1972 (in Chinese).
9. **Anon.,** Research Group 301, A sharp increase of the frequency of pollen-plant induction in wheat with potato medium, *Acta Genet. Sin.*, 3, 25, 1976.
10. **Bajaj, Y. P. S.,** *In vitro* induction of haploids in wheat (*Triticum aestivum* L.), *Crop Improv.*, 4, 54, 1977.
11. **Bajaj, Y. P. S.,** Ed., *Wheat*, Biotechnology in Agriculture and Forestry, Vol. 13, Springer-Verlag, Berlin, 1990.
12. **Ball, S. T., Zhou, H., and Konzak, C. F.,** Sucrose concentration and its relationship to anther culture in wheat, *Crop Sci.*, 32, 149, 1992.
13. **Bandyopadhyay, S. and Ghosh, P. D.,** Cell divisions in newly formed cells from leaf mesophyll protoplasts of wheat (*Triticum aestivum* cv. Sonalika), *Curr. Sci.*, 55, 1084, 1986.
14. **Bapat, S. A., Joshi, C. P., and Mascarenhas, A. F.,** Occurrence and frequency of precocious germination of somatic embryos is a genotype-dependent phenomenon in wheat, *Plant Cell Rep.*, 7, 538, 1988.
15. **Bapat, S. A., Rawal, S. K., and Mascarenhas, A. F.,** Isozyme profiles during ontogeny of somatic embryos in wheat (*Triticum aestivum* L.), *Plant Sci.*, 82, 235, 1992.
16. **Barnabás, B., Pfahler, P. L., and Kovács, G.,** Direct effect of colchicine on the microspore embryogenesis to produce dihaploid plants in wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 81, 675, 1991.
17. **Baroncelli, S., Bulatti, M., Bennici, A., Foroughi-Wehr, G., Mix, B., Gaul, H., Tagliasacchi, A. M., Lolero, M., and Giorgi, B.,** Genetic control of *in vitro* and *in vivo* growth in hexaploid wheat. I. Behaviour of ditelocentric lines, *Z. Pflanzenzuecht.*, 80, 109, 1978.
18. **Bartok, T. and Sági, F.,** A new, endosperm-supported callus induction method for wheat (*Triticum aestivum* L.), *Plant Cell Tiss. Organ Cult.*, 22, 37, 1990.
19. **Batty, N. P. and Evans, J. M.,** Biological ballistics — no longer a shot in the dark, *Transgen. Res.*, 1, 107, 1992.
20. **Baum, M., Lagudah, E. S., and Appels, R.,** Wide crosses in cereals, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 43, 117, 1992.
21. **Bedo, Z., Karsai, I., Balla, L., and Barnabás, B.,** Some possibilities of efficient haploid production in wheat, *Proc. 7th Int. Wheat Genetics Symp.* Vol. 2, Institute of Plant Science Research, Cambridge, 1989, 1043.
22. **Bhojwani, S. S. and Hayward, C.,** Some observations and comments on tissue culture of wheat, *Z. Pflanzenphysiol.*, 85, 341, 1977.
23. **Bjornastad, A., Opsahl-Ferstad, H. G., and Aasmo, M.,** Effects of donor plant environment and light during incubation on anther cultures of some spring wheat (*Triticum aestivum* L.) cultivars, *Plant Cell Tiss. Organ Cult.*, 17, 27, 1989.
24. **Borrelli, G. M., Fonzo, N. Di., and Lupotto, E.,** Effect of cefotaxime on callus culture and plant regeneration in durum wheat, *J. Plant Physiol.*, 140, 372, 1992.
25. **Breiman, A., Felsenburg, T., and Galun, E.,** *Nor* loci analysis in progenies of plants regenerated from the scutellar callus of bread-wheat, *Theor. Appl. Genet.*, 73, 827, 1987.
26. **Breiman, A., Felsenburg, T., and Galun, E.,** Is *Nor* region variability in wheat invariably caused by tissue culture?, *Theor. Appl. Genet.*, 77, 809, 1989.
27. **Brown, C., Brooks, F. J., Pearson, D., and Mathias, R. J.,** Control of embryogenesis and organogenesis in immature wheat embryo callus using increased medium osmolarity and abscisic acid, *J. Plant Physiol.*, 133, 727, 1989.
28. **Burström, H.,** Studies on growth and metabolism of roots. V. Cell elongation and dry matter content, *Physiol. Plant.*, 4, 199, 1951.
29. **Bullock, W. P., Baenziger, P. S., Schaeffer, G. W., and Botino, P. J.,** Anther culture of wheat (*Triticum aestivum* L.) F_1's and their reciprocal crosses, *Theor. Appl. Genet.*, 62, 155, 1982.
30. **Callis, J., Fromm, M., and Walbot, V.,** Introns increase gene expression in cultured maize cells, *Genes Dev.*, 1, 183, 1987.
31. **Carman, J. G.,** Improved somatic embryogenesis in wheat by partial simulation of the in-ovulo oxygen, growth-regulator and desiccation environments, *Planta*, 175, 417, 1988.
32. **Carman, J. G., Jefferson, N. E., and Campbell, W. F.,** Induction of embryogenic *Triticum aestivum* L. calli. I. Quantification of genotype and culture medium effects, *Plant Cell Tiss. Organ Cult.*, 10, 101, 1987.
33. **Carman, J. G., Jefferson, N. E., and Campbell, W. F.,** Induction of embryogenic *Triticum aestivum* L. calli. II. Quantification of organic addenda and other culture variable effects, *Plant Cell Tiss. Organ Cult.*, 10, 115, 1987.
34. **Carver, B. F. and Johnson, B. B.,** Partitioning of variation derived from tissue culture of winter wheat, *Theor. Appl. Genet.*, 78, 405, 1989.
35. **Chang, Y. F., Wang, W. C., Warfield, C. Y., Nguyen, H. T., and Wong, J. R.,** Plant regeneration from protoplasts isolated from long-term cell cultures of wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 9, 611, 1991.

36. **Chen, Y. R. and Wang, P.**, A synthetic medium (C_1 medium) for wheat anther cultures, *Hereditas (Beijing)*, 1, 31, 1979 (in Chinese).
37. **Chen, T. H. H., Kartha, K. K., and Gusta, L. V.**, Cryopreservation of wheat suspension culture and regenerable callus, *Plant Cell Tiss. Organ Cult.*, 4, 101, 1985.
38. **Chen, Y.**, Anther and pollen culture of rice, in *Haploids of Higher Plants In Vitro*, Hu, H. and Yang, H., Eds., China Academic Publishers, Beijing, 1986, 3.
39. **Chen, D. F. and Dale, P. J.**, A comparison of methods for delivering DNA to wheat: the application of wheat dwarf virus DNA to seeds with exposed apical meristems, *Transgen. Res.*, 1, 93, 1992.
40. **Chibbar, R. N., Kartha, K. K., Leung, N., Qureshi, J., and Caswell, K.**, Transient expression of marker genes in immature zygotic embryos of spring wheat (*Triticum aestivum*) through microprojectile bombardment, *Genome*, 34, 453, 1991.
41. **Chin, J. C. and Scott, K. J.**, Studies on the formation of roots and shoots in wheat callus cultures, *Ann. Bot.*, 41, 473, 1977.
42. **Christou, P.**, Genetic transformation of crop plants using microprojectile bombardment, *Plant J.*, 2, 275, 1992.
43. **Chu, C. C. and Hill, R. D.**, An improved anther culture method for obtaining higher frequency of pollen embryoids in *Triticum aestivum* L., *Plant Sci.*, 55, 175, 1988.
44. **Chu, C. C., Hill, R. D., and Brule-Babel, A. L.**, High frequency of pollen embryoid formation and plant regeneration in *Triticum aestivum* L. on monosaccharide containing media, *Plant Sci.*, 66, 255, 1990.
45. **Chu, C. C., Wang, C. C., Sun, C. S., Chien, N. F., Yin, K. C., and Hsu, C.**, Investigations on the induction and morphogenesis of wheat (*Triticum vulgare*) pollen plants, *Acta Bot. Sin.*, 15, 1, 1973 (in Chinese).
46. **Chu, C. C., Wang, C. C., Sun, C. S., Chien, N. F., Yin, K. C., Chu, C. Y., and Bi, F. Y.**, Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources, *Sci. Sin.*, 18, 659, 1975 (in Chinese).
47. **Chuang, C. C., Ouyang, T. W., Chia, Hsu., Chou, S. M., and Ching, C. K.**, A set of potato media for wheat anther culture, in *Proc. Symp. Plant Tissue Culture*, Science Press, Peking, 1978, 51.
48. **Cooper, D. B., Sears, R. G., Lookhart, G. L., and Jones, B. L.**, Heritable somaclonal variation in gliadin proteins of wheat plants derived from immature embryo callus culture, *Theor. Appl. Genet.*, 71, 784, 1986.
49. **Craig, I. L.**, Haploid plants (2n = 21) from *in vitro* anther culture of *Triticum aestivum*, *Can. J. Genet. Cytol.*, 16, 697, 1974.
50. **Cure, W. W. and Mott, R. L.**, A comparative anatomical study of organogenesis in cultured tissues of maize, wheat and oats, *Physiol. Plant.*, 42, 91, 1978.
51. **Dale, P. J., Marks, M. S., Brown, M. M., Woolston, C. J., Gunn, H. V., Mullineaux, P. M., Lewis, D. M., Kemp, J. M., Chen, D. F., Gilmour, D. M., and Flavell, R. B.**, Agroinfection of wheat: inoculation of *in vitro* grown seedlings and embryos, *Plant Sci.*, 63, 237, 1989.
52. **Daniell, H., Krishnan, M., and McFadden, B. F.**, Transient expression of B-glucuronidase in different cellular compartments following biolistic delivery of foreign DNA into wheat leaves and calli, *Plant Cell Rep.*, 9, 615, 1991.
53. **Datta, S. K. and Wenzel, G.**, Isolated microspore derived plant formation via embryogenesis in *Triticum aestivum* L., *Plant Sci.*, 48, 49, 1987.
54. **Datta, S. K., Potrykus, I., Bolik, M., and Wenzel, G.**, Culture of isolated pollen of wheat (*Triticum aestivum* L.), in *Wheat*, Biotechnology in Agriculture and Forestry, Vol. 13, Bajaj, Y. P. S., Ed., Springer-Verlag, Berlin, 1990, 435.
55. **De Buyser, J., Hachemi-Rachedi, S., Lemee, M.-L., Sejourne, S., Marcotte, J. L., and Henry, Y.**, Aneuploid analysis of anther culture response in wheat, *Plant Breed.*, 109, 339, 1992.
56. **Dekeyser, R. A., Claes, B., De Rycke, R. M. U., Habets, M. E., Van Montagu, M. C., and Caplan, A. B.**, Transient gene expression in intact and organized rice, *Plant Cell*, 2, 591, 1990.
57. **De La Roche, A. I., Keller, W. A., Singh, J., and Simlnovitch, D.**, Isolation of protoplasts from unhardened and hardened tissues of winter rye and wheat, *Can. J. Bot.*, 55, 1181, 1977.
58. **Delporte, F., Mingeo, C., Salomon, F., and Jacquemin, J. M.**, Expression of zein family protein in wheat endosperm (*Triticum aestivum* L.) after transformation by copollination with plasmidic DNA, in *Eucarpia Symposium on Genetic Manipulation in Plant Breeding*, abstract I/P4, Tarragona, 1991.
59. **Deng, W. Y., Lin, X. Y., and Shao, Q. Q.**, *Agrobacterium tumefaciens* can transform *Triticum aestivum* and *Hordeum vulgare* of Gramineae, *Sci. China (Ser. B)*, 33, 27, 1990.
60. **Dennis, E. S. and Brettell, R. I. S.**, DNA methylation of maize transposable elements is correlated with activity, *Philos. Trans. Soc. London Ser. B.*, 326, 217, 1990.
61. **Devaux, P.**, Significant higher proportions of 1BL-1RS wheat-rye translocations lines among doubled haploid progenies derived from F_1 hybrids between translocated and non-translocated wheat cultivars, in *Progress in Plant Cellular and Molecular Biology*, Proc. 7th Int. Congr. on Plant Tissue and Cell Culture, Nijkamp, H. J. J., van der Plas, L. H., and van Aartrijk, J., Eds., Kluwer Academic Press, Amsterdam, 1990, 442.
62. **Dhir, S. K., Pajeau, M. E., Fromm, M. E., and Fry, J. E.**, Anthocyanin genes as visual markers for wheat transformation, in Gordon Conference "Plant Cell and Tissue Culture" Plant Transgenes: Tools for Discovery and Design, June 13–18, 1993.
63. **Dhir, S. K., Pajeau, M. E., Fromm, M. E., and Fry, J. E.**, Anthocyanin genes as visual markers for wheat transformation, in Guthrie Centenary Conference, Sydney, Australia, Sept. 11–16, 1993.
64. **Dudits, D., Nemet, G., and Haydu, Z.**, Study of callus growth and organ formation in wheat (*Triticum aestivum*) tissue cultures, *Can. J. Bot.*, 53, 957, 1975.
65. **Dudits, D. and Nemet, G.**, Semaine d'etude cerealcultivé, Gembloux, Belgium, 1976, 127.
66. **Du, J., Liu, H., Xie, D. X., Hua, X. J., and Fan, Y. L.**, Injection of exogenous DNA into young floral tillers of wheat, *Genet. Manipulat. Plants*, 5, 8, 1989 (in Chinese).
67. **Dutta Gupta, S. and Ahmed, R.**, Role of inoculum weight on tissue growth and plantlet regeneration in *Triticum aestivum* L., *Ind. J. Exp. Biol.*, 25, 235, 1987.
68. **Dutta Gupta, S. and Ghosh, P. D.**, Chromosome analysis in callus culture of *Triticum aestivum*, *Nucleus*, 26, 16, 1983.
69. **Eapen, S. and Rao, P. S.**, Factors controlling callus initiation, growth and plant regeneration in breadwheat (*Triticum aestivum* L.), *Proc. Ind. Acad. Sci. (Plant Sci.)*, 94, 33, 1985.
70. **Eapen, S. and Rao, P. S.**, Factors controlling pollen embryogenesis in *Triticale* and wheat, *Proc. Ind. Natl. Sci. Acad.*, 51, 353, 1985.

71. **Ekiz, H. and Konzak, C. F.,** Nuclear and cytoplasmic control of anther culture response in wheat: I. Analyses of alloplasmic lines, *Crop Sci.*, 31, 1421, 1991.
72. **Ekiz, H. and Konzak, C. F.,** Nuclear and cytoplasmic control of anther culture response in wheat: II. Common wheat X alloplasmic lines, *Crop Sci.*, 31, 1427, 1991.
73. **Ekiz, H. and Konzak, C. F.,** Nuclear and cytoplasmic control of anther culture response in wheat: III. Common wheat crosses, *Crop Sci.*, 31, 1431, 1991.
74. **Elena, E. B. and Ginzo, H. D.,** Effect of auxin levels on shoot formation with different embryo tissues from a cultivar and a commercial hybrid of wheat (*Triticum aestivum* L.), *J. Plant Physiol.*, 132, 600, 1988.
75. **Evans, P. K., Keats, A. G., and Cocking, E. C.,** Isolation of protoplasts from cereal leaves, *Planta*, 104, 178, 1972.
76. **Fadel, F. and Wenzel, G.,** Medium-genotype-interaction on androgenetic haploid production in wheat, *Plant Breed.*, 105, 278, 1990.
77. **Fadel, F. and Wenzel, G.,** *In vitro* selection for tolerance to *Fusarium* in F_1 microspore populations of wheat, *Plant Breed.*, 110, 89, 1993.
78. **Felsenburg, T., Feldman, M., and Galun, E.,** Aneuploid and alloplasmic lines as tools for the study of nuclear and cytoplasmic control of culture ability and regeneration of scutellar calli from common wheat, *Theor. Appl. Genet.*, 74, 802, 1987.
79. **Feng, G. H. and Ouyang, J.,** The effects of KNO_3 concentration in callus induction medium for wheat anther culture, *Plant Cell Tiss. Organ Cult.*, 12, 3, 1988.
80. **Foroughi-Wehr, B. and Zeller, F. J.,** *In vitro* microspore reaction of different German wheat cultivars, *Theor. Appl. Genet.*, 79, 77, 1990.
81. **Fujii, T.,** Callus formation in wheat anthers, *Wheat Inf. Serv.*, 31, 1, 1970.
82. **Galiba, G. and Erdei, L.,** Dependence of wheat callus growth, differentiation and mineral content on carbohydrate supply, *Plant Sci.*, 45, 65, 1986.
83. **Galiba, G. and Yamada, Y.,** A novel method for increasing the frequency of somatic embryogenesis in wheat tissue culture by NaCl and KCl supplementation, *Plant Cell Rep.*, 7, 55, 1988.
84. **Gamborg, O. L. and Eveleigh, D. E.,** Culture methods and detection of glucanases in suspension cultures of wheat and barley, *Can. J. Biochem.*, 416, 417, 1968.
85. **Gosch-Wackerle, G., Avivi, L., and Galun, E.,** Induction, culture and differentiation of callus from immature rachises, seeds and embryos of *Triticum*, *Z. Pflanzenphysiol.*, 91, 267, 1979.
86. **Graves, A. E., Goldman, S. L., Banks, S. W., and Graves, A. C. F.,** Scanning electron microscope studies of *Agrobacterium tumefaciens*. Attachment to *Zea mays*, *Gladiolus* sp. and *Triticum aestivum*, *J. Bacteriol.*, 170, 2395, 1988.
87. **Grossmann, K., Sauerbrey, E., and Jung, J.,** Influence of growth retardants and ethylene-generating compounds on culture response of leaf explants from wheat (*Triticum aestivum* L.), *J. Plant Physiol.*, 135, 725, 1990.
88. **Guha, S. and Maheshwari, S. C.,** *In vitro* production of embryos from anthers of *Datura*, *Nature*, 204, 497, 1964.
89. **Guha, S. and Maheshwari, S. C.,** Cell division and differentiation of embryos in the pollen grains of *Datura in vitro*, *Nature*, 212, 97, 1966.
90. **Guo, G. Q., Xia, G. M., Li, Z. Y., and Chen, H. M.,** Direct somatic embryogenesis and plant regeneration from protoplast-derived cells of wheat (*Triticum aestivum* L.), *Sci. Sin.*, 9, 970, 1990.
91. **Guo, G. Q., Xia, G. M., Li, Z. Y., and Chen, H. M.,** Direct somatic embryogenesis and plant regeneration from protoplast-derived cells of wheat (*Triticum aestivum*), *Sci. China (Ser. B)*, 34, 438, 1991.
92. **Guenzi, A. C., Mornhinweg, D. W., and Johnson, B. B.,** Genetic analysis of a grass dwarf mutation induced by wheat callus culture, *Theor. Appl. Genet.*, 84, 952, 1992.
93. **Gupta, S., Ghosal, K. K., and Gadgil, V. N.,** Haploid tissue culture of *Triticum aestivum* var. Sonalika and *Pisum sativum* var. B 22, *Ind. Agric.*, 16, 277, 1972.
94. **Harris, R., Wright, M., Byrne, M., Varnum, J., Brightwell, B., and Schubert, K.,** Callus formation and plantlet regeneration from protoplasts derived from suspension cultures of wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 7, 337, 1988.
95. **Hartmann, C., De Buyser, J., Henry, Y., Falconet, D., Lejeune, B., Benslimane, A. A., Quétier, F., and Rode, A.,** Time-course of mitochondrial genome variation in wheat embryogenic somatic tissue cultures, *Plant Sci.*, 53, 191, 1987.
96. **Hartmann, C., Henry, Y., De Buyser, J., Aubry, C., and Rode, A.,** Identification of new mitochondrial genome organizations in wheat plants regenerated from somatic tissue cultures, *Theor. Appl. Genet.*, 77, 169, 1989.
97. **Hashim, Z. N., Campbell, W. F., and Carman, J. G.,** Normalization of the DNA content of telophase cells from wheat calli by nutrient modifications, *Theor. Appl. Genet.*, 82, 413, 1991.
98. **Hautpmann, R. M., Ozias-Akins, P., Vasil, V., Tabaeizadeh, Z., Rogers, S. G., Horsch, R. B., Vasil, I. K., and Fraley, R. T.,** Transient expression of electroporated DNA in monocotyledonous and dicotyledonous species, *Plant Cell Rep.*, 6, 265, 1987.
99. **Hauptmann, R. M., Vasil, V., Ozias-Akins, P., Tabaeizadeh, Z., Rogers, S. G., Fraley, R. T., Horsch, R. B., and Vasil, I. K.,** Evaluation of selectable markers for obtaining stable transformants in the Gramineae, *Plant Physiol.*, 86, 602, 1988.
100. **Hayashi, Y. and Shimamoto, K.,** Wheat protoplast culture: embryogenic colony formation from protoplasts, *Plant Cell Rep.*, 7, 414, 1988.
101. **He, D. G. and Ouyang, J.,** Callus and plantlet formation from cultured wheat anthers at different developmental stages, *Plant Sci. Lett.*, 33, 71, 1984.
102. **He, D. G., Tanner, G., and Scott, K. J.,** Somatic embryogenesis and morphogenesis in callus derived from the epiblast of immature embryos of wheat (*Triticum aestivum*), *Plant Sci.*, 45, 119, 1986.
103. **He, D. G., Yang, Y. M., Bertram, J., and Scott, K. J.,** The histological development of the regenerative tissue derived from cultured immature embryos of wheat (*Triticum aestivum* L.), *Plant Sci.*, 68, 103, 1990.
104. **He, D. G., Yang, Y. M., and Scott, K. J.,** A comparison of scutellum callus and epiblast callus induction in wheat: the effect of genotype, embryo age and medium, *Plant Sci.*, 57, 225, 1988.
105. **He, D. G., Yang, Y. M., and Scott, K. J.,** Plant regeneration from protoplasts of wheat (*Triticum aestivum* cv. Hartog), *Plant Cell Rep.*, 11, 16, 1992.
106. **Henry, Y. and De Buyser, J.,** Float culture of wheat anthers, *Theor. Appl. Genet.*, 60, 77, 1981.

107. **Henry, Y., De Buyser, J., Guenegou, T., and Ory, C.,** Wheat microspore embryogenesis during *in vitro* anther culture, *Theor. Appl. Genet.*, 67, 439, 1984.
108. **Henry, Y. and De Buyser, J.,** Effect of the 1B/1R translocation on anther culture ability in wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 4, 307, 1985.
109. **Henry, Y. and De Buyser, J.,** Wheat anther culture: agronomic performance of doubled haploid lines and the release of a new variety "Florin", in *Wheat*, Biotechnology in Agriculture and Forestry, Vol. 13, Bajaj, Y. P. S., Eds., Springer-Verlag, Berlin, 1990, 285.
110. **Heberle-Bors, E. and Odenbach, W.,** *In vitro* pollen embryogenesis and cytoplasmic male sterility in *Triticum aestivum*, *Z. Pflanzenzuecht.*, 95, 14, 1985.
111. **Hess, D., Dressler, K., and Nimmrichter, R.,** Transformation experiments by pipetting *Agrobacterium* into the spikelets of wheat (*Triticum aestivum* L.), *Plant Sci.*, 72, 233, 1990.
112. **Heszky, L. and Mesch, J.,** Anther culture investigations in cereal gene bank collection, *Z. Pflanzenzuecht.*, 77, 187, 1976.
113. **Heyne, E. G.,** Ed., *Wheat and Wheat Improvement*, 2nd ed., American Society of Agronomy, Madison, 1987.
114. **Heyser, J. W., Nabors, M. W., MacKinnon, C., Dykes, T. A., DeMott, K. J., Kautzman, D. C., and Mujeeb-Kazi, A.,** Long-term, high-frequency plant regeneration and the induction of somatic embryogenesis in callus cultures of wheat (*Triticum aestivum* L.), *Z. Pflanzenzuecht.*, 94, 218, 1985.
115. **Higgins, P. and Mathias, R. J.,** The effect of the 4B chromosomes of hexaploid wheat on the growth and regeneration of callus cultures, *Theor. Appl. Genet.*, 74, 439, 1987.
116. **Hoess, R., Wierzbicki, A., and Abremski, K.,** Isolation and characterization of intermediates in site-specific recombination, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 6840, 1987.
117. **Hu, H.,** Wheat: improvement through anther culture, in *Biotechnology in Agriculture and Forestry*, Vol. 2, Crops I. and Bajaj, Y. P. S., Ed., Springer-Verlag, Berlin, 1986, 55.
118. **Huang, B.,** Effects of incubation temperature on microspore callus production and plant regeneration in wheat anther cultures, *Plant Cell Tiss. Organ Cult.*, 9, 45, 1987.
119. **Hunsinger, H. and Schauz, K.,** The influence of dicamba on somatic embryogenesis and frequency of plant regeneration from cultured immature embryos of wheat (*Triticum aestivum* L.), *Plant Breed.*, 98, 119, 1987.
120. **Ishii, C.,** Callus induction and shoot differentiation of wheat, oat and barley, *Proc. 5th Intl. Congr. Plant Tissue Cell Culture*, 185, 1982.
121. **Johnston, S. A.,** Biolistic transformation: microbes to mice, *Nature*, 346, 776, 1990.
122. **Jones, A. M. and Petolino, J. F.,** Effects of donor plant genotype and growth environment on anther culture of soft-red winter wheat (*Triticum aestivum* L.), *Plant Cell Tiss. Organ Cult.*, 8, 215, 1987.
123. **Jones, A. M. and Petolino, J. F.,** Effects of support medium on embryo and plant production from cultured anthers of soft-red winter wheat (*Triticum aestivum* L.), *Plant Cell Tiss. Organ Cult.*, 12, 255, 1988.
124. **Kaleikau, E. K., Sears, R. G., and Gill, B. S.,** Monosomic analysis of tissue culture response in wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 78, 625, 1989.
125. **Kaleikau, E. K., Sears, R. G., and Gill, B. S.,** Control of tissue culture response in wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 78, 783, 1989.
126. **Karp, A. and Maddock, S. E.,** Chromosome variation in wheat plants regenerated from cultured immature embryos, *Theor. Appl. Genet.*, 67, 249, 1984.
127. **Karp, A., Wu, Q. S., Steele, S. H., and Jones, M. G. K.,** Chromosome variation in dividing protoplasts and cell suspensions of wheat, *Theor. Appl. Genet.*, 74, 140, 1987.
128. **Kasha, K. J., Ziauddin, A., and Cho, U. H.,** Haploids in cereal improvement: anther and microspore culture, in *Gene Manipulation in Plant Improvement*. II, Gustafson, J. P., Ed., Plenum Press, New York, 1990, 213.
129. **Kendall, E. J., Kartha, K., Qureshi, J. A., and Chermak, P.,** Cryopreservation of immature spring wheat zygotic embryos using an abscisic acid pretreatment, *Plant Cell Rep.*, 12, 89, 1993.
130. **Kelly, S. L.,** The Effect of Sucrose, Dextran and Glucose on Anther Culture of Winter Wheat (*Triticum aestivum*), M. Sc. thesis, University of Guelph, Canada, 1986.
131. **Kloti, A., Iglesias, V. A., Wunn, J., Burkhardt, P. K., Datta, S. K., and Potrykus, I.,** Gene transfer by electroporation into intact scutellum cells of wheat embryos, *Plant Cell Rep.*, 12, 671, 1993.
132. **Konzak, C. F., Ekiz, H., Zhou, H., and Davis, M. A.,** Nuclear and cytoplasmic control of anther culture responses in wheat-potential of cytoplasm genetic variability. Special lecture by C.F.K., in *Proc. Symp. U.S.-Japan Assoc. Wheat Breed. Genet.*, Sasakuma, T. and Kinoshita, T., Eds., 1991, 251.
133. **Konzak, C. F. and Zhou, H.,** Anther culture methods for doubled haploid production in wheat, *Cereal Res. Commun.*, 19, 147, 1991.
134. **Langridge, P., Lazzeri, P., and Lörz, H.,** A segment of rye chromosome 1 enhances growth and embryogenesis of calli derived from immature embryos of wheat, *Plant Cell Rep.*, 10, 148, 1991.
135. **Langridge, P., Brettschneider, R., Lazzeri, P., and Lörz, H.,** Transformation of cereals via *Agrobacterium* and the pollen pathway: a critical assessment, *Plant J.*, 2, 631, 1992.
136. **Larkin, P. J. and Scowcroft, W. R.,** Somaclonal variation — a novel source of variability from cell cultures for plant improvement, *Theor. Appl. Genet.*, 60, 197, 1981.
137. **Larkin, P. J., Ryan, S. A., Brettell, R. I. S., and Scowcroft, W. R.,** Heritable somaclonal variation in wheat, *Theor. Appl. Genet.*, 67, 443, 1984.
138. **Lashermes, P., Engin, G., and Ortiz-Ferrara, G.,** Anther culture of wheat (*Triticum aestivum*) adapted to dry areas of West Asia and North Africa, *J. Genet. Breed.*, 45, 33, 1991.
139. **Last, D. I. and Brettell, R. I. S.,** Embryo yield in wheat anther culture is influenced by the choice of sugar in the culture medium, *Plant Cell Rep.*, 9, 14, 1990.
140. **Last, D. I., Brettell, R. I. S., Chamberlain, D. A., Chaudhury, A. M., Larkin, P. J., Marsh, E. L., Peacock, W. J., and Dennis, E. S.,** pEmu: an improved promoter for gene expression in cereal cells, *Theor. Appl. Genet.*, 81, 581, 1991.
141. **Lazar, M. D., Collins, G. B., and Vian, W. E.,** Genetic and environmental effects on the growth and differentiation of wheat somatic cell cultures, *J. Heredity*, 74, 353, 1983.
142. **Lazar, M. D., Baenziger, P. S., and Schaeffer, G. W.,** Combining abilities and heritability of callus formation and plantlet regeneration in wheat (*Triticum aestivum* L.) anther cultures, *Theor. Appl. Genet.*, 68, 131, 1984.
143. **Lazar, M. D. Schaeffer, G. W., and Baenziger, P. S.,** Cultivar and cultivar X environment effects on the develop-

ment of callus and polyhaploid plants from anther cultures of wheat, *Theor. Appl. Genet.*, 67, 273, 1984.

144. **Lazar, M. D., Schaeffer, G. W., and Baenziger, P. S.,** The physical environment in relation to high frequency callus and plantlet development in anther cultures of wheat (*Triticum aestivum* L.) cv. Chris, *J. Plant Physiol.*, 121, 103, 1985.
145. **Lazar, M. D., Chen, T. H. H., Gusta, L. V., and Kartha, K. K.,** Somaclonal variation for freezing tolerance in a population derived from norstar winter wheat, *Theor. Appl. Genet.*, 75, 480, 1988.
146. **Lazar, M. D., Schaeffer, G. W., and Baenziger, P. S.,** The effects of interactions of culture environment with genotype on wheat (*Triticum aestivum*) anther culture response, *Plant Cell Rep.*, 8, 525, 1990.
147. **Lee, B. T., Murdoch, K., Topping, J., de Both, M. T. J., Wu, Q. S., Karp, A., Steele, S., Symonds, C., Kreis, M., and Jones, M. G. K.,** Isolation, culture and morphogenesis from wheat protoplasts, and study of expression of DNA constructs by direct gene transfer, *Plant Cell Tiss. Organ Cult.*, 12, 223, 1988.
148. **Lee, B., Murdoch, K., Topping, J., Kreis, M., and Jones, M. G. K.,** Transient gene expression in aleurone protoplasts isolated from developing caryopses of barley and wheat, *Plant Mol. Biol.*, 13, 21, 1989.
149. **Lettre, J. J., Kelly, S. L., Marsolais, A. A., and Kasha, J. J.,** Wheat anther culture using liquid media, in *Wheat*, Biotechnology in Agriculture and Forestry, Vol. 13, Bajaj, Y. P. S., Ed., Springer-Verlag, Berlin, 1990, 416.
150. **Li, H., Qureshi, J. A., and Kartha, K. K.,** The influence of different temperature treatments on anther cultures response of spring wheat (*Triticum aestivum* L.), *Plant Sci.*, 57, 55, 1988.
151. **Li, X. H., Yan, Q. S., Huang, M. J., Sun, Y. N., and Li, W. B.,** Divisions of cells regenerated from mesophyll protoplasts of wheat (*Triticum aestivum* L.), in *Advances in Protoplast Research*, Ferenczy, L. and Farkas, G. L., Eds., Pergamon Press, Oxford, 1980, 281.
152. **Li, Z. Y., Xia, G. M., and Chen, H. M.,** Somatic embryogenesis and plant regeneration from protoplasts isolated from embryogenic cell suspension of wheat (*Triticum aestivum* L.), *Plant Cell Tiss. Organ Cult.*, 28, 79, 1992.
153. **Li, Z. Y., Xia, G. M., Chen, H. M., and Guo, C. Q.,** Plant regeneration from protoplasts derived from embryogenic suspension cultures of wheat (*Triticum aestivum* L.), *J. Plant Physiol.*, 139, 714, 1992.
154. **Liang, G. H., Sangduen, N., Heyne, E. G., and Sears, R. G.,** Polyhaploid production through anther culture in common wheat, *J. Heredity*, 73, 360, 1982.
155. **Liang, G. H., Xu, A., and Tang, H.,** Direct generation of wheat haploids via anther culture, *Crop Sci.*, 27, 336, 1987.
156. **Ling, D. X., Luckett, D. J., and Darvey, N. L.,** Low-dose gamma irradiation promotes wheat anther culture response, *Aust. J. Bot.*, 39, 467, 1991.
157. **Lonsdale, D., Onde, S., and Cuming, A.,** Transient expression of exogenous DNA in intact, viable wheat embryos following particle bombardment, *J. Exp. Bot.*, 41, 1161, 1990.
158. **Lörz, H., Baker, B., and Schell, J.,** Gene transfer to cereal cells mediated by protoplast transformation, *Mol. Gen. Genet.*, 199, 178, 1985.
159. **Lörz, H., Gobel, E., and Brown, P.,** Advances in tissue culture and progress towards genetic transformation of cereals, *Plant Breed.*, 100, 1, 1988.
160. **LoSchiavo, F., Pitto, L., Giuliano, G., Torti, G., Nuti-Ronchi, V., Marazziti, D., Vergara, R., Orselli, S., and Terzi, M.,** DNA methylation of embryogenic carrot cell cultures and its variations as caused by mutation, differentiation, hormones and hypomethylating drugs, *Theor. Appl. Genet.*, 77, 325, 1989.
161. **Lu, C. S., Sharma, H. C., and Ohm, H. W.,** Wheat anther culture: effect of genotype and environmental conditions, *Plant Cell Tiss. Organ Cult.*, 24, 233, 1991.
162. **Luckett, D. J., Venkatnagappa, S., Darvey, N. L., and Smithard, R. A.,** Anther culture of Australian wheat germplasm using modified C17 medium and membrane rafts, *Aust. J. Plant Physiol.*, 18, 357, 1991.
163. **Luckett, D. J. and Darvey, N. L.,** Utilisation of microspore culture in wheat and barley improvement, *Aust. J. Bot.*, 40, 807, 1992.
164. **Ludwig, S. R., Bowen, B., Beach, L., and Wessler, S. R.,** A regulatory gene as a novel visible marker for maize transformation, *Science*, 24, 449, 1990.
165. **Maddock, S. E.,** Suspension and protoplast culture of hexaploid wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 6, 23, 1987.
166. **Maddock, S. E., Lancaster, V. A., Risiott, R., Franklin, J.,** Plant regeneration from cultured immature embryos and inflorescences of 25 cultivars of wheat (*Triticum aestivum*), *Ind. J. Exp. Bot*, 34, 915, 1983.
167. **Maddock, S. E. and Semple, J. T.,** Field assessment of somaclonal variation in wheat, *J. Exp. Bot.*, 37, 1065, 1986.
168. **Magnusson, I. and Bornman, C. H.,** Anatomical observations on somatic embryogenesis from scutellar tissues of immature zygotic embryos of *Triticum aestivum*, *Physiol. Plant*, 63, 137, 1985.
169. **Mahalakshmi, A., Khurana, J. P., Maheshwari, S. C., and Khurana, P.,** Rapid induction of somatic embryogenesis by 2,4-D in leaf base cultures of wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 1993, in press.
170. **Mahalakshmi, A., Maheshwari, S. C., and Khurana, P.,** High frequency divisions in leaf base protoplasts of wheat (*Triticum aestivum* L.), *J. Plant Biochem. Biotech.*, 2, 61, 1993.
171. **Maheshwari, N., Rajyalakshmi, K., Chowdhary, C. N., Grover, A., Tyagi, A. K., and Maheshwari, S. C.,** *In vitro* culture of wheat and rice for understanding the molecular basis of somatic embryogenesis and for transformation, in *The Impact of Biotechnology in Agriculture*, Sangwan, R. S. and Sangwan-Norreel, B. S., Eds., Kluwer Academic Publishers, Netherlands, 1990, 191.
172. **Marburger, J. E., Sammons, D. J., and Schaeffer, G. W.,** Effect of a modified potato medium on anther culture of wheat, *Crop Sci.*, 27, 351, 1987.
173. **Maretzki, A.,** Tissue culture: its prospects and problems, in *Sugarcane Improvement Through Breeding*, Heinz, D. J., Ed., Elsevier, Amsterdam, 1987, 343.
174. **Marks, M. S., Kemp, J. M., Wooston, C. J., and Dale, P. J.,** Agroinfection of wheat: a comparison of *Agrobacterium* strains, *Plant Sci.*, 63, 247, 1989.
175. **Marsolais, A. A., Seguin-Swartz, G., and Kasha, K. J.,** The influence of anther cold pretreatments and donor plant genotypes on in vitro androgenesis in wheat (*Triticum aestivum* L.), *Plant Cell Tiss. Organ Cult.*, 3, 69, 1984.
176. **Mascarenhas, A. F., Hendre, R. R., Seetharama Rao, B., and Jagannathan, V.,** Tissue cultures of maize, wheat, jawar and rice, *Ind. J. Exp. Biol.*, 7, 65, 1969.
177. **Mascarenhas, A. F., Hendre, R. R., Pathak, M., Nadgir, A. L., and Jagannathan, V.,** Tissue cultures of maize, wheat,

rice and sorghum. III. Growth and nutrition of root cultures of maize, wheat and sorghum in agitated liquid media, *Ind. J. Exp. Biol.*, 13, 112, 1975.

178. **Mascarenhas, A. F., Pathak, M., Hendre, R. R., Ghugale, D. D., and Jagannathan, V.,** Tissue cultures of maize, wheat, rice and sorghum. IV. Studies of organ differentiation in tissue cultures of maize, wheat and rice, *Ind. J. Exp. Biol.*, 13, 116, 1975.
179. **Mascarenhas, A. F., Pathak, M., Hendre, R. R., and Jagannathan, V.,** Tissue cultures of maize, wheat, rice and sorghum. I. Initiation of viable callus and root cultures, *Ind. J. Exp. Biol.*, 13, 103, 1975.
180. **Mathias, R. J. and Atkinson, E.,** *In vitro* expression of genes affecting whole plant phenotype — the effect of Rht/Gai alleles on the callus culture response of wheat (*Triticum aestivum* L. em. Thell), *Theor. Appl. Genet.*, 75, 474, 1988.
181. **Mathias, R. J. and Boyd, L. A.,** Cefotaxime stimulates callus growth, embryogenesis and regeneration in hexaploid bread wheat (*Triticum aestivum* L. em. Thell), *Plant Sci.*, 46, 217, 1986.
182. **Mathias, R. J., Brown, C., Pearson, D., and Brooks, F. J.,** Cereal transformation and tissue culture studies, in *Annual Report*, AFRC/IPSR Cambridge, U.K., 1987, 16.
183. **Mathias, R. J. and Simpson, E. S.,** The interaction of genotype and culture medium on the tissue culture responses of wheat (*Triticum aestivum* L. em. Thell) callus, *Plant Cell Tiss. Organ Cult.*, 7, 31, 1986.
184. **Matzeit, V., Schaefer, S., Kammann, M., Schalk, H. J., Schell, J., and Gronenborn, B.,** Wheat dwarf virus vectors replicate and express foreign genes in cells of monocotyledonous plants, *Plant Cell*, 3, 247, 1991.
185. **McElroy, D., Zhang, W., Cao, J., and Wu, R.,** Isolation of an efficient actin promoter for use in rice transformation, *Plant Cell*, 2, 163, 1990.
186. **McGregor, L. J. and McHughen, A.,** The influence of various cultural factors on anther culture of four cultivars of spring wheat (*Triticum aestivum* L.), *Can. J. Plant Sci.*, 70, 183, 1990.
187. **McHughen, A.,** Rapid regeneration of wheat *in vitro*, *Ann. Bot.*, 51, 851, 1983.
188. **Mejza, S. J., Morgant, V., DiBona, D. E., and Wong, J. R.,** Plant regeneration from isolated microspores of *Triticum aestivum*, *Plant Cell Rep.*, 12, 149, 1993.
189. **Mendel, R. R. and Teeri, T. H.,** Cereal crops (III), barley, wheat, rye, oats, in *Transformation of Plants and Soil Microorganism*, Wang, K., Herrera-Estrella, A., and Van Montagu, M., Eds., Cambridge University Press, New York, 1993, 1.
190. **Messens, E., Dekeyser, R., and Stachel, S. E.,** A nontransformable *Triticum monococcum* monocotyledonous culture produces the potent *Agrobacterium* vir-inducing compound ethyl ferulate, *Proc. Natl. Acad. Sci. U.S.A.*, 87, 4368, 1990.
191. **Mohmand, A. S. and Nabors, M. W.,** Somaclonal variant plants of wheat derived from mature embryo explants of three genotypes, *Plant Cell Rep.*, 8, 558, 1990.
192. **Mohmand, A. S. and Nabors, M. W.,** Comparison of two methods for callus culture and plant regeneration in wheat (*Triticum aestivum*), *Plant Cell Tiss. Organ Cult.*, 26, 185, 1991.
193. **Mooney, P. A. and Goodwin, P. B.,** Adherence of *Agrobacterium tumefaciens* to the cells of immature wheat embryos, *Plant Cell Tiss. Organ Cult.*, 25, 199, 1991.
194. **Mooney, P. A., Goodwin, P. B., Dennis, E. S., and Llewellyn, D. J.,** *Agrobacterium tumefaciens*-gene transfer into wheat tissues, *Plant Cell Tiss. Organ Cult.*, 25, 209, 1991.
195. **Morère-Le Paven, M. C., Henry, Y., De Buyser, J., Corre, F., Hartmann, C., and Rode, A.,** Organ/tissue-specific changes in the mitochondrial genome organization of in vitro cultures derived from different explants of a single wheat variety, *Theor. Appl. Genet.*, 85, 1, 1992.
196. **Morère-Le Paven, M. C., De Buyser, J., Henry, Y., Corre, F., Hartmann, C., and Rode, A.,** Multiple patterns of mtDNA reorganization in plants regenerated from different in vitro cultured explants of a single wheat variety, *Theor. Appl. Genet.*, 85, 9, 1992.
197. **Morrish, F. M. and Fromm, M. E.,** Cereal transformation methods, *Curr. Opin. Biotechnol.*, 3, 141, 1992.
198. **Muller, G., Borschel, H., Vahl, U., Wiberg, A., Hartel, H., and Damisch, W.,** Die Nutzung der Antherkulturmethode im Zuchtprozess von Winterweizen. I. Zur Androgenesefähigkeit von 1B-1R Weizen-Translokationsformen, *Plant Breed.*, 102, 196, 1989.
199. **Muller, G., Bohme, T., Borschel, H., Vahl, U., and Wiberg, A.,** Die Nutzung der Antherenkulturmethode im Zuchtprozess von Winterweizen. III. Zur Antherenkultureignung von Winterweizen — F_1-Populationen mit beiden heterozygoten Chromosomenpaaren 1AL-1AS/1AL-1RS und 1BL-1BS/1BL-1RS, *Plant Breed.*, 104, 272, 1990.
200. **Mulligan, R. C., Howard, B. H., and Berg, P.,** Synthesis of rabbit B-globin in cultured monkey kidney cells following infection with a SV40 B-globin recombinant genome, *Nature*, 277, 108, 1979.
201. **Murata, M.,** Effects of auxin and cytokinin on induction of sister chromatid exchanges in cultured cells of wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 78, 521, 1989.
202. **Nabors, M. W., Heyser, J. W., Dykes, T. A., and DeMott, K. J.,** Long-duration, high-frequency plant regeneration from cereal tissue cultures, *Planta*, 157, 385, 1983.
203. **Oard, J. H., Paige, D., and Dvorak, J.,** Chimeric gene expression using maize intron in cultured cells of bread wheat, *Plant Cell Rep.*, 8, 156, 1989.
204. **Oard, J. H., Paige, D. F., Simmonds, J. A., and Gradziel, T. M.,** Transient gene expression in maize, rice and wheat cells using an airgun apparatus, *Plant Physiol.*, 92, 334, 1990.
205. **Ogura, H.,** Considerations on the involvements of genetical differences with callus formation in wheat, *Wheat Inf. Serv.*, 44, 5, 1977.
206. **O'Hara, J. F. and Street, H. E.,** Wheat callus culture: the initiation, growth and organogenesis of callus derived from various explant sources, *Ann. Bot.*, 42, 1029, 1978.
207. **Okuno, T. and Furusawa, I.,** A simple method for the isolation of intact mesophyll protoplasts from cereal plants, *Plant Cell Physiol.*, 18, 1357, 1977.
208. **Oono, K.,** Putative homozygous mutations in regenerated plants of rice, *Mol. Gen. Genet.*, 198, 377, 1985.
209. **Opik, H.,** The ultrastructure of coleoptile cells in dry rice (*Oryza sativa* L.) grains after anhydrous fixation with osmium tetraoxide vapour, *New Phytol.*, 85, 521, 1980.
210. **Orshinsky, B. R., McGregor, L. J., Johnson, G. I. E., Hucl, P., and Kartha, K. K.,** Improved embryoid induction and green shoot regeneration from wheat anthers cultured in medium with maltose, *Plant Cell Rep.*, 9, 365, 1990.
211. **Ou-Lee, T. M., Turgeon, R., and Wu, R.,** Expression of a foreign gene linked to either a plant-virus or a *Drosophila* promoter, after electroporation of protoplasts of rice, wheat and sorghum, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 6815, 1986.

212. **Ouyang, J. W., Hu, H., Chuang, C. C., and Tseng, C. C.,** Induction of pollen plants from anthers of *Triticum aestivum* L. cultured *in vitro*, *Sci. Sin.*, 16, 79, 1973.
213. **Ouyang, J. W., Zhou, S. M., and Jia, S. E.,** The response of anther culture to culture temperature in *Triticum aestivum*, *Theor. Appl. Genet.*, 66, 101, 1983.
214. **Ouyang, J. W.,** Induction of pollen plants in *Triticum aestivum*, in *Haploids of Higher Plants In Vitro*, Hu, H. and Yang, H., Eds., China Academic Publishers, Beijing, 1986, 26.
215. **Ouyang, J. W., He, D. G., Feng, G. H., and Jia, S. E.,** The response of anther culture to culture temperature varies with growth conditions of anther-donor plants, *Plant Sci.*, 49, 145, 1987.
216. **Ozias-Akins, P. and Vasil, I. K.,** Plant regeneration from cultured immature embryos and inflorescences of *Triticum aestivum* L. (wheat): evidence for somatic embryogenesis, *Protoplasma*, 110, 95, 1982.
217. **Ozias-Akins, P. and Vasil, I., K.,** Callus induction and growth from the mature embryo of *Triticum aestivum* (wheat), *Protoplasma*, 115, 104, 1983.
218. **Ozias-Akins, P. and Vasil, I. K.,** Improved efficiency and normalization of somatic embryogenesis in *Triticum aestivum* (wheat), *Protoplasma*, 117, 40, 1983.
219. **Pan, J. L. and Gao, G. H.,** The production of wheat pollen embryos and the influence of some factors on its frequency of induction, *Acta Bot. Sin.*, 20, 122, 1978 (in Chinese).
220. **Papenfuss, J. M. and Carman, J. G.,** Enhanced regeneration from wheat callus cultures using dicamba and kinetin, *Crop. Sci.*, 27, 588, 1987.
221. **Pauk, J. and Szarka, B.,** Protoplast isolation and culture investigations in common wheat (*Triticum aestivum* L.), *Abstr. 8th Int. Protoplast Symp. Physiol. Plant.*, 82, A4/27, 1991.
222. **Pei, W. and Yurong, C.,** Effects of growth conditions of anther-donor plants on the production of pollen plants in wheat anther culture, *Acta. Gen. Sin.*, 7, 64, 1980.
223. **Peri, A., Kless, H., Blumenthal, A., Galili, G., and Galun, E.,** Improvement of plant regeneration and GUS expression in scutellar wheat calli by optimization of culture conditions and DNA-microprojectile delivery procedures, *Mol. Gen. Genet.*, 235, 279, 1992.
224. **Phillips, R. L., Kaeppler, S. M., and Peschke, V. M.,** Do we understand somaclonal variation? in *Progress in Plant Cellular and Molecular Biology*, Nijkamp, H. J. J., Van Der Plas, L. H. W., and Van Aartrijk, J., Eds., Kluwer Academic Publishers, Netherlands, 1990, 136.
225. **Picard, E. and De Buyser, J.,** Obtention de plantules haploides de *Triticum aestivum* L. a partir de culture d'antheres *in vitro*, *C. R. Acad. Sci. (Paris)*, 277, 1463, 1973.
226. **Picard, E. and De Buyser, J.,** Noveaux resultats concernant la culture d'antheres de *Triticum aestivum* L. Conditions de regeneration des plantes haploides et production de lignees entierement homozygotes, *C. R. Acad. Sci. (Paris)*, 281, 989, 1975.
227. **Picard, E. and De Buyser, J.,** High production of embryoids in anther culture of pollen derived homozygous spring wheats, *Ann. Amelior. Plant*, 27, 483, 1977.
228. **Picard, E., Hours, C., Gregoire, S., Phan, T. H., and Meunier, J. P.,** Significant improvement of androgenetic haploid and doubled haploid induction from wheat plants treated with a chemical hybridization agent, *Theor. Appl. Genet.*, 74, 289, 1987.
229. **Picard, E., Jacquemin, J. M., Granier, F., Bobin, M., and Forgeois, P.,** Genetic transformation of (*Triticum aestivum*) by plasmid DNA uptake during pollen tube germination, in *Proc. 7th Int. Wheat Genet. Symp.*, Vol. 1, Miller, T. E. and Koebner, R. M. D., Eds., Institute of Plant Science Research, Cambridge, 1988, 779.
230. **Picard, E., Rode, A., Doussinault, G., Rousset, M., and Rives, M.,** Wheat (*Triticum aestivum*): in vitro production and utilization of doubled haploids, in *Haploids in Crop Improvement Biotechnology in Agriculture and Forestry*, Vol. 12, Bajaj, Y. P. S., Ed., Springer-Verlag, Berlin, 1990, 101.
231. **Purnhauser, L., Medgyesy, P., Czako, M., Dix, P. J., and Marton, L.,** Stimulation of shoot regeneration in *Triticum aestivum* and *Nicotiana plumbaginifolia* Viv. tissue cultures using the ethylene inhibitor $AgNO_3$, *Plant Cell Rep.*, 6, 1, 1987.
232. **Qiao, Y. M., Cattaneo, M., Locatelli, F., and Lupotto, E.,** Plant regeneration from long-term suspension culture derived protoplasts of hexaploid wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 11, 262, 1992.
233. **Qureshi, J. A., Kartha, K. K., Abrams, S. R., and Steinhauer, L.,** Modulation of somatic embryogenesis in early and late-stage embryos of wheat (*Triticum aestivum* L.) under the influence of (±)-abscisic acid and its analogs, *Plant Cell Tiss. Organ Cult.*, 18, 55, 1989.
234. **Rajyalakshmi, K., Dhir, S. K., Maheshwari, N., and Maheshwari, S. C.,** Callusing and regeneration of plantlets via somatic embryogenesis from inflorescence cultures of *Triticum aestivum* L.: role of genotype and long-term retention of morphogenic potential, *Plant Breed.*, 101, 80, 1988.
235. **Rajyalakshmi, K., Grover, A., Maheshwari, N., Tyagi, A. K., and Maheshwari, S. C.,** High-frequency regeneration of plantlets from the leaf-bases via somatic embryogenesis and comparison of polypeptide profiles from morphogenic and non-morphogenic calli in wheat (*Triticum aestivum*), *Physiol. Plant.*, 82, 617, 1991.
236. **Redway, F. A., Vasil, V., Lu, D., and Vasil, I, K.,** Identification of callus types for long-term maintenance and regeneration from commercial cultivars of wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 79, 607, 1990.
237. **Redway, F. A., Vasil, V., and Vasil, I. K.,** Characterization and regeneration of wheat (*Triticum aestivum* L.) embryogenic cell suspension cultures, *Plant Cell Rep.*, 8, 714, 1990.
238. **Reed, J., Nahory, J., Jayne, S., Chang, Y. F., Stein, J., Grater, T., Miner, D., Walker, L., Wong, J., and Crossland, L.,** Production of transgenic wheat by particle bombardment using methotrexate as a selection agent, in Proceedings of Keystone Meeting, 1994, 303.
239. **Ren, Y. G., Jia, J. F., Li, M. Y., and Zhen, G. C.,** Plant regeneration from protoplasts isolated from callus cultures of immature inflorescenes of wheat (*Triticum aestivum* L.), *Chin. Sci. Bull.*, 34, 1648, 1990.
240. **Reynolds, T. L. and Kitto, S. L.,** Identification of embryoid-abundant genes that are temporally expressed during pollen embryogenesis in wheat anther cultures, *Plant Physiol.*, 100, 1744, 1992.
241. **Rode, A., Hartmann, C., De Buyser, J., and Henry, Y.,** Evidence for a direct relationship between mitochondrial genome organization and regeneration ability in hexaploid wheat somatic tissue cultures, *Curr. Genet.*, 14, 387, 1988.
242. **Rode, A., Hartmann, C., Falconet, D., Lejeune, B., Quétier, F., Benslimane, A., Henry, Y., and De Buyser, J.,** Exten-

sive mitochondrial DNA variation in somatic tissue cultures initiated from immature embryos, *Curr. Genet.*, 12, 369, 1987.

243. **Ryan, S. A. and Scowcroft, W. R.**, A somaclonal variant of wheat with additional B-amylase isozymes, *Theor. Appl. Genet.*, 73, 459, 1987.
244. **Ryan, S. A., Larkin, P. J., and Ellison, F. W.**, Somaclonal variation in some agronomic and quality characters in wheat, *Theor. Appl. Genet.*, 74, 77, 1987.
245. **Ryschka, S., Ryschka, U., and Schulze, J.**, Anatomical studies on the development of somatic embryoids in wheat and barley explants, *Biochem. Physiol. Pflanz.*, 187, 31, 1991.
246. **Sági, L. and Barnabás, B.**, Evidence for cytoplasmic control of in vitro microspore embryogenesis in the anther cul-

loid plant development from anthers and *in vitro* embryo culture of wheat, *Crop Sci.*, 19, 697, 1979.

250. **Schmid, J.**, In vitro production of haploids in *Triticum spelta*, in *Wheat*, Biotechnology in Agriculture and Forestry, Vol. 13, Bajaj, Y. P. S., Ed., Springer-Verlag, Berlin, 1990, 363.
251. **Schmid, J. and Keller, E. R.**, Effect of a gametocide on the induction of haploids in *Triticum aestivum*, in *Genetic Manipulation in Plant Breeding*, Horn, W., Jensen, C. J., Odenbach, W., and Schieder, O., Eds., Walter de Gruyter, Berlin, 1986, 347.
252. **Sears, R. G. and Deckard, E. L.**, Tissue culture variability in wheat: callus induction and plant regeneration, *Crop Sci.*, 22, 546, 1982.
253. **Shang, X. M. and Wang, W. C.**, DNA amplification, chromatin variations and polytene chromosomes in differentiating cells of common bread wheat in vitro and roots of regenerated plants in vivo, *Genome*, 34, 799, 1991.
254. **Shimada, T., Sasakuma, T., and Tsunewaki, K.**, *In vitro* culture of wheat tissues. I. Callus formation, organ redifferentiation, and single cell culture, *Can. J. Genet. Cytol.*, 11, 294, 1969.
255. **Shimada, T. and Makino, T.**, *In vitro* culture of wheat. III. Anther culture of the A genome aneuploids in common wheat, *Theor. Appl. Genet.*, 46, 407, 1975.
256. **Shimada, T.**, Plant regeneration from the callus induced from wheat embryo, *Jpn. J. Genet.*, 53, 371, 1978.
257. **Shimada, T. and Yamada, Y.**, Wheat plants regenerated from embryo cell cultures, *Jpn. J. Genet.*, 54, 379, 1979.
258. **Shimada, T., Seki, M., and Morikawa, H.**, Transient expression of B-glucuronidase (GUS) gene in wheat pollen embryos via microprojectile bombardment, *Wheat Inf. Serv.*, 72, 106, 1991.
259. **Simonson, R. L. and Baenziger, P. S.**, The effect of gelling agents on wheat anther and immature embryo culture, *Plant Breed.*, 109, 211, 1992.
260. **Simmond, J.**, Improved androgenesis of winter cultivars of *Triticum aestivum* L. in response to low temperature treatment of donor plants, *Plant Sci.*, 65, 225, 1989.
261. **Simmonds, J., Stewart, P., and Simmonds, D.**, Regeneration of *Triticum aestivum* apical explants after microinjection of germ line progenitor cells with DNA, *Physiol. Plant.*, 85, 197, 1992.
262. **Simmonds, J. A. and Grainger, J. L.**, The toxicity of antibiotics to protoplast cultures of *Triticum aestivum* L., *Plant Sci.*, 89, 209, 1993.
263. **Szakács, E., Kovács, G., Pauk, J., and Barnabás, B.**, Substitution analysis of callus induction and plant regeneration from anther culture in wheat (*Triticum aestivum* L.), *Plant Cell Rep.*, 7, 127, 1988.
264. **Taylor, M. G., Vasil, V., and Vasil, I. K.**, Enhanced GUS gene expression in cereal/grass cell suspensions and immature embryos using the maize ubiquitin-based plasmid pAHC25, *Plant Cell Rep.*, 12, 491, 1993.

268. **Usami, S., Morikawa, S., Takebe, I., and Machida, Y.**, Absence in monocotyledonous plants of the diffusible plant factors inducing T-DNA circularization and *vir* gene expression in *Agrobacterium*, *Mol. Gen. Genet.*, 209, 221, 1987.
269. **Usami, S., Okamoto, S., Takebe, I., and Machida, Y.**, Factor inducing *Agrobacterium tumefaciens vir* gene expression is present in monocotyledonous plants, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 3748, 1988.
270. **Vasil, I. K.**, Developing call and tissue culture systems for the improvement of cereal and grass crops, *J. Plant Physiol.*, 128, 193, 1987.
271. **Vasil, I. K. and Vasil, V.**, Regeneration in cereal and other grass species, in *Cell Culture and Somatic Cell Genetic Variability*, Vasil, I. K., Ed., Academic Press, Orlando, FL, 1986, 121.
272. **Vasil, V., Brown, S. M., Re, D., Fromm, M. E., and Vasil, I. K.**, Stably transformed callus lines from microprojectile bombardment of cell suspension cultures of wheat, *Biol Technology*, 9, 743, 1991.
273. **Vasil, V., Castillo, A. M., Fromm, M. E., and Vasil, I. K.**, Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus, *Bio/Technology*, 10, 667, 1992.
274. **Vasil, V., Redway, F., and Vasil, I. K.**, Regeneration of plants from embryogenic suspension culture protoplasts of wheat (*Triticum aestivum* L.), *Bio/Technology*, 8, 429, 1990.
275. **Vasil, V., Srivastava, V., Castillo, A. M., Fromm, M. E., and Vasil, I. K.**, Rapid production of transgenic wheat plants by direct bombardment of cultured immature embryos, *Biol Technology*, 11, 1553, 1993.
276. **Wang, B. K. and Yu, Y. A.**, Study of effect of radiation on anther culture in wheat, in *Application of Atomic Energy in Agriculture*, Vol. 3, 1984 (in Chinese).
277. **Wang, C. C., Chu, C. C., Sun, C. S., Wu, S. H., Yin, K. C., and Hsu, C.**, The androgenesis in wheat (*Triticum aestivum*) anthers cultured *in vitro*, *Sci. Sin.*, 16, 218, 1973 (in Chinese).

278. **Wang, H. B., Li, X. H., Sun, Y. R., Chen, J., Zhu, Z., Fang, R., Wang, P., and Wei, J. Q.,** Culture of wheat protoplast, *Sci. Sin. Ser. B.*, 33, 294, 1990.
279. **Wang, J. J., Hu, O. F., Wang, H. M., and Tang, Y. L.,** Studies on increasing the induction frequency of foreign callus in wheat, *Hereditas (Beijing)*, 1, 28, 1981 (in Chinese).
280. **Wang, R. and Chen, Y. P.,** Preliminary study on prediction of height of pollen plant H_2 generation in winter wheat grown in field, *Acta Agron. Sin.*, 9, 283, 1983 (in Chinese).
281. **Wang, W. C. and Nguyen, H. T.,** A novel approach for efficient plant regeneration from long-term suspension culture of wheat, *Plant Cell Rep.*, 8, 639, 1990.
282. **Wang, W. C., Shang, X. M., Yucel, M., and Nguyen, H. T.,** Selection of cultured wheat cells for tolerance to high temperature stress, *Crop Sci.*, 33, 315, 1993.
283. **Wang, Y. C., Klein, T. M., Fromm, M., Cao, J., Sanford, J. C., and Wu, R.,** Transient expression of foreign genes in rice, wheat and soybean cells following particle bombardment, *Plant Mol. Biol.*, 11, 433, 1988.
284. **Wei, Z. M.,** Pollen callus culture in *Triticum aestivum*, *Theor. Appl. Genet.*, 63, 71, 1982.
285. **Weeks, J. T., Blechl, A. E., and Anderson, O. D.,** Bialaphos selection of regenerable wheat callus, *In Vitro Cell. Anim. Devel. Biol.*, 29a(3), 64A, 1993.
286. **Weeks, J. T., Anderson, O. D., and Blechl, A. E.,** Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*), *Plant Physiol.*, 102, 1077, 1993.
287. **Wernicke, W. and Milkovits, L.,** Developmental gradients in wheat leaves — response of leaf segments in different genotypes cultured *in vitro*, *J. Plant Physiol.*, 115, 49, 1984.
288. **Wernicke, W. and Milkovits, L.,** The regeneration potential of wheat shoot meristems in the presence and absence of 2,4-dichlorophenoxyacetic acid, *Protoplasma*, 131, 131, 1986.
289. **Werr, W. and Lörz, H.,** Transient gene expression in a Gramineae cell line. A rapid procedure for studying plant promoters, *Mol. Gen. Genet.*, 202, 471, 1986.
290. **Winfield, M., Davey, M. R., and Karp, A.,** A comparison of chromosome instability in cell suspensions of diploid,
293. **Yang, D. L.,** Preliminary report on wheat haploid breeding research — Changwei Institute of Agricultural Sciences, *Chin. J. Bot.*, 2, 7, 1975 (in Chinese; cited in Ouyang, 1981).
294. **Yang, Y. M., He, D. G., and Scott, K. J.,** Establishment of embryogenic suspension cultures of wheat by continuous callus selection, *Aust. J. Plant Physiol.*, 18, 445, 1991.
295. **Yin, D. C., Wei, Q. J., Yu, Q. C., and Weng, L.,** Effect of gamma radiation on wheat pollen development in anther culture, in *Proc. Int. Symp. Genetic Manipulation in Crops*, Published for IRRI and Academia Sinica, Cassell Tycosly, London, 1984.
296. **Yurkova, G. N., Levenko, B. A., and Novozhilov, O. V.,** Induction of plant regeneration in wheat tissue culture, *Biochem. Physiol. Pflanz.*, 176, 236, 1981.
297. **Zaghmout, O. M.-F. and Trolinder, N. L.,** Simple and efficient method for directly electroporating *Agrobacterium* plasmid DNA into wheat callus cells, *Nucl. Acid. Res.*, 21, 1048, 1993.
298. **Zaghmout, O. M.-F. and Trolinder, N. L.,** Factors affecting transient gene expression in protoplasts isolated from very slowly growing embryogenic callus cultures of wheat (*Triticum aestivum* L.), *Theor. Appl. Genet.*, 86, 721, 1993.
299. **Zamora, A. B. and Scott, K. J.,** Callus formation and plant regeneration from wheat leaves, *Plant Sci. Lett.*, 29, 183, 1983.
300. **Zhang, B. H. and Meng, X. G.,** Introducing a medium for wheat anther culture, *Hereditas (Beijing)*, 4, 31, 1982 (in Chinese).
301. **Zhang, Y. L. and Li, S.,** Anther culture of monosomics in *Triticum aestivum*, *Hereditas (Beijing)*, 6, 7, 1983 (in Chinese).
302. **Zhou, H., Zheng, Y., and Konzak, C. F.,** Osmotic potential of media affecting green plant percentage in wheat anther culture, *Plant Cell Rep.*, 10, 63, 1991.
303. **Zhou, H., Ball, S. T., and Konzak, C. F.,** Functional properties of ficoll and their influence on anther culture responses of wheat, *Plant Cell Tissue Organ Culture*, 30, 77, 1992.
304. **Zhou, H., Stiff, C. M., and Konzak, C. F.,** Stably transformed callus of wheat by electroporation-induced direct

12
DNA Transfer and Gene Expression in Transgenic Cereals

RICHARD I.S. BRETTELL AND FIONA R. MURRAY

CSIRO Division of Plant Industry, Institute of Plant Production and Processing, GPO Box 1600, Canberra ACT 2601, Australia

Introduction

This decade has seen rapid advances in the development of methods for the routine transfer of genes to cereals, and there are now examples of genetic transformation for all major cereal crop species. Until recently, the two methods which have found wide application are direct gene transfer to protoplasts, and bombardment of intact tissues with DNA-coated microprojectiles. There is now substantial evidence that Agrobacterium can provide an alternative means of transferring genes to rice, and has further potential as a vector for cereal transformation.

For practical application of the technology it is essential to have transformed plants which show continued expression and stable inheritance of transgenes that alter the plant phenotype. Studies of the behaviour of transgenes in cereals have been hampered by the relatively low transformation rates obtained using current transformation methods. However, continued development of existing strategies together with the promise of alternative methods suggests this difficulty will be overcome by enhanced efficiency of gene transfer.

The significance of cereals as the principal source of food throughout the world means that they are an obvious target for genetic engineering, although the acceptance of genetically modified cereal products by the consumer remains to be tested. Despite the technological limitations described in this review, there are now examples where genes have been introduced to alter an agronomic or quality trait of a cereal. The characteristics that are receiving most attention are resistance to pests and diseases, and components of grain quality. Both are amenable to manipulation by the introduction of single genes. An example of pest resistance is the introduction of the Bt toxin gene from *Bacillus thuringiensis* into maize, providing protection against damage caused by the European corn borer (Koziel *et al.*, 1993; Armstrong *et al.*, 1995). The transfer to rice of a chitinase gene with activity against fungal cell walls

Abbreviations: GUS, β-glucuronidase; *gus*, gene coding for β-glucuronidase (*uidA* or *gusA*); HPT, hygromycin phosphotransferase; *hpt*, gene coding for hygromycin phosphotransferase (including *hph*)

Biotechnology and Genetic Engineering Reviews – Vol. 13, December 1995
0264–8725/94/12/1–38 $20.00 + $0.00

has produced lines with enhanced resistance to sheath blight caused by the fungus *Rhizoctonia solani* (Lin *et al.*, 1995). Specific pathogen resistance genes have recently been cloned and characterised (Staskawicz *et al.*, 1995) and provide a further opportunity for engineering disease resistance. Cereal genes in this category that have been isolated include those corresponding to the *Xa-21* locus for resistance to bacterial blight in rice (Wang *et al.*, 1995) and the *Cre-3* locus for resistance to cereal cyst nematode in wheat (E.Lagudah and O. Moullet, personal communication). In terms of grain quality, the target genes depend on the end use of the grain. In wheat, attention is focussed on genes altering protein and starch composition (Lazzeri and Shewry, 1993; Anderson *et al.*, 1994). In barley, the principal targets are genes that have a role in defining malting quality, such as those coding for α-amylase and β-glucanase (McElroy and Jacobsen, 1995).

This review will first consider the merits of the different methods that have been applied for transforming cereals, before discussing the conditions necessary to ensure stable expression of the transferred genetic material.

Methods for transforming cereals

MICROPROJECTILE BOMBARDMENT

Compared to other methods for plant transformation, microprojectile bombardment or biolistics is a relatively recent innovation. Microscopic particles of tungsten or gold are coated with DNA and then fired into target cells (Sanford *et al.*, 1987). In a proportion of the cells the DNA will be transferred to the nucleus (Yamashita, Jida and Morikawa, 1991; Hunold, Bronner and Hahne, 1994) providing an opportunity for integration into the genome. The particles need to be sufficiently small to penetrate individual cells without destroying their integrity and viability. Initial experiments were made using gunpowder charges to accelerate the DNA-coated metal particles, however consistent results were obtained using helium gas at pressure and this is the preferred propellent in most devices used today (Sanford *et al.*, 1991; Birch and Franks, 1991; Finer *et al.*, 1992; Brown *et al.*, 1994 ; Nabulsi *et al.*, 1994; Songstad *et al.*, 1995). Another approach has been to use a high-voltage electrical discharge through a water droplet to create a shock wave and accelerate a thin sheet carrying the DNA-coated particles (McCabe *et al.*, 1988). Overall, microprojectile bombardment has so far proved to be the most versatile method for cereal transformation and has been used to transform all major species (*Table 1*).

Particle bombardment efficiently delivers DNA into the surface layers of the target material. In multicellular material, potentially transformed cells are connected to and surrounded by non-transformed cells. Consequently, the successful recovery of transformed cultures depends on an efficient and discriminating means of selecting for the proliferation of those cells which carry the introduced DNA. As with other methods of transformation, selection has been achieved using herbicides or antibiotics. Among the widely applied selectable marker genes in cereals, is the *bar* gene from *Streptomyces hygroscopicus* which codes for phosphinothricin N-acetyl transferase (PAT) and confers resistance to phosphinothricin based herbicides such as bialaphos and Basta (Spencer *et al.*, 1990; Wilmink and Dons, 1993; Dennehey *et al.*, 1994). Alternative strategies have employed mutant plant genes for resistance to herbicides

Table 1. Chronology of development of methods for the genetic transformation of cereals. The dates refer to the first substantiated example of stable gene transfer using a given method for the different species. References are provided in the text.

	Microprojectile bombardment	Direct gene transfer to protoplasts	Agrobacterium	Other methods
Rice	1991	1988	1994	
Wheat	1992	1994		
Barley	1994	1995		
Maize	1990	1988		1992
Sorghum	1993			
Oats	1992			
Rye	1994			

such as the sulphonylureas (Fromm *et al.*, 1990; Wilmink and Dons, 1993; Chamberlain *et al.*, 1994) or glyphosate (Vasil *et al.*, 1991). Effective selection of transformed tissues and plants has also been achieved with antibiotics such as kanamycin, geneticin, paromomycin and hygromycin B and corresponding bacterial genes for resistance, *nptII* and *hpt* (Vasil *et al.*, 1991; Bower and Birch, 1992; Li *et al.*, 1993; Hagio *et al.*, 1995; Torbert *et al.*, 1995).

The recovery of transformed plants from the bombarded tissues depends on the ability to regenerate plants from the target cells. Initial success was achieved using established embryogenic suspension cultures as target tissue, for example with maize (Gordon-Kamm *et al.*, 1990; Fromm *et al.*, 1990), rice (Cao *et al.*, 1992), wheat (Vasil *et al.*, 1992) and oats (Somers *et al.*, 1992). However, such long term cultures have a number of disadvantages; they are difficult to establish and the longer they are maintained, the greater is the likelihood that deleterious mutations will have accumulated in the cell lineages which will eventually produce the regenerated plants. In subsequent experiments, embryogenic callus tissues from solid culture media have also been found to provide suitable target materials (Bower and Birch, 1992; Somers *et al.*, 1994; Wan, Widholm and Lemaux, 1995). To obviate the difficulties associated with maintaining embryogenic cultures, there is currently a trend towards using material directly from the plant for bombardment. Immature embryos have been found to be a suitable source of actively dividing tissues for different species such as rice (Christou, Ford and Kofron, 1991, maize (Koziel *et al.*, 1993), barley (Ritala *et al.*, 1994; Wan and Lemaux, 1994), sorghum (Casas *et al.*, 1993), wheat (Vasil *et al.*, 1993; Weeks, Anderson and Blechl, 1993; Becker, Brettschneider and Lörz, 1994; Nehra *et al.*, 1994), triticale (Zimny *et al.*, 1995) and rye (Castillo, Vasil and Vasil, 1994). The target is the part of the embryo, the scutellum, which in the intact cereal grain is the tissue in contact with the developing endosperm. The cells in the surface layers of the scutellum will undergo somatic embryogenesis when a cereal embryo at an appropriate stage of development is placed on a culture medium with the embryo axis in direct contact with medium. Histological studies in barley have confirmed that cells in the scutellar epithelium and sub-epithelial layers participate in the production of embryogenic tissue, and thus provide a suitable target for transformation (Ryschka, Ryschka and Schulze, 1991). An alternative approach has been to target the developing shoot meristem of very young embryos in maize, resulting in the production of

chimaeric meristems from which germline transformants can be recovered (Lowe *et al.*, 1995). This was achieved by culturing sections of the shoot apex on medium containing cytokinin to stimulate shoot proliferation. The method appears to be adaptable to a range of genotypes, including elite inbred lines.

Frequencies of transformation reported for microparticle bombardment are variable, but most commonly lie within the range of 0.1%–1.0% for the percentage of shot explants yielding a transformed cell line. The numbers can often be confounding, as multiple plantlets can be regenerated from a single embryogenic culture. It is therefore important to distinguish between clones and independently transformed plants. The highest frequencies, greater than 5%, have been quoted for rice (Christou, Ford and Kofron, 1991), while wheat transformation appears to be at least ten times less efficient (Weeks, Anderson and Blechl, 1993; Becker, Brettschneider and Lörz, 1994).

Another target that has been used with some success is the immature inflorescence which has been shown to be capable of initiating embryogenic cultures in a number of cereal species (Brettell, Wernicke and Thomas, 1980; Ozias-Akins and Vasil, 1982; Rangan and Vasil, 1983). Barcelo *et al.* (1994) were able to recover transgenic tritordeum plants from inflorescences given a short period of preculture, with the highest frequency of 17 transformants for 178 inflorescences achieved when the tissue was placed in culture the day prior to bombardment. An alternative approach has been to target microspores. These are immature pollen grains, which, although they only have a haploid chromosome complement, are able to divide and regenerate plantlets. This approach has been tried in a number of laboratories, particularly with barley, and there is evidence that it can be used for the recovery of transformed plants. In one example, transformed plants were recovered at a frequency of 1 in 10^7 microspores bombarded (Jähne *et al.*, 1994). It has yet to be demonstrated whether this approach of using gametic cells will be generally applicable for transforming a range of cereal species.

DIRECT GENE TRANSFER TO PROTOPLASTS

Protoplasts are single plant cells from which the cell wall has been removed, and they can provide a large and uniform population of target cells for the introduction of DNA. Plants were first regenerated from tobacco protoplasts more than twenty years ago (Nagata and Takebe, 1971). The technique has since been extended to a range of cereals, following initial success with rice (Abdullah, Cocking and Thompson, 1986; Yamada, Yang and Tang, 1986; Kyozuka, Hayashi and Shimamoto, 1987). However, the recovery of plants from protoplasts is still technically demanding and cannot be achieved routinely for all species.

Populations of protoplasts are prepared by digesting plant tissues in a solution containing a mixture of cellulytic and pectolytic enzymes. DNA, as circular or linearised plasmid, can be introduced into isolated protoplasts following treatment with polyethylene glycol to facilitate the transport of macromolecules through the plasma membrane (Krens *et al.*, 1982; Potrykus *et al.*, 1985; Maas and Werr, 1989). Alternatively the protoplasts can be subjected to electroporation whereby an electric field is used to drive the uptake of DNA (Fromm *et al.*, 1985; Nagata, 1989; Larkin *et al.*, 1990). As techniques have been developed to regenerate plants from cultured

cereal protoplasts, direct gene transfer has been used to produce transgenic plants of rice (Toriyama *et al.*, 1988; Shimamoto *et al.*, 1989; Zhang *et al.*, 1988) and maize (Rhodes *et al.*, 1988).

Application of direct gene transfer to protoplasts is still severely limited by the difficulties encountered in regenerating plants from the cultured protoplasts. In cereals, regeneration of plants has only been achieved from protoplasts isolated from embryogenic cultures, and generally from those kept in liquid suspension. There are no confirmed reports of regeneration of plants from protoplasts isolated directly from an intact cereal plant. Embryogenic suspension cultures are difficult to establish and maintain, and the length of time required to establish suitable cultures appears to favour the accumulation of deleterious mutations. Thus for example, a common feature of barley regenerated from protoplasts is a high frequency of albino, chloroplast-deficient and therefore inviable plants. With wheat and barley, successful regeneration of plants from protoplasts has been achieved independently in a number of laboratories (Harris *et al.*, 1988; Ren *et al.*, 1989; Vasil, Redway and Vasil, 1990; Chang *et al.*, 1991; Jähne, Lazzeri and Lörz, 1991; He, Yang and Scott, 1992; Li *et al.*, 1992; Qiao *et al.*, 1992; Ahmed and Sagi, 1993; Pauk *et al.*, 1994), although the plants recovered in many of these cases lacked vigour and were sterile. While transformed plants have been obtained from protoplasts (He, Yang and Scott, 1994; Funatsuki *et al.*, 1995), there are to our knowledge still no examples where significant numbers of fertile transgenic wheat or barley plants have been recovered by this method. In maize, successful regeneration of plants is limited to a few genotypes but has led to the recovery of transformed plants (Rhodes *et al.*, 1988; Golovkin *et al.*, 1993; Sukhapinda *et al.*, 1993). Rice, on the other hand, appears to be more flexible, with protoplast transformation finding wide application (Toriyama *et al.*, 1988; Zhang *et al.*, 1988; Shimamoto *et al.*, 1989; Datta *et al.*, 1990, 1992; Fujimoto *et al.*, 1993; Shimamoto *et al.*, 1993, Uchimiya *et al.*, 1993; Chamberlain *et al.*, 1994).

AGROBACTERIUM

Genetic manipulation of plant tissues is far from being a recent innovation. Crown gall disease, which affects a number of dicotyledonous species, results from the introduction of foreign DNA into plant cells at sites of wounding, from a Ti (tumour inducing) plasmid harboured by the pathogenic bacterium *Agrobacterium tumefaciens* (Binns and Thomashow, 1988). The Ti-plasmid contains genes which alter the plant host's metabolism in favour of the pathogen. This naturally occurring genetic engineer has been exploited to mediate the introduction into target plant cells of a number of 'genes of interest' following their insertion into the segment of DNA, the T-DNA, transferred from the Ti-plasmid to the host plant.

Agrobacterium has found wide application as a vector for plant transformation, including species that are outside the natural host range of this bacterium. However, cereals have shown a notable reluctance to submit to transformation by Agrobacterium, although it has been demonstrated that infectious viral sequences can be introduced into cereal tissues through a process described as agroinfection (Grimsley *et al.*, 1987; Dale *et al.*, 1989). In these examples genomes from Geminiviruses, maize streak virus or wheat dwarf virus, were cloned into the T-DNA of *Agrobacterium tumefaciens*. Inoculation of a host plant with the bacteria resulted in systemic viral infection.

Attempts to modify this process and use viral vectors for the stable transformation of cereal tissues were unsuccessful, and it was concluded by some authors that the chances of transforming cereals with Agrobacterium were minimal, on the grounds that cereals lack the necessary wound response (Potrykus, 1990). It was argued that transformation would only occur at a wound site and any competent cells at or adjacent to the wound would be unlikely to survive the damage necessary to expose the cells to Agrobacterium.

There have been a number of isolated examples of cereal transformation using Agrobacterium; for example in rice (Raineri *et al.*, 1990; Chan, Lee and Chang, 1992; Chan *et al.*, 1993), maize (Gould *et al.*, 1991), wheat and barley (Deng, Lin and Shao, 1990; Mooney *et al.*, 1991). A lack of confirmation of these results in other laboratories cast some doubt on the validity of these experiments. However, in the past year a comprehensive study has been made in which transformation of japonica rice was achieved through co-cultivation of immature embryos, scutellum callus or suspension cells with *Agrobacterium tumefaciens* (Hiei *et al.*, 1994). Supporting the contention that DNA transfer was occurring in an analogous manner to that in dicotyledonous plants, sequence data showed that the boundaries of the T-DNA were similar to those found in transgenic tobacco with respect to short duplicated sequences. Expression and inheritance of the transgenes, coding for GUS and HPT, were demonstrated over three generations, providing further evidence for stable integration of the transgenes.

These results have since been extended to maize (Y. Hiei and co-workers, unpublished data), and provide an exciting prospect for transformation of other small grain cereals. Agrobacterium mediated transformation would provide an attractive alternative to microprojectile bombardment, particularly in species such as wheat and barley where high rates of transformation have been claimed by few laboratories. In the case of japonica rice, Hiei *et al.* (1994) reported an efficiency of transformation of scutellum-derived callus pieces of between 12% and 29%, which is comparable to that obtained for dicotyledonous plants. This efficiency of transformation would certainly rival the best that can be achieved for rice transformation by microprojectile bombardment.

OTHER METHODS FOR CEREAL TRANSFORMATION

Plant cells present many obstacles to the introduction of exogenous DNA, not least of which is the presence of a thick cellulosic cell wall. In addition extracellular nucleases, which degrade any unprotected DNA molecules, can be produced in abundance. Depending on the tissue, plant cells with large vacuoles and other compartments may also have correspondingly small nuclei as a target for gene transfer. In the absence of really efficient methods for cereal transformation, a number of other approaches have been tried in order to overcome these barriers to introducing exogenous gene sequences into the nucleus where integration into genomic DNA can occur.

One of the more promising approaches has been the use of silicon carbide fibres with an average diameter less than 1 μm (Kaeppler *et al.*, 1990; Wang *et al.*, 1995). The method involves vortexing a mixture of DNA, silicon carbide fibres and plant cells. The silicon carbide fibres act to pierce the plant cell wall allowing entry of DNA. This method has been used to produce fertile, transgenic maize plants following treatment of embryogenic suspension cultures (Frame *et al.*, 1994). However, it is not

clear that the method offers any obvious advantages over particle bombardment with the exception of the low cost of setting up the system. Another development has been the use of fine laser beams to puncture holes in the cell wall and plasma membrane, and this technique has recently been applied to rice transformation (Guo, Liang and Berns, 1995)

Among other methods which have been applied to cereals are injection of developing inflorescences with DNA solutions (de la Pena, Lörz and Schell, 1987), application of DNA to florets near the time at which pollination occurs providing an opportunity for transformation by the so called 'pollen tube pathway' (Luo and Wu, 1988), treating floral organs with Agrobacterium (Hess, Dressler and Nimmrichter, 1990), electrophoresis of DNA into seed tissues (Ahokas, 1989), and electroporation of intact tissues and cells (Li *et al.*, 1991; D'Halluin *et al.*, 1992; Klöti *et al.*, 1993; Zaghmout and Trolinder, 1993; Zhou, Stiff and Konzak, 1993; Laursen *et al.*, 1994). Of these methods, electroporation of intact cells and tissues appears to be the most promising. The other approaches have a poor record: the experiments have not been satisfactorily reproduced, and the evidence for integration of the DNA into the plant genome has been lacking. Moreover, most observations have been confined to first generation plants, and inheritance of putative transgenes has not been examined. It has been suggested that some of the results may have been due to artefact, such as the transformation of endophytic microorganisms suggested by Langridge *et al.* (1992). The market is always open for novel techniques for cereal transformation, but those investing effort in applying a new method should be cautious until it has been satisfactorily reproduced in other laboratories. Meanwhile it appears that existing methods can provide routes for achieving reliable and routine transfer of genes into cereal crop species.

Gene expression in transgenic cereals

STRATEGIES FOR INTRODUCING GENES INTO CEREALS

None of the methods described in the preceding section for delivering DNA into plant cells are so efficient that transformed plants can be routinely recovered without some additional process to enrich or select for the transformed cells. The selection of transformed cells is generally achieved in tissue culture through the use of a selectable marker gene in combination with the corresponding selective agent. For cereals, both antibiotics and herbicides have been shown to be effective (Wilmink and Dons, 1993).

The strategy for transfer of a gene of interest is to introduce this gene at the same time as the selectable marker gene. For methods such as microprojectile bombardment and direct gene transfer to protoplasts which do not involve the construction of a specialised plasmid vector, the two sets of DNA sequence can simply be mixed prior to the transformation. For example, in the case of electroporation of protoplasts, the two types of DNA are added and mixed into the electroporation buffer just before the electric field is applied. Under these conditions, a high proportion of the transformants selected with the selectable marker gene are also found to carry one or more copies of the non-selected gene of interest. This frequency of co-transformation is commonly recorded as greater than 50% and in some studies with microprojectile bombardment has been reported to be as high as 85%–90% (Barcelo *et al.*, 1994; Wan and Lemaux,

1994). To ensure co-integration of a gene of interest and the selectable marker gene, other researchers have adopted the strategy of using single plasmids containing the two genes (e.g. Fromm *et al.*, 1990; Becker, Brettschneider and Lörz, 1994; Cooley, Ford and Christou, 1995; Lin *et al.*, 1995).

For all methods of plant transformation, the transfer of exogenous DNA is essentially a random process with respect to the site of insertion into the genome. However, there is evidence from work with Agrobacterium that integration occurs preferentially into transcriptionally active sites (Koncz *et al.*, 1989). For the few examples in cereals in which a thorough study has been made of the inheritance of transgenes, the data indicate that multiple copies are frequently inserted at a single locus (Spencer *et al.*, 1992; Register *et al.*, 1994; Cooley, Ford and Christou, 1995; Peng *et al.*, 1995).

In the development of a new transformation procedure, the inclusion of a reporter gene can help gauge the efficiency of gene transfer. The most widely used reporter is the β-glucuronidase (*gus*) gene whose product hydrolyses a range of β-glucuronide substrates. Thus transformed cells can be visualised following incubation in the presence of 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc) which is cleaved to produce an insoluble blue precipitate (Jefferson, Kavanagh and Bevan, 1987). However, there are differences in opinion as to its reliability as a marker in cereals. In rice, the *gus* gene has been used to visualise the tissue specificity of expression determined by a range of promoter sequences (Kyozuka *et al.* 1993, 1994; Itoh *et al.*, 1995; Terada *et al.*, 1995). However, in sorghum and wheat, there are instances in which gene activity has not been detected even though presence of the *gus* gene has been confirmed by gel blot hybridisation (Casas *et al.*, 1993; I.K. Vasil, personal communication). Alternative reporters include firefly luciferase (de Wet *et al.*, 1987; Chia, Chan and Chua, 1994), aequorin from the jellyfish *Aequorea victoria* (Baulcombe, Chapman and Cruz, 1995; Niedz, Sussman and Satterlee, 1995), and *trans*-acting factors that regulate anthocyanin biosynthesis (Goff *et al.*, 1990; Bodeau and Walbot, 1995). However, the ideal reporter gene is still elusive. This would exhibit low background activity, have only moderate stability *in vivo*, have no detrimental effects on metabolism, and be easy to assay quantitatively (McElroy and Brettell, 1994)

REGULATION OF GENE EXPRESSION IN TRANSGENIC CEREALS

The application of gene transfer technology for cereal crop improvement will depend on a clear understanding of the molecular elements that regulate gene expression in plants. Promoters that have found general use in broad-leaved plants for driving constitutive levels of gene expression do not necessarily perform as well in cereals. For example, when a *gus* gene controlled by the 35S promoter from cauliflower mosaic virus was introduced by electroporation into protoplasts isolated from five different cereal cell lines, GUS activities ranged from 0.3% to 10.8% of the activity in *Nicotiana plumbaginifolia* protoplasts (Last *et al.*, 1991). The low efficacy of the 35S promoter in cereals is also seen with DNA introduced by microprojectile bombardment, both in transient expression (Schledzewski and Mendel, 1994) and in the selection of transformed cultures (Z.Y. Li, N.M. Upadhyanya, A.J. Gibbs and P.M. Waterhouse, unpublished results). This suggests that there may be differences between cereals and dicotyledons in the recognition of promoter sequences, due for example to

different affinities for transcription factors. This is further supported by observations that some cereal promoters show poor expression in cells of dicotyledons (Ellis *et al.*, 1987; Yamaguchi-Shinozaki *et al.*, 1990).

Promoter sequences are now available which have been found to give suitable levels of constitutive expression in transformed cereal tissues (McElroy and Brettell, 1994). Not surprisingly many of these have been derived from cereal genes. Examples that have been widely tested are promoters from the rice actin 1, *Act1*, gene (Zhang, McElroy and Wu, 1991) and a maize ubiquitin gene, *Ubi1* (Christensen, Sharrock and Quail, 1992). The strategy of including an intron in the transcribed portion of the gene has commonly been used to further enhance expression (Callis, Fromm and Walbot, 1987; Luehrsen and Walbot, 1991; McElroy *et al.*, 1991; Li *et al.*, 1995). Levels of gene transcription can also be increased by the addition of *cis*-acting elements such as in the pEmu promoter where ocs-elements from Agrobacterium have been combined with the *Adh1* promoter from maize (Last *et al.*, 1991).

The application of genetic engineering to modify characters such as grain quality and disease resistance in cereals will depend on promoters that control gene expression in a tissue specific manner. Tissue specific gene expression is desirable because the expression of transgenes in tissues where they are not required may drain the resources of the plant and result in deleterious effects such as stunting, increased susceptibility to pathogen attack and reduction in yield. For disease resistance, expression of resistance genes in tissues colonised by the pathogen may be sufficient for protection. Thus for barley yellow dwarf virus which is limited to phloem tissues of the host plant, a promoter such as the *rolC* promoter from *Agrobacterium rhizogenes* might be a suitable candidate when constructing genes that interfere with viral replication. This promoter shows a specificity of expression limited to vascular and embryogenic tissues (Matsuki *et al.*, 1989).

Similarly for genes influencing grain composition, such as those coding for seed protein or for enzymes involved in starch biosynthesis, it will be necessary to limit expression of transgenes to the developing endosperm. In a recent series of experiments with wheat, Blechl, Weeks and Anderson (1995) demonstrated expression of an intro-duced chimaeric Dy10-Dx5 high molecular weight glutenin gene linked to a wheat glutenin promoter. Following transformation by microprojectile bombardment, the majority of lines resistant to the selective agent bialaphos exhibited co-expression of the new storage protein gene in endosperm. The transgene was inherited over two genera-tions and levels of novel protein produced were comparable to those produced by the native glutenin genes. This result has been substantiated by studies on tritordeum, a hexaploid hybrid between tetraploid durum wheat and a diploid wild barley. A tritor-deum line was transformed with a glutenin Dx5 gene, again under the control of a wheat glutenin promoter (P.A. Lazzeri, P. Barcelo, F. Barro, A.S. Tatham, R. Fido and P.R. Shewry, unpublished results). The expression of the glutenin subunit is shown in *Figure 1*.

In contrast to broad-leaved dicot plants, there are relatively few studies of the regulation of gene expression in transgenic cereals (McElroy and Brettell, 1994). In those cases where tissue specific promoters have been introduced from a dicotyledon, the pattern of gene expression in the cereal is consistent with that observed in the source plant. Thus a tomato *rbcS* promoter was active in mesophyll cells of transgenic rice (Kyozuka *et al.*, 1993), although activities were less than those recorded for a rice

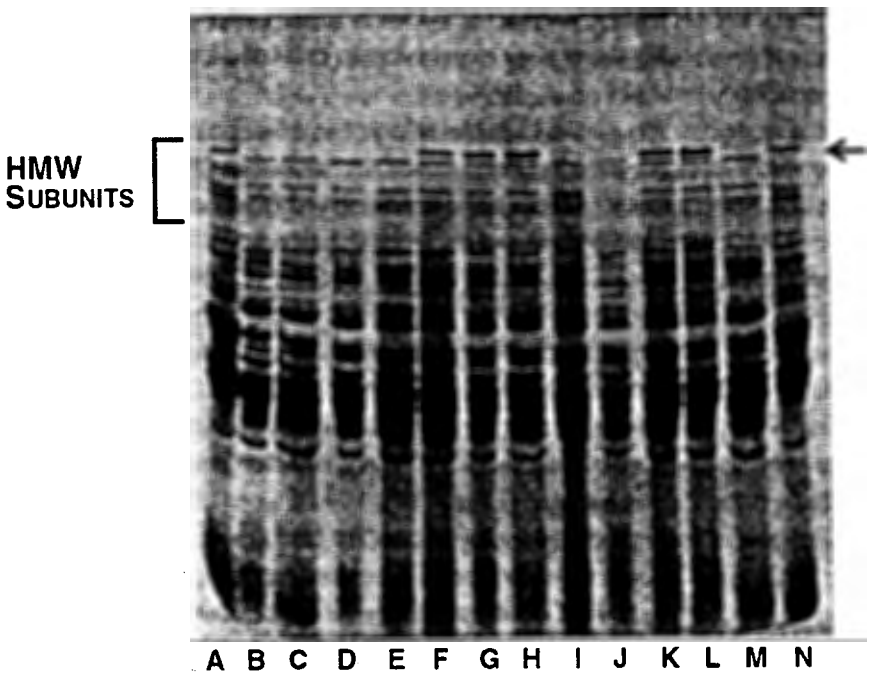


Figure 1. Expression of a wheat high molecular weight glutenin subunit in seeds of transgenic tritordeum. Lanes A and N are bread wheat controls, the presence of the 1Dx5 subunit is seen in lanes F, G, H, K and L (courtesy Paul A. Lazzeri).

rbcS promoter attached to the same *gus* reporter gene. The promoter of a barley aleurone-specific gene has been shown to confer aleurone cell-specific expression in transgenic rice (Kalla *et al.*, 1994). The regulation of expression by the promoter from a rice α-amylase gene has similarly been studied in transgenic rice (Itoh *et al.*, 1995). To ensure consistent levels of transgene expression, the available data would support a strategy of seeking homologous promoters for cereals, before utilising promoters from other plant genera.

GENE INACTIVATION IN TRANSGENIC CEREALS

Many of the descriptions of cereal transformation are confined to the expression of selectable marker and reporter genes, and to material growing in the glasshouse for one or two generations following gene transfer. As transgenic plants are grown more widely and examined under field conditions, it is found that examples are accumulating for transgenes being subject to progressive inactivation. This phenomenon has been discussed in detail in a recent review (Finnegan and McElroy, 1994), and it is clear that cereals provide no exception. For one study where a set of transgenic rice plants carrying multiple copies of the *nptII* gene was examined in the field, no plants could be found which expressed the antibiotic resistance coded by the transgene (Schuh *et al.*, 1993). In oats, six out of fifteen lines with a transgene for GUS showed aberrant segregation ratios with a higher than expected number of plants scoring negative for activity of the enzyme (Somers *et al.*, 1994). In a study of the inheritance of *gus* and *nptII* genes in rice, irregular expression of both genes was found in two of three families examined (Peng *et al.*, 1995). With transformed rice plants shown to contain at least one copy of both *gus* and *bar*, over 90% of the plants expressed bar but only 50% expressed gus (Cooley, Ford and Christou, 1995). In wheat, a study of six

independent lines transformed with *gus* and *bar* revealed differences in the stability of expression of the two transgenes (V. Srivastava, V. Vasil and I.K. Vasil, unpublished results). One line showed unstable expression of *bar*, but for five of the six lines GUS activity was not detectable in the T_2 generation even though presence of the *gus* gene was confirmed.

There is clear evidence that copy number and the position of a transgene in the plant genome will influence the level at which it is expressed (Hobbs, Kpodar and Delong, 1990; Linn *et al.*, 1990; Assaad, Tucker and Signer, 1993). In a recent study of transgenic rice, all plants with one or two copies of *gus* driven by the 35S promoter showed expression, whereas none of the plants with more than ten integrated copies of the gene showed GUS activity (Cooley, Ford and Christou, 1995). The insertion of multiple copies of a given sequence is associated with reduced gene expression and is commonly observed when DNA is introduced by any of the techniques that employ direct gene transfer, e.g. PEG treatment of protoplasts, particle bombardment (Peng *et al.*, 1990; Gordon-Kamm *et al.*, 1990; Somers *et al.*, 1992; Barcelo *et al.*, 1994; Register *et al.*, 1994; Cooley, Ford and Christou, 1995; Dalton *et al.*, 1995). The use of these techniques may therefore be a contributory factor to the high incidence of uneven expression of transgenes in cereals. Another feature which may determine the tendency of a transgene to be silenced is the extent to which it can be recognised as foreign in its new position in the plant genome. Disruption of compositionally homogeneous chromatin by the integration of foreign DNA may mark a region of the chromosome for inactivation (Meyer and Heidmann, 1994). The size and GC content of the transgene, the structure of the promoter, and the 3' sequences may all be important parameters in this regard. Such factors may explain the high incidence of inactivation seen in cereals for the *gus* reporter gene, compared to some of the selectable marker genes commonly used.

Inactivation of transgenes is frequently associated with specific methylation of DNA (Finnegan, Brettell and Dennis, 1993; Meyer and Heidmann, 1994). The occurrence of methylation as a cause or consequence of transgene inactivation, however, is still a matter of debate. However, the following experiments support a direct role for methylation. The long term expression and methylation status of *hpt* and *gus* genes were examined in transgenic pearl millet cells obtained by microprojectile bombardment. During long term culture expression of the *hpt* gene was maintained while there was a gradual decrease in measurable GUS activity, which could be recovered by exposing cells for two weeks to 10μM of the demethylating agent 5-azacytidine (Lambé, Dinant and Matagne, 1995). The inactivation of the *gus* gene was correlated with progressive methylation, revealed by isoschizomeric restriction enzymes, which differ in their ability to cut at sites with methylated cytosine residues. A further example suggesting the involvement of methylation is provided by experiments with rice. In one family of japonica rice transformed with the *gus* gene driven by a modified 35S promoter, consistent uneven expression of GUS was observed in leaves and roots when visualised with a histochemical X-Gluc stain. Rows and blocks of cells with intense blue staining were interspersed with non-staining cells, resulting in a 'spotty' phenotype (B.Witrzens and R. Brettell, unpublished). When seeds of this rice line, homozygous for the *gus* transgene, were germinated on culture medium containing the demethylating agent 5-azacytidine at a concentration of 60μM, uniform blue staining was observed in root tissues in contact with the medium indicating

that GUS activity had been restored. The effect was, however, transitory and was not observed in the leaf tissues.

The precise mechanisms by which foreign DNA sequences are recognised and inactivated remain obscure. The situation is complicated in that gene silencing can be due to transcriptional inactivation or to post-transcriptional events (Flavell, 1994; Matzke and Matzke, 1995). In other plant species, transcriptional inactivation is likewise frequently associated with specific methylation of the transgene, as seen in experiments with *Petunia hybrida* where a chimaeric maize *A1* gene was introduced into a white flowered strain to produce plants with brick-red flower colour (Meyer *et al.*, 1987). Variegation and loss of flower colour was observed among plants in a field trial, and inactivation of the transgene was correlated with increased cytosine methylation in the 35S promoter region derived from cauliflower mosaic virus (Linn *et al.*, 1990; Meyer and Heidmann, 1994). There may be multiple mechanisms by which plants are able to recognise and specifically inactivate exogenous DNA sequences, and these may involve specific DNA-DNA or DNA-RNA interactions (Matzke and Matzke, 1995).

A better understanding of transgene inactivation will help formulate strategies for improving transgene stability under field conditions (Finnegan and McElroy, 1994, 1996 [in the press]). These include selection for single copy transgene insertion events and development of site-specific recombination systems. The inclusion of matrix attachment regions with the introduced gene(s) may mitigate the effects of sequences adjacent to the site of integration (Allen *et al.*, 1993; Mlynárová *et al.*, 1994). Another strategy that is being developed is transposon-mediated delivery of a transgene. For example a gene of interest can be introduced into barley on the *Ds* transposable element from maize. In the presence of an active maize *Ac* element, transposition of single copies of the gene can be induced (D. McElroy, personal communication). Such a system has the dual advantages of providing a means of ensuring single copy integrations, while at the same time enabling the gene of interest to be separated from the selectable marker gene whose presence in the field may not be desirable.

Concluding remarks

The past five years have seen rapid advances in the application of gene transfer technology to cereals. Methods have been developed for genetic transformation of all major crop species, and information is becoming available about the behaviour of transgenes in transgenic cereal plants. Relative efficiencies of transformation appear to be greater with rice and maize than with wheat, barley and oats; however, this may to some extent reflect the amount of research effort expended on the different species.

While there are only a limited number of examples of genes other than selectable markers being expressed in transgenic cereals, the information to hand suggests that the constraints on stable expression are not fundamentally different to those observed in broad-leaved plant species such as tobacco for which transgenic plants have been available for more than a decade. A proportion of transgenes seem to be subject to inactivation, which may depend on position of integration in the genome, copy number and sequence organisation of the introduced DNA. Further refinement of the transformation technology will make it feasible to generate greater numbers of independent

transformants and select those that have the required levels of stable expression.

The agronomic performance of transgenic cereals is insufficiently researched, but will be an important consideration in future studies. By analogy with other species, in most cases it will be necessary to undertake a programme of back-crossing for the following reasons. There is wide genotypic variation in tissue culture response among the cereals, meaning that a particular method of transformation can be limited to certain cultivars or strains. Further, the accumulation of undesirable somaclonal variation (Larkin and Scowcroft, 1981) is associated with an extensive phase of tissue culture which is a feature of all current methods of cereal transformation. Thus, even if transformation can be achieved for the cultivar of interest, it is recommended that the primary transformant is crossed to parental seed grown material to remove any deleterious mutations.

The next decade will be decisive in determining the contribution gene transfer technology can make to crop improvement. The technology must be viewed as an adjunct to, rather than a replacement of, existing breeding methods. The precise tailoring of individual genes provides a powerful additional tool, but does not obviate the necessity of exhaustive testing of any new breeding material generated. The adoption of genetic engineering for cereals will also strongly depend on public acceptance of the technology, and demonstration that genetic modification can result in safe cereal products which provide, directly or indirectly, most of the nutrition for the world's population.

Acknowledgements

We would like to thank John Stead for critical reading of the manuscript. We are also grateful to Paul Lazzeri, Indra Vasil, David Somers and David McElroy for providing us with unpublished data for the preparation of this article.

References

ABDULLAH, R., COCKING, E.C., AND THOMPSON, J.A. (1986). Efficient plant regeneration from rice protoplasts through somatic embryogenesis. *Bio/Technology* **4**, 1087–1090.

AHMED, K.Z. AND SÁGI, F. (1993). Culture of and fertile plant regeneration from regenerable embryogenic suspension cell-derived protoplasts of wheat (*Triticum aestivum* L.). *Plant Cell Reports* **12**, 175–179.

AHOKAS, H. (1989). Transfection of germinating barley seed electrophoretically with exogenous DNA. *Theoretical and Applied Genetics* **77**, 469–472.

ALLEN, G.C., HALL, G.E., CHILDS, L.C., WEISSINGER, A.K., SPIKER, S., AND THOMPSON, W.F. (1993). Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. *Plant Cell* **5**, 603–613.

ANDERSON, O.D., BLECHL, A.E., GREENE, F.C. AND WEEKS, J.T. (1994). Progress towards genetic engineering of wheat with improved quality. In *Improvement of Cereal Quality by Genetic Engineering* (R.J. Henry and J.A. Ronalds, Eds), pp.87–95. Plenum Press, New York.

ARMSTRONG, C.L., PARKER, G.B., PERSHING, J.C., BROWN, S.M., SANDERS, P.R., DUNCAN, D.R., STONE, T., DEAN, D.A., DEBOER, D.L., HART, J., HOWE, A.R., MORRISH, F.M., PAJEAU, M.E., PETERSEN, W.L., REICH, B.J., RODRIGUEZ, R., SANTINO, C.G., SATO, S.J., SCHULER, W., SIMS, S.R., STEHLING, S., TAROCHIONE, L.J. AND FROMM, M.E. (1995). Field evaluation of European corn borer control in progeny of 173 transgenic corn events expressing an insecticidal protein from *Bacillus thuringiensis*. *Crop Science* **35**, 550–557.

ASSAAD, F., TUCKER, K.L. AND SIGNER, E.R. (1993). Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. *Plant Molecular Biology* **22**, 1067–1085.

BARCELO, P., HAGEL, C., BECKER, D., MARTIN, A. AND LÖRZ, H. (1994). Transgenic cereal (tritordeum) plants obtained at high efficiency by microprojectile bombardment of inflorescence tissue. *The Plant Journal* **5**, 583–592.

BAULCOMBE, D.C., CHAPMAN, S. AND CRUZ, S.S. (1995). Jellyfish green fluorescent protein as a reporter for virus infections. *The Plant Journal* **7**, 1045–1053.

BECKER, D., BRETTSCHNEIDER, R. AND LÖRZ, H. (1994). Fertile transgenic wheat from microprojectile bombardment of scutellar tissue. *The Plant Journal* **5**, 299–307.

BINNS, A.N. AND THOMASHOW, M.F. (1988). Cell biology of *Agrobacterium* infection and transformation of plants. *Annual Review of Microbiology* **42**, 575–606.

BIRCH, R.G. AND FRANKS, T. (1991). Development and optimisation of microprojectile systems for plant genetic transformation. *Australian Journal of Plant Physiology* **18**, 453–469.

BLECHL, A.E., WEEKS, T.W. AND ANDERSON, O.D. (1995). Expression of a novel high molecular weight glutenin subunit gene in transgenic wheat. *Proceedings of Conference on Value-Added Cereals Through Biotechnology*, Saskatoon, 10–13th June 1995, p.39.

BODEAU, J.P. AND WALBOT, V. (1995). Genetic regulation of anthocyanin biosynthesis in embryogenic maize callus. *Maydica* **40**, 77–83.

BOWER, R. AND BIRCH, R.G. (1992) Transgenic sugarcane plants via microprojectile bonbardment. *The Plant Journal* **2**, 409–416.

BRETTELL, R.I.S., WERNICKE, W. AND THOMAS, E. (1980). Embryogenesis from cultured immature inflorescences of *Sorghum bicolor*. *Protoplasma* **104**, 141–148.

BROWN, D.C.W., TIAN, L., BUCKLEY, D.J., LEFEBVRE, M., MCGRATH, A. AND WEBB, J. (1994). Development of a simple particle bombardment device for gene transfer into plant cells. *Plant Cell, Tissue and Organ Culture* **37**, 47–53.

CALLIS, J., FROMM, M. AND WALBOT, V. (1987). Introns increase gene expression in cultured maize cells. *Genes and Development* **1**, 1183–1200.

CAO, J., DUAN, X., MCELROY, D. AND WU, R. (1992). Regeneration of herbicide resistant transgenic rice plants following microparticle-mediated transformation of suspension culture cells. *Plant Cell Reports* **11**, 586–591.

CASAS, A.M., KONONOWICZ, A.K., ZEHR, U.B., TOMES, D.T., AXTELL, J.D., BUTLER, L.G., BRESSAN, R.A. AND HASEGAWA, P.M. (1993). Transgenic sorghum plants via microprojectile bombardment. *Proceedings of the National Academy of Sciences USA* **90**, 11212–11216.

CASTILLO, A.M., VASIL, V., AND VASIL, I.K. (1994). Rapid production of fertile transgenic plants of rye (*Secale cereale* L.). *Bio/Technology* **12**, 1366–1371.

CHAMBERLAIN, D. A., BRETTELL, R.I.S., LAST, D. I., WITRZENS, B., MCELROY, D., DOLFERUS, R. AND DENNIS, E. S. (1994). The use of the Emu promoter with antibiotic and herbicide resistance genes for the selection of transgenic wheat callus and rice plants. *Australian Journal of Plant Physiology* **21**, 95–112.

CHAN, M.T., LEE, T.M. AND CHANG, H.H. (1992). Transformation of indica rice (*Oryza sativa* L.) mediated by *Agrobacterium tumefaciens*. *Plant and Cell Physiology* **33**, 577–583.

CHAN, M.T., CHANG, H.H., HO, S.L., TONG, W.F. AND YU, S.M. (1993). *Agrobacterium*-mediated production of transgenic rice plants expressing a chimeric α-amylase promoter/β-glucuronidase gene. *Plant Molecular Biology* **22**, 491–506.

CHANG, Y.F., WANG, W.C., WARFIELD, C.Y., NGUYEN, H.T. AND WONG, J.R. (1991). Plant regeneration from protoplasts isolated from long-term cell cultures of wheat (*Triticum aestivum* L.). *Plant Cell Reports* **9**, 611–614.

CHIA, T-F., CHAN, Y-S. AND CHUA, N-H. (1994). The firefly luciferase gene as a non-invasive reporter for *Dendrobium* transformation. *The Plant Journal* **6**, 441–446.

CHRISTENSEN, A.H., SHARROCK, R.A. AND QUAIL, P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology* **18**, 675–689.

CHRISTOU, P., FORD, T.L. AND KOFRON, M. (1991). Production of transgenic rice (*Oryza sativa* L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Bio/Technology* **9**, 957–962.

COOLEY, J., FORD, T. AND CHRISTOU, P. (1995). Molecular and genetic characterization of elite transgenic rice plants produced by electric-discharge particle acceleration. *Theoretical and Applied Genetics* **90**, 97–104.

DALE, P.J., MARKS, M.S., BROWN, M.M., WOOLSTON, C.J., GUNN, H.V., MULLINEAUX, P.M., LEWIS, D.M., KEMP, J.M., CHEN, D.F., GILMOUR, D.M. AND FLAVELL, R.B. (1989). Agroinfection of wheat: inoculation of *in vitro* grown seedlings and embryos. *Plant Science* **63**, 237–245.

DALTON, S.J., BETTANY, A.J.E., TIMMS, E. AND MORRIS, P. (1995). The effect of selection pressure on transformation frequency and copy number in transgenic plants of tall fescue (*Festuca arundinacea* Schreb). *Plant Science* **108**, 63–70.

DATTA, S.K., DATTA, K., SOLTANIFAR, N., DONN, G. AND POTRYKUS, I. (1992). Herbicide-resistant Indica rice plants from IRRI breeding line IR72 after PEG-mediated transformation of protoplasts. *Plant Molecular Biology* **20**, 619–629.

DATTA, S.K., PETERHANS, A., DATTA, K. AND POTRYKUS, I. (1990). Genetically engineered fertile Indica-rice recovered from protoplasts. *Bio/Technology* **8**, 736–740.

DE LA PENA, A., LÖRZ, H., AND SCHELL, J. (1987). Transgenic plants obtained by injecting DNA into young floral tillers. *Nature* **325**, 274–276.

DE WET, J.R., WOOD, K.V. DELUCA, M., HELINSKI, D.R. AND SUBRAMANI, S. (1987). Firefly luciferase gene: structure and expression in mammalian cells. *Molecular and Cellular Biology* **7**, 725–737.

DENG, W.Y., LIN, X.Y., AND SHAO, Q.Q. (1990). *Agrobacterium tumefaciens* can transform *Triticum aestivum* and *Hordeum vulgare* of Gramineae. *Science in China* (Series B) **33**, 27–34.

DENNEHEY, B.K., PETERSEN, W.L., FORD-SANTINO, C., PAJEAU, M. AND ARMSTRONG, C.L. (1994). Comparison of selective agents for use with the selectable marker gene *bar* in maize transformation. *Plant Cell, Tissue and Organ Culture* **36**, 1–7.

D'HALLUIN, K., BONNE, E., BOSSUT, M., DE BEUCKELEER, M. AND LEEMANS, J. (1992). Transgenic maize plants by tissue electroporation. *The Plant Cell* **4**, 1495–1505.

ELLIS, J.G., LLEWELLYN, D.J. DENNIS, E.S. AND PEACOCK, W.J. (1987). Maize *Adh-1* promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. *The EMBO Journal* **6**, 11–16.

FINNEGAN, J. AND MCELROY, D. (1994) Transgene inactivation: plants fight back! *Bio/Technology* **12**, 883–888.

FINNEGAN, J. AND MCELROY, D. (1996) Transgene stability. In *Transgenic Plants: A Production System for Industrial and Pharmaceutical Proteins* (M.R.L. Owen and J. Pen, Eds) in the press. John Wiley & Sons, Chichester.

FINER, J.J., VAIN, P., JONES, M.W. AND MCMULLEN, M.D. (1992). Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Reports* **11**, 323–328.

FINNEGAN, E.J., BRETTELL, R.I.S. AND DENNIS, E.S. (1993). The role of DNA methylation in the regulation of plant gene expression. In *DNA Methylation: Molecular Biology and Biological Significance*. (J. P. Jost and H. P. Saluz, Eds), pp. 218–261. Birkhäuser Verlag, Basle, Switzerland.

FLAVELL, R.B. (1994). Inactivation of gene expression in plants as a consequence of novel sequence duplications. *Proceedings of the National Academy of Sciences, USA* **91**, 3490–3496.

FRAME, B.R., DRAYTON, P.R., BAGNALL, S.V., LEWNAU, C.J., BULLOCK, W.P., WILSON, H.M., DUNWELL, J.M., THOMPSON, J.A. AND WANG, K. (1994). Production of fertile transgenic maize plants by silicon carbide whisker-mediated transformation. *The Plant Journal* **6**, 941–948.

FROMM, M.E., MORRISH, F., ARMSTRONG, C., WILLIAMS, R., THOMAS, J. AND KLEIN, T.M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Bio/Technology* **8**, 833–839.

FROMM, M., TAYLOR, L.P. AND WALBOT, V. (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proceedings of the National Academy of Sciences, USA* **82**, 5824–5828.

FUJIMOTO, H., ITOH, K., YAMAMOTO, M., KYOZUKA, J., AND SHIMAMOTO, K. (1993). Insect resistant rice generated by introduction of a modified δ-endotoxin gene of *Bacillus thuringiensis*. *Bio/Technology* **11**, 1151–1155.

FUNATSUKI, H., KURODA, H., KIHARA, M., LAZZERI, P.A., MÜLLER, E., LÖRZ, H. AND KISHINAMI, I. (1995). Fertile transgenic barley generated by direct DNA transfer to

protoplasts. *Theoretical and Applied Genetics* **91**, 707–712.

GOFF, S.A., KLEIN, T.M., ROTH, B.A., FROMM, M.E., CONE, K.C., RADICELLA, J.P. AND CHANDLER, V.L. (1990). Trans-activation of anthocyanin biosynthetic genes following transfer of *B* regulatory genes into maize tissues. *The EMBO Journal* **9**, 2517–2522.

GOLOVKIN, M.V., ABRAHAM, M., MOROCZ, S., BOTTKA, S., FEHER, A. AND DUDITS, D. (1993). Production of transgenic maize plants by direct DNA uptake into embryogenic proroplasts. *Plant Science* **90**, 41–52.

GORDON-KAMM, W.J., SPENCER, T.M., MANGANO, M.L., ADAMS, T.R., DAINES, R.J., START, W.G., O'BRIEN, J.V., CHAMBERS, S.A., ADAMS, W.R., WILLETTS, N.G., RICE, T.B., MACKEY, C.J., KRUEGER, R.W., KAUSCH, A.P. AND LEMAUX, P.G. (1990). Transformation of maize cells and regeneration of fertile transgenic plants. *The Plant Cell* **2**, 603–618.

GOULD, J., DEVEY, M., HASEGAWA, O., ULIAN, E.C., PETERSON, G. AND SMITH, R.H. (1991). Transformation of *Zea mays* L. using *Agrobacterium tumefaciens* and the shoot apex. *Plant Physiology* **95**, 426–434.

GRIMSLEY, N.H., HOHN, T., DAVIES, J.W. AND HOHN, B. (1987). *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. *Nature* **325**, 177–179.

GUO, Y., LIANG, H. AND BERNS, M.W. (1995). Laser-mediated gene transfer in rice. *Physiologia Plantarum* **93**, 19–24.

HAGIO, T., HIRABAYASHI, T., MACHII, H. AND TOMOTSUNE, H. (1995). Production of fertile transgenic barley (*Hordeum vulgare* L.) plant using the hygromycin-resistance marker. *Plant Cell Reports* **14**, 329–334.

HARRIS, R., WRIGHT, M., BYRNE, M., VARNUM, J., BRIGHTWELL, B. AND SCHUBERT, K. (1988). Callus formation and plantlet regeneration from protoplasts derived from suspension cultures of wheat (*Triticum aestivum* L.). *Plant Cell Reports* **7**, 337–340.

HE, D.G., YANG, Y.M. AND SCOTT, K.J. (1992). Plant regeneration from protoplasts of wheat (*Triticum aestivum* cv. Hartog). *Plant Cell Reports* **11**, 16–19.

HESS, D., DRESSLER, K., AND NIMMRICHTER, R. (1990). Transformation experiments by pipetting *Agrobacterium* into the spikelets of wheat (*Triticum aestivum* L.). *Plant Science* **72**, 233–244.

HIEI, Y., OHTA, S., KOMARI, T., AND KUMASHIRO, T. (1994). Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal* **6**, 271–282.

HOBBS, S.L.A., KPODAR, P. AND DELONG, C.M.O. (1990). The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. *Plant Molecular Biology* **15**, 851–864.

HUNOLD, R., BRONNER, R., AND HAHNE, G. (1994). Early events in microprojectile bombardment: cell viability and particle location. *The Plant Journal* **5**, 593–604.

ITOH, K., YAMAGUCHI, J., HUANG, N., RODRIGUEZ, R.L., AKAZAWA, T. AND SHIMAMOTO, K. (1995). Developmental and hormonal regulation of rice α-amylase (*RAmy1A*)-*gusA* fusion genes in transgenic rice seeds. *Plant Physiology* **107**, 25–31.

JÄHNE, A., BECKER, D., BRETTSCHNEIDER, R. AND LÖRZ, H. (1994). Regeneration of transgenic, microspore-derived, fertile barley. *Theoretical and Applied Genetics* **89**, 525–533.

JÄHNE, A., LAZZERI, P.A. AND LÖRZ, H. (1991). Regeneration of fertile plants from protoplasts derived from embryogenic cell suspensions of barley (*Hordeum vulgare* L.). *Plant Cell Reports* **10**, 1–6.

JEFFERSON, R.A., KAVANAGH, T.A. AND BEVAN, M.W. (1987). GUS fusions: B-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal* **6**, 3901–3907.

KAEPPLER, H.F., GU, W., SOMERS, D.A., RINES, H.W. AND COCKBURN, A.F. (1990). Silicon carbide fiber-mediated DNA delivery into plant cells. *Plant Cell Reports* **8**, 415–418.

KALLA, R., SHIMAMOTO, K., POTTER, R., NIELSEN, P.S., LINNESTAD, C. AND OLSEN, O.A. (1994). The promoter of the barley aleurone-specific gene encoding a putative 7 Kda lipid transfer protein confers aleurone cell-specific expression in transgenic rice. *The Plant Journal* **6**, 849–860.

KLÖTI, A., IGLESIAS, V.A., WÜNN, J., BURKHARDT, P.K., DATTA, S.K. AND POTRYKUS, I. (1993). Gene transfer by electroporation into intact scutellum cells of wheat embryos. *Plant Cell Reports* **12**, 671–675.

KONCZ, C., MARTINI, N., MAYERHOFER, R., KONCZ-KALMAN, Z., KÖRBER, H., REDEI, G.P.

AND SCHELL, J. (1989). High-frequency T-DNA-mediated gene tagging in plants. *Proceedings of the National Academy of Sciences, USA* **86**, 8467–8471.

KOZIEL, M.G., BELAND, G.L., BOWMAN, C., CAROZZI, N.B., CRENSHAW, R., CROSSLAND, L., DAWSON, J., DESAI. N., HILL, M., KADWELL, S., LAUNIS, K., LEWIS, K., MADDOX, D., MCPHERSON, K., MEGHJI, M.R., MERLIN, E., RHODES, R., WARREN, G.W., WRIGHT, M., AND EVOLA, S.V. (1993). Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Bio/Technology* **11**, 194–200

KRENS, F. A., MOLENDIJK, L., WULLEMS, G. J. AND SCHILPEROORT, R. A. (1982). *In vitro* transformation of plant protoplasts with Ti-plasmid DNA. *Nature* **296**, 72–74.

KYOZUKA, J., HAYASHI, Y. AND SHIMAMOTO, K. (1987) High frequency plant regeneration from rice protoplasts by novel nurse culture methods. *Molecular and General Genetics* **206**, 408–413.

KYOZUKA, J., MCELROY, D., HAYAKAWA, T., XIE, Y., WU, R. AND SHIMAMOTO, K. (1993). Light-regulated and cell-specific expression of tomato *rbcS-gusA* and rice *rbcS-gusA* fusion genes in transgenic rice. *Plant Physiology* **102**, 991–1000.

KYOZUKA, J., OLIVE, M., PEACOCK, W.J., DENNIS, E.S. AND SHIMAMOTO, K. (1994). Promoter elements required for developmental expression of the maize*Adh1* gene in transgenic rice. *The Plant Cell* **6**, 799–810.

LAMBÉ, P., DINANT, M., AND MATAGNE, R.F. (1995). Differential long-term expression and methylation of the hygromycin phosphotransferase (hph) and β-glucuronidase (GUS) genes in transgenic pearl millet (*Pennisetum glaucum*) callus. *Plant Science* **108**, 51–62.

LANGRIDGE, P., BRETTSCHNEIDER, R., LAZZERI, P. AND LÖRZ, H. (1992). Transformation of cereals via *Agrobacterium* and the pollen pathway: a critical assessment. *The Plant Journal* **2**, 631–638.

LARKIN, P.J. AND SCOWCROFT, W.R. (1981). Somaclonal variation – a novel source of variability from cell cultures for plant improvement. *Theoretical and Applied Genetics* **60**, 197–214.

LARKIN, P.J., TAYLOR, B.H., GERSMANN, M. AND BRETTELL, R.I.S. (1990). Direct gene transfer to protoplasts. *Australian Journal of Plant Physiology* **17**, 291–302.

LAST, D.I., BRETTELL, R.I.S., CHAMBERLAIN, D.A., CHAUDHURY, A.M., LARKIN, P.J., MARSH, E.L., PEACOCK, W.J. AND DENNIS, E.S. (1991). pEmu: an improved promoter for gene expression in cereal cells. *Theoretical and Applied Genetics* **81**, 581–588.

LAURSEN, C.M., KRZYZEK, R.A., FLICK, C.E., ANDERSON, P.C. AND SPENCER, T.M. (1994). Production of fertile transgenic maize by electroporation of suspension culture cells. *Plant Molecular Biology* **24**, 51–621.

LAZZERI, P.A. AND SHEWRY, P.R. (1993). Biotechnology of cereals. *Biotechnology and Genetic Engineering Reviews* **11**, 79–146.

LI, B.J., XU, X.P., SHI, H.P. AND KE, X.Y. (1991). Introduction of foreign genes into the seed embryo cells of rice by electro-injection and the regeneration of transgenic rice plants. *Science in China* **34**, 923–930.

LI, L., QU, R., DE KOCHKO, A., FAUQUET, C. AND BEACHY, R.N. (1993) An improved rice transformation system using the biolistic method. *Plant Cell Reports* **12**, 250–255.

LI, Y.Z., MA, H.M., ZHANG, J.L., WANG, Z.Y. AND HONG, M.M. (1995). Effects of the first intron of rice waxy gene on the expression of foreign genes in rice and tobacco protoplasts. *Plant Science* **108**, 181–190.

LI, Z.Y., XIA, G.M., CHEN, H.M. AND GUO, G.Q. (1992). Plant regeneration from protoplasts derived from embryogenesis (sic) suspension cultures of wheat (*Triticum aestivum* L.). *Journal of Plant Physiology* **139**, 714–718.

LIN, W., ANURATHA, C.S., DATTA, K., POTRYKUS, I., MUTHUKRISHNAN, S., AND DATTA, S.K. (1995). Genetic engineering of rice for resistance to sheath blight. *Bio/Technology* **13**, 686–691.

LINN, F., HEIDMANN, I., SAEDLER, H. AND MEYER, P. (1990). Epigenetic changes in the expression of the maize *A1* gene in *Petunia hybrida*: role of numbers of integrated gene copies and state of methylation. *Molecular and General Genetics* **222**, 329–336.

LOWE, K., BOWEN, B., HOERSTER, G., ROSS, M., BOND, D., PIERCE, D. AND GORDON-KAMM, W. (1995). Germline transformation of maize following manipulation of chimeric shoot meistems. *Bio/Technology* **13**, 677–682.

LUEHRSEN, K.R. AND WALBOT, V. (1991). Intron enhancement of gene expression and the splicing efficiency of introns in maize cells. *Molecular and General Genetics* **225**, 81–93.

LUO, Z AND WU, R. (1988) A simple method for transformation of rice via the pollen-tube pathway. *Plant Molecular Biology Reporter* **6**, 165–174.

MCCABE, D.E., SWAIN, W.F., MARTINELL, B.J. AND CHRISTOU, P. (1988). Stable transformation of soybean (*Glycine max*) by particle accelerator. *Bio/Technology* **6**, 923–926.

MCELROY, D., BLOWERS, A.D., JENES, B. AND WU, R. (1991). Construction of expression vectors based on the rice actin 1 (*Act1*) 5' region for use in monocot transformation. *Molecular and General Genetics* **231**, 150–160.

MCELROY, D. AND BRETTELL, R.I.S. (1995). Foreign gene expression in transgenic cereals. *Trends in Biotechnology* **12**, 62–68.

MCELROY, D. AND JACOBSEN, J. (1995). What's brewing in barley biotechnology? *Bio/Technology* **13**, 245–249.

MAAS, C. AND WERR, W. (1989). Mechanism and optimized conditions for PEG mediated DNA transfection into plant protoplasts. *Plant Cell Reports* **8**, 148–151.

MATSUKI, R., ONODERA, H., YAMAGUCHI, T. AND UCHIMIYA, H. (1989). Tissue-specific expression of the *rolC* promoter of the Ri plasmid in transgenic rice plants. *Molecular and General Genetics* **220**, 12–16.

MATZKE, M.A. AND MATZKE, A.J.M. (1995). How and why do plants inactivate homologous (trans)genes? *Plant Physiology* **107**, 679–685.

MEYER. P., HEIDMANN, I., FORKMANN, G. AND SAEDLER, H. (1987). A new petunia flower colour generated by transformation of a mutant with a maize gene. *Nature* **330**, 677–678.

MEYER, P AND HEIDMANN, I. (1994). Epigenetic variants of a transgenic petunia line show hypermethylation in transgene DNA: an indication for specific recognition of foreign DNA in transgenic plants. *Molecular and General Genetics* **243**, 390–399.

MLYNÁROVÁ, L., LOONEN, A., HELDENS, J., JANSEN, R.C., KEIZER, P., STIEKEMA, W.J. AND NAP, J-P. (1994). Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. *The Plant Cell* **6**, 417–422

MOONEY, P.A., GOODWIN, P.B., DENNIS, E.S. AND LLEWELLYN, D.J. (1991). *Agrobacterium tumefaciens*-gene transfer into wheat tissues. *Plant Cell, Tissue and Organ Culture* **25**, 209–218.

NABULSI, S.M., PAGE, N.W., DUVAL, A.L., SEABROOK, Y.A. AND SCOTT, K.J. (1994). A gas-driven gene gun for microprojectile methods of genetic engineering. *Meas. Sci. Technol.* **5**, 267–274

NAGATA, T. (1989). Cell biological aspects of gene delivery into plant protoplasts by electroporation. *International Review of Cytology* **116**, 229–255.

NAGATA, T. AND TAKEBE, I. (1971). Plating of isolated mesophyll protoplasts on agar medium. *Planta* (Berlin) **99**, 12–20.

NEHRA, N., CHIBBAR, R.N., LEUNG, N., CASWELL, K., MALLARD, C., STEINHAUER, L., BAGA, M., AND KARTHA, K.K. (1994). Self-fertile transgenic wheat plants regenerated from isolated scutellar tissues following microprojectile bombardment with two distinct gene constructs. *The Plant Journal* **5**, 285–297.

NIEDZ, R.P., SUSSMAN, M.R. AND SATTERLEE, J.S. (1995). Green fluorescent protein – an in vivo reporter of plant gene expression. *Plant Cell Reports* **14**, 403–406.

OZIAS-AKINS, P. AND VASIL, I.K. (1982). Plant regeneration from cultured immature embryos and inflorescences of *Triticum aestivum* L. (Wheat): Evidence for somatic embryogenesis. *Protoplasma* **110**, 95–105

PAUK, J., KERTÉSZ, Z., JENES, B., PURNHAUSER, L., MANNINEN, O., PULLI, S., BARABÁS, Z. AND DUDITS, D. (1994). Fertile wheat (*Triticum aestivum* L.) regenerants from protoplasts of embryogenic suspension culture. *Plant Cell, Tissue and Organ Culture* **38**, 1–10.

PENG, J., LYZNIK, L.A., LEE, L. AND HODGES, T.K. (1990). Co-transformation of indica rice protoplasts with *gusA* and *neo* genes. *Plant Cell Reports* **9**, 168–172.

PENG, J.Y., WEN, F.J., LISTER, R.L. AND HODGES, T.K. (1995). Inheritance of *gusA* and *neo* genes in transgenic rice. *Plant Molecular Biology* **27**, 91–104.

POTRYKUS, I. (1990). Gene transfer to cereals: an assessment. *Bio/Technology* **8**, 535–542.

POTRYKUS, I., SAUL, M.W., PETRUSKA, J., PASZKOWSKI, J. AND SHILLITO, R.D. (1985). Direct gene transfer to cells of a graminaceous monocot. *Molecular and General Genetics* **199**, 183–188.

QIAO, Y.M., CATTANEO, M., LOCATELLI, F. AND LUPOTTO, E. (1992). Plant regeneration from long term suspension culture-derived protoplasts of hexaploid wheat (*Triticum aestivum* L.). *Plant Cell Reports* **11**, 262–265.

RAINERI, D.M., BOTTINO, P., GORDON, M.P. AND NESTER, E.W. (1990). *Agrobacterium*-mediated transformation of rice. (*Oryza sativa* L.) *Bio/Technology* **8**, 33–37.

RANGAN, T.S. AND VASIL I.K. (1983). Somatic embryogenesis and plant regeneration in tissue cultures of *Panicum miliaceum* L. and *Panicum miliare* Lamk. *Zeitschrift fiir Pflanzenphysiologie* **109**, 49–53.

REGISTER III, J.C., PETERSON, D.J., BELL, P.J., BULLOCK, W.P., EVANS, I.J., FRAME, B., GREENLAND, A.J., HIGGS, N.S., JEPSON, I., JIAO, S., LEWNAU, C.J., SILLICK, J.M. AND WILSON, H.M. (1994). Structure and function of selectable and non-selectable transgenes in maize after introduction by particle bombardment. *Plant Molecular Biology* **25**, 951–961.

REN, J.G., JIA, J.F., LI, M.Y. AND ZHEN, G.C. (1989). Plantlet regeneration from protoplasts isolated from callus cultures of immature inflorescences of wheat (*Triticum aestivum* L.). *Chinese Science Bulletin* **34**, 1648–1652.

RHODES, C. A., PIERCE, D.A., METTLER, I. J., MASCARENHAS, D. AND DETMER, J.J. (1988). Genetically transformed maize plants from protoplasts. *Science* **240**, 204–207.

RITALA, A., ASPEGREN, K., KURTÉN, U., SALMENKALLIO-MARTTILA, M., MANNONEN, L., HANNUS, R., KAUPPINEN, V., TEERI, T.H. AND ENARI, T-M. (1994). Fertile transgenic barley by particle bombardment of immature embryos. *Plant Molecular Biology* **24**, 317–325.

RYSCHKA, S., RYSCHKA, U. AND SCHULZE, J. (1991). Anatomical studies on the development of somatic embryoids in wheat and barley explants. *Biochem. Physiol. Pflanz.* **187**, 31–41.

SANFORD, J.C., KLEIN, T.M., WOLF, E.D. AND ALLEN, N. (1987). Delivery of substances into cells and tissues using a particle bombardment process. *Particulate Science and Technology* **5**, 27–37.

SANFORD, J.C., DEVIT, M.J., RUSSELL, J.A., SMITH, F.D., HARPENDING, P.R., ROY, M.K. AND JOHNSON, S.A. (1991). An improved, helium-driven biolistic device. *Technique* **3**, 3–16.

SCHLEDZEWSKI, K. AND MENDEL, R.R. (1994). Quantitative transient gene expression: comparison of the promoters for maize polyubiquitin 1, rice actin 1, maize-derived *Emu* and CaMV 35S in cells of barley, maize and tobacco. *Transgenic Research* **3**, 249–255.

SCHUH, W., NELSON, M.R., BIGELOW, D.M., ORUM, T.V., ORTH, C.E., LYNCH, P.T., EYLES, P.S., BLACKHALL, N.W., JONES, J., COCKING, E.C. AND DAVEY, M.R. (1993) The phenotypic characterization of R_2 generation transgenic rice plants under field conditions. *Plant Science* **89**, 69–79.

SHIMAMOTO, K., MIYAZAKI, C., HASHIMOTO, H., IZAWA, T., ITOH, K., TERADA, R., INAGAKI, Y. AND IIDA, S. (1993). *Trans*-activation and stable integration of the maize transposable element *Ds* cotransfected with the *Ac* transposase gene in transgenic rice plants. *Molecular and General Genetics* **239**, 354–360.

SHIMAMOTO, K., TERADA, R., IZAWA, T. AND FUJIMOTO, H. (1989). Fertile rice plants regenerated from transformed protoplasts. *Nature* **338**, 274–276.

SOMERS, D. A., RINES, H.W., GU, W., KAEPPLER, H. F. AND BUSHNELL, W. R. (1992). Fertile, transgenic oat plants. *Bio/Technology* **10**, 1589–1594.

SOMERS, D.A., TORBERT, K.A., PAWLOWSKI, W.P. AND RINES, H.W. (1994). Genetic engineering of oat. In *Improvement of Cereal Quality by Genetic Engineering* (R.J. Henry and J.A. Ronalds, Eds), pp.37–46. Plenum Press, New York.

SONGSTAD, D.D., SOMERS, D.A. AND GRIESBACH, R.J. (1995). Advances in alternative DNA delivery techniques. *Plant Cell, Tissue and Organ Culture* **40**, 1–15.

SPENCER, T.M., O'BRIEN, J.V., START, W.G., ADAMS, T.R., GORDON-KAMM, W.J. AND LEMAUX, P.G. (1992). Segregation of transgenes in maize. *Plant Molecular Biology* **18**, 201–210.

STASKAWICZ, B.J., AUSUBEL, F.M., BAKER, B.J., ELLIS, J.G. AND JONES, J.D.G. (1995). Molecular genetics of plant disease resistance. *Science* **268**, 661–667.

SUKHAPINDA, K., KOZUCH, M.E., RUBIN-WILSON, B., AINLEY, W.M. AND MERLO, D.J. (1993). Transformation of maize (*Zea mays* L.) protoplasts and regeneration of haploid transgenic plants. *Plant Cell Reports* **13**, 63–68.

TERADA, R., NAKAYAMA, T., IWABUCHI, M. AND SHIMAMOTO, K. (1995). A type I element composed of the hexamer (ACGTCA) and octamer (CGCGGATC) motifs plays a role(s)

in meristematic expression of a wheat histone H3 gene in transgenic rice plants. *Plant Molecular Biology* **27**, 17–26.

TORBERT, K.A., RINES, H.W. AND SOMERS, D.A. (1995). Use of paromomycin as a selective agent for oat transformation. *Plant Cell Reports* **14**, 635–640.

TORIYAMA, K., ARIMOTO, Y., UCHIMIYA, H. AND HINATA, K. (1988). Transgenic rice plants after direct gene transfer into protoplasts. *Bio/Technology* **6**, 1072–1074.

UCHIMIYA, H., IWATA, M., NOJIRI, C., SAMARAJEEWA, P.K., TAKAMATSU, S., OOBA, S., ANZAI, H, CHRISTENSEN, A.H., QUAIL, P.H. AND TOKI, S. (1993). Bialaphos treatment of transgenic rice plants expressing a *bar* gene prevents infection by the sheath blight pathogen (*Rhizoctonia solani*). *Bio/Technology* **11**, 835–836.

VASIL, V., REDWAY, F. AND VASIL, I.K. (1990). Regeneration of plants from embryogenic suspension culture protoplasts of wheat (*Triticum aestivum* L.) *Bio/Technology* **8**, 429–434.

VASIL, V., BROWN, S.M., RE, D., FROMM, M.E. AND VASIL, I.K. (1991). Stably transformed callus lines from microprojectile bombardment cell suspension cultures of wheat. *Bio/Technology* **9**, 743–747.

VASIL, V., CASTILLO, A.M., FROMM, M.E. AND VASIL, I.K. (1992). Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology* **10**, 667–674.

VASIL, V., SRIVASTAVA, V., CASTILLO, A.M., FROMM, M.E. AND VASIL, I. K. (1993). Rapid production of transgenic wheat plants by direct bombardment of cultured immature embryos. *Bio/Technology* **11**, 1553–1559.

WAN, Y. AND LEMAUX, P.G. (1994). Generation of large numbers of independently transformed fertile barley plants. *Plant Physiology* **104**, 37–48.

WAN, Y., WIDHOLM, J.M. AND LEMAUX, P.G. (1995). Type I callus as a bombardment target for generating fertile transgenic maize (*Zea mays* L.). *Planta* **196**, 7–14.

WANG, G.L., HOLSTEN, T.E., SONG, W.Y., WANG, H.P. AND RONALD, P.C. (1995). Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus. *The Plant Journal* **7**, 525–533.

WANG, K., DRAYTON, P., FRAME, B., DUNWELL, J., AND THOMPSON, J. (1995). Whisker-mediated plant transformation – an alternative terchnology. *In Vitro Cellular and Developmental Biology – Plant* **31**, 101–104.

WEEKS, J. T., ANDERSON, O. D. AND BLECHL, A. E. (1993). Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). *Plant Physiology* **102**, 1077–1084.

WILMINK, A. AND DONS, J. J. M. (1993). Selective agents and marker genes for use in transformation of monocotyledonous plants. *Plant Molecular Biology Reporter* **11**, 165–185.

YAMADA, Y., YANG, Z.Q. AND TANG, D.T. (1986). Plant regeneration from protoplast-derived callus or rice (*Oryza sativa* L.) *Plant Cell Reports* **4**, 85–88.

YAMAGUCHI-SHINOZAKI, K., MINO, M., MUNDY, J., AND CHUA, N.H. (1990). Analysis of an ABA-responsive rice gene promoter in transgenic tobacco. *Plant Molecular Biology* **15**, 905–912.

YAMASHITA, T., IIDA, A. AND MORIKAWA, H. (1991). Evidence that more than 90% of β-glucuronidase-expressing cells after particle bombardment directly receive the foreign gene in their nucleus. *Plant Physiology* **97**, 829–831.

ZAGHMOUT, O.M.F. AND TROLINDER, N.L. (1993). Simple and efficient method for directly electroporating *Agrobacterium* plasmid DNA into wheat callus cells. *Nucleic Acids Research* **21**, 1048.

ZHANG, H.M., YANG, H., RECH, E.L., GOLDS, T.J., DAVIS, A.S., MULLIGAN, B.J., COCKING, E.C. AND DAVEY, M.R. (1988). Transgenic rice plants produced by electroporation-mediated plasmid uptake into protoplasts. *Plant Cell Reports* **7**, 379–384.

ZHANG, W., MCELROY, D. AND WU, R. (1991). Analysis of rice *Act1* 5' region activity in transgenic rice plants. *Plant Cell* **3**, 1155–1165.

ZHOU, H., STIFF, C.M. AND KONZAK, C.F. (1993). Stably transformed callus of wheat by electroporation-induced direct gene transfer. *Plant Cell Reports* **12**, 612–616.

ZIMNY, J., BECKER, D., BRETTSCHNEIDER, R. AND LÖRZ, H. (1995). Fertile transgenic triticale (x Triticosecale Wittmack). *Molecular Breeding* **1**, 155–164.

reviews

Approaches to evaluating the transgenic status of transformed plants

James C. Register III

Analysis of transgenic plants is used today to answer the same questions as ten years ago. Since then, a wide variety of new molecular and biochemical technologies, as well as genetic tools, have become available. While there are many technical manuals describing how to use these tools, much less has been written about when they should and should not, be employed. This gap is important, as poor analysis can easily sidetrack or derail a project. This review addresses appropriate and inappropriate applications of several technologies commonly used for molecular and biochemical analysis of transformed plants.

The field of plant genetic engineering has evolved rapidly over the past decade. Despite the extent of this change, the need to determine whether or not transformed material is transgenic remains. While this issue may seem straightforward, the frequency with which insufficient or inappropriate data are used to support claims of transformation demonstrates that obtaining clear answers is not as trivial as it might seem. Although there are numerous manuals detailing how to carry out essentially any molecular technique, far less attention has been paid to describing the types of analytical approaches that provide appropriate data for answering a variety of commonly asked questions, including whether transformation of a plant has been achieved. This review will address appropriate and inappropriate applications of several technologies commonly used for molecular and biochemical analysis of nucleic acids and proteins in transformed plants.

Definitions

Before proceeding, a few definitions should be established for the purposes of this review. Transform(ed) will be used to refer to the act of introducing DNA into plant cells as well as to plant cells after DNA introduction. Transgenic will not be used interchangeably with transformed but will be used to refer to plant cells in which the DNA of interest (the transgene) is integrated into the host genome. Regenerated transgenic plants will be referred to as the T_0 generation. Plant transcription unit (PTU) refers to all the vector components that contribute to transgene expression (i.e. promoter, coding sequence, polyadenylation site and any other regulatory sequences employed). Because it is not uncommon for several copies of a transgene to insert at the same site (regardless of the transformation approach; e.g. Refs 1–4), copy number and number of insertion sites of a transgene are not the same and these terms will not be used interchangeably.

J. C. Register III (register@phibred.com) is at Pioneer Hi-Bred International Inc., Trait and Technology Development, PO Box 1004, Johnston, IA 50131-1004, USA.

Analytical approaches

Molecular and biochemical analysis of transformed plant material is done for two reasons. First, to determine whether the material is transgenic and, second, to characterize the material (e.g. determine copy number and/or the complexity of the DNA insert, and evaluate transgene expression). A wide variety of analytical tools are available for routine transgenic plant analysis. Some of the commonly used technologies and applications are listed in Table 1.

Transgene presence and structure

Polymerase chain reaction

PCR technology (for reviews see Refs 5,6) is perhaps the most popular method for screening material for the presence or absence of transgene sequences. It is particularly useful for rapid analysis of large numbers of samples. Despite its popularity, PCR has several inherent characteristics that limit the conclusions that can be drawn. First, a positive PCR result shows only that sequences homologous to the primers exist in a sample in close enough proximity for a product to be made. The results do not indicate whether the template DNA source is the intended sample material or

Table 1. Commonly used analytical technologies and appropriate applications

Technique	Applications
PCR	Rapid screening of transformants to focus subsequent analysis Identifying transgenics in segregating populations
Quantitative PCR	Estimating copy number (total) Hemizygote/homozygote discrimination (potentially)
Southern hybridization	Screening for presence/absence of specific sequences Assessing complexity of integration/transgene rearrangements Estimating copy number (rearranged and unrearranged) Estimating whether transgenes inserted at single or multiple loci 'Fingerprinting' identity of independent events
Reverse transcription PCR	Transgene expression: accumulation of specific mRNA – can be (semi-)quantitative and used for determining relative levels
Northern hybridization[a]	Transgene expression: quantifying specific mRNA accumulation Estimating mRNA size
ELISA[b]	Transgene expression: rapid screening for presence/absence of specific protein Can be very high throughput Quantifying specific protein accumulation
Western immunoblot	Transgene expression: screening for presence/absence of specific protein Approximate quantitation of specific protein Estimating protein size, evaluating protein processing

[a]The ribonuclease protection assay can be used in many cases to obtain analogous results. Compared with northern hybridization, it is generally regarded as more sensitive and offers the potential for greater throughput but it cannot be used for estimating RNA sizes.
[b]Western immunoassays of protein dot blots can provide some of the advantages of an ELISA as compared with traditional western immunoassays of electrophoresed proteins (e.g. increased throughput), potentially with reduced assay development requirements.

due to contamination, nor whether the template DNA is integrated into the plant genome. Information about the sequence outside the primers is also not provided. When products are analyzed using electrophoresis, verification of product size can be made. Southern analysis of the products can confirm that the sequence in between the primers is as expected, but this step is time consuming and thus rarely carried out. Quantitative PCR (for reviews see Refs 7,8) opens up the opportunity to use this technology for providing information historically obtained using Southern hybridization, such as estimates of copy number. The requirements for synthesizing standards, and use of additional standards and controls[9,10], as well as the fact that Southern hybridization analysis provides some information that PCR does not, have kept this approach from being used routinely.

PCR as used for transgenic plant analysis has different requirements than PCR used for cloning. Foremost is the robustness required of the reactions – a 10% failure rate is unacceptable for analytical purposes. Several steps, in addition to those routinely recommended for all PCR applications, can be taken to maximize robustness when setting up and monitoring reactions during analysis. For example, keeping reaction products less than ~1000 bp in length, using relatively long primers (25–30 nucleotides) and using primers with a high T_m (~65–68°C) can help ensure robustness and specificity. Ensuring that primers are sufficiently sensitive for reproducible detection of single copy sequences is crucial, and primer location affects the utility of the information provided by PCR (Fig. 1). Finally, during analysis, amplification of a known endogenous single copy gene, in addition to the transgene, provides a useful control for monitoring template DNA quality.

The extreme sensitivity of PCR results in the potential for contamination and false positive results. A variety of approaches to minimizing PCR contamination, both physical and biochemical, have been described[5,11–13]. In my experience, attention to primer design (Fig. 1), experimental design and logistics (how and where the work is carried out) is sufficient to minimize the occurrence of false positive results.

Southern hybridization

Southern analysis (for reviews see Refs 5,14,15) is one of the more powerful tools available for molecular characterization of transgenic plants. Depending

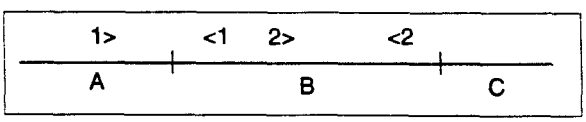


Figure 1
PCR primer location. The drawing shows a generic plant transcription unit in which A, B and C represent the promoter, coding sequence and polyadenylation site, respectively. They may also be thought of as any other juxtaposed DNA sequences in a vector. In this example, primers (1> and <1) are shown flanking a junction between DNA sequences not normally found in nature. This approach has two advantages over locating both primers within a single stretch of DNA (such as shown for primers 2> and <2): some information about the presence or absence of promoter, as well as the coding sequence, is provided and it may help avoid the occurrence of false positive results.

on how restriction digests are designed, information regarding the complexity of transgene insertion(s), the number of transgene copies present, the integrative status and the number of chromosomal sites where the transgene(s) has inserted can be obtained.

For analysis of copy number and integration complexity, digests should be designed to excise the entire DNA fragment for which information is desired. Most often, this will be a PTU, since it contains both the coding sequence and accompanying regulatory sequences. While it is not uncommon in the literature to see only the coding sequence analyzed, this approach does not provide as much information as analysis of the entire PTU. This approach can be used whether *Agrobacterium* or direct DNA delivery (DDD, e.g. electroporation or particle gun bombardment) was used to produce the material being analyzed. The vector DNA used for transformation should be digested and analyzed alongside sample genomic DNA to provide the necessary controls for evaluating rearrangements and copy number. Sample bands of a different size from the control band are considered rearranged; however, this analysis does not characterize the rearrangement(s). Copy number reconstructions can be made by diluting vector DNA to appropriate concentrations in a matrix of DNA from untransformed plants. Hybridizing signal strengths of these controls can then be compared with those in samples to estimate copy number. Because of the sample-to-sample variability between DNA preparations (with regard to degradation and digestibility) and the imprecision of genome size estimates (a review of the literature indicates up to a threefold spread in estimates for some species; compare for instance Refs 16,17), only rough estimates (± approximately 2–3-fold) can be made using this approach. Accuracy can be improved somewhat by also probing for an endogenous gene having a known copy number[18], but because this is more time consuming, and many labs do not have access to such probes, this approach is often not taken. Examples of the kinds of results that may be obtained from this analysis are shown in Fig. 2, lanes 5–8.

Southern analysis can be used to provide evidence for transgene integration, including insertion at multiple sites, although misleading results can be obtained. The analysis is most straightforward when the material being analyzed was produced using *Agrobacterium*. In this case the T-DNA right border typically provides a defined junction between vector and plant DNA which can be exploited as shown in Fig. 3a. In this border analysis, the presence of multiple bands provides evidence for multiple insertion sites, but not proof, as multiple bands can also result from transgene rearrangement and multimerization at single insertion sites. The situation is less straightforward for material produced by DDD as vector–plant DNA junctions are more random. Two approaches can be taken (Figs 3b and 3c). In one (Fig. 3b) the restriction enzyme cuts only within the plant DNA and not the vector used for transformation. The presence of multiple DNA bands is diagnostic for integration at multiple sites. Because the resulting transgene-containing fragments can be fairly large (particularly if the DNA used for transformation was larger than approximately 5–8 kb), multiple fragments may not be resolved using standard agarose electrophoresis conditions. An alternative is to use a restriction enzyme that cuts once within the transgene (Fig. 3c). This approach tends to produce smaller transgene-containing fragments but multiple bands due to transgene rearrangements and multimerization at single insertion sites can also occur. Examples of the kinds of results that can be obtained using these approaches are shown in Fig. 2, lanes 9–12.

In the approaches described above, a sample DNA fragment size different from that produced by the vector (particularly a higher molecular mass band) is indicative of, but does not prove, integration. Similarly, transgene probe hybridization to unrestricted high molecular mass DNA provides evidence for, but not proof of, transgene integration. Additional evidence, such as mendelian inheritance or cloned vector–plant DNA junctions, is required to demonstrate integration unambiguously. The importance of assessing the transgenic status of material using several complementary methods when the transformation approach being used is not well characterized has been illustrated by Langridge *et al.*[19].

Southern analysis of transgenes can be particularly problematic when PTU sequences are present in the untransformed host genome. In some cases this problem can be circumvented by using a probe to a part of the PTU not present in the genome, but this approach is not always feasible. In cases where the probe hybridizes to endogenous, as well as transgene, DNA, three precautions are recommended. First, try and use restriction enzymes that produce as simple a pattern as possible when the untransformed DNA is analyzed. Second, these fragment size(s) should be sufficiently different from those of the transgene fragment such that interpretation of the band pattern in transformed material is not ambiguous. Finally, untransformed genomic DNA controls need to be included on every row of a gel (this is always a good idea but is particularly crucial under these circumstances).

Transgene expression

Transgene expression studies typically examine either protein or RNA accumulation. While such analysis does a great deal to characterize transformants, the results do not directly address their transgenic status. Thus, correlating expression with stable transformation is inappropriate for all but the best-characterized transformation systems. The three questions asked most frequently in expression studies are: how much?, when? and where? The techniques described in this section can be used to provide detailed answers to how much? and when? and crude (i.e. resolution at the organ and sometimes tissue levels, but not at the cell level) answers to where? Experiments addressing how much? often require statistical analysis of data. Because transgene expression data from T_0 plants often do not have a normal distribution,

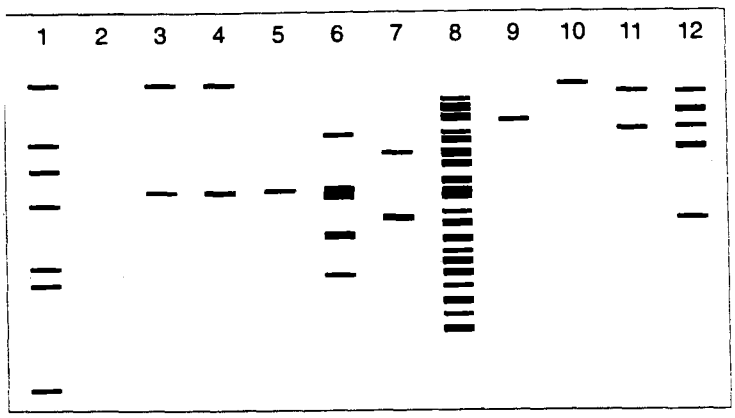


Figure 2

Representative Southern analysis results. This drawing shows cartoon examples of Southern data for either plant transcription unit (PTU) characterization (lanes 2–8) or analysis of the number of integration sites (lanes 9–12). Lane 1 represents a λ HindIII digest, which produces fragments of ~23.1, 9.4, 6.6, 4.4, 2.3, 2.0 and 0.6 kb. Lanes 2–4 represent copy number reconstructions for PTU analysis of a ~5 kb PTU, having 0, 1 and 5 copy equivalents, respectively. Lane 5 shows the expected result for a low (one to a very few) copy unrearranged PTU insertion. Lane 6 shows the expected result for material having both unrearranged (high copy number) and rearranged transgene copies. Lane 7 shows an example of a result from material having only rearranged transgene copies. Lane 8 shows an example of the extremely complex high copy number patterns that can occur, particularly with material produced using direct DNA delivery. Lanes 9 and 10 show examples of results from integration analyses that, as described in the text, indicate single insertion sites. In particular, the high molecular mass band in lane 10 may be misleading as multiple DNA fragments can co-migrate in this portion of the gel. Lanes 11 and 12 show patterns consistent with more than one insertion locus. As noted in the text, such results can also be misleading as, depending on the restriction digest approach used, these patterns (particularly as shown in lane 11) may result from rearrangement of multiple tandem copies. Patterns similar to that shown in lane 12 do correlate well with the presence of multiple insertion sites and are often seen with material having a PTU profile like that shown in lane 8.

statistical tools not typically employed by molecular biologists or biochemists must be considered[20]. Detailed tissue and cell localization studies require other approaches such as *in situ* hybridization (for reviews see Refs 21,22), which are beyond the scope of this review. Some marker genes also offer the potential for histological analysis (see below).

RNA analysis

In most cases, analysis of transgene expression focuses on proteins or other end products since the desired phenotype results from accumulation of a specific protein. When protein analysis is not possible, analysis of RNA transcripts is often used. Even in cases where protein analysis is carried out, RNA analysis can provide useful information about transcript accumulation and stability that may help explain unexpected phenotypes. Techniques such as reverse transcription (RT) PCR[23], northern analysis[14,15,24] and the ribonuclease protection assay (RPA; Ref. 25) can be used to measure steady-state RNA levels. Analysis of RNA synthesis or degradation rates requires nuclear transcription run-on analysis[26].

RT-PCR can be used as a quick and relatively high-throughput screen for the presence or absence of a specific transcript. The advantages of RT-PCR, as compared with northern analysis or RPA, are the small amount of material needed, the high sensitivity and throughput and the ease of sample preparation. Unlike northern analysis, however, RT-PCR results provide no information about transcript size. If RT-PCR is to be used for quantitative analysis, additional effort and controls are required[27,28]. The ability to discriminate between products resulting from amplification of sample genomic DNA versus cDNA synthesized from sample RNA *in vitro* is critical when using RT-PCR. Thorough DNase digestion of RNA preparations is always required and, whenever possible, primers should be located such that amplification of sample DNA is either impossible or results in a product having a different size than a cDNA amplification product[29,30].

Immunoassays

Immunoassays[31] are the most commonly used approach for studying accumulation of proteins in transgenic plants. There are two broad types of immunoassay typically employed: immunoblot hybridization (western) and the enzyme-linked immunosorbent assay (ELISA).

Western immunoassays and ELISAs have largely complementary strengths and weaknesses. Advantages offered by westerns over ELISAs include less assay development (ELISA development can be nearly as much art as science) and the ability to assess protein molecular masses. In addition, acceptable results can often be obtained using relatively impure antibody preparations. However, westerns are much more time consuming to run than are ELISAs and quantitative results are approximate. Historically, westerns were often less sensitive than ELISAs; however, with the appearance of robust chemiluminescent detection systems (e.g. Ref. 32) the magnitude of this difference has decreased. Indeed, the costs of assay development and adopting the quality assurance measures required to obtain the most accurate ELISA results (e.g. Refs 33,34) may more than offset the advantages for many applications.

Foremost among the technical problems typically encountered when using immunoassays is an unacceptable background[31,35,36]. Certain backgrounds may be due to the specific interaction of an antibody (either primary or secondary, if more than one antibody is used) with a protein in the extract, while others are nonspecific. The two most common sources of nonspecific background problems are antibody dilutions and blocking. Backgrounds can often be reduced or eliminated by decreasing antibody concentrations, and the optimal blocking reagent and blocking conditions should be determined empirically for each protein.

Transgene product activity

In most transgenic plants, the desired phenotype is due to a specific activity of the protein product. This activity may or may not be enzymatic and may be measurable using a biochemical assay and/or some

type of bioassay (e.g. resistance to an insect or pathogen pressure). In any event, it is important to remember that, since this activity is, in many cases, the most direct link to the desired *in planta* phenotype, its measurement may be crucial to understanding the phenotype. While the myriad of potential assays that could be used for such analysis cannot be covered adequately here, such analysis should be included as a part of a comprehensive molecular and biochemical analytical package. In an analogous manner, such analysis of endogenous enzyme activity can be useful when antisense (or sense) gene suppression is being attempted. It is important to remember that activity assays are not measuring the same thing as an immunoassay and so activity results will not necessarily correlate with protein accumulation levels.

Marker enzymes

A number of scorable marker enzymes have been used extensively in plants, most notably β-glucuronidase[37,38] and luciferase[39]. Recently, an autofluorescent protein from jellyfish, the green fluorescence protein, was described[40], and its use in plants seems certain to increase (e.g. Ref. 41). These markers are extremely useful in transgene expression studies (some for both quantitative and histochemical analyses), for promoter characterization and for quickly screening large numbers of transformants so that subsequent detailed analysis can focus on the most appropriate subset of plants.

The general utility of marker genes and the ease with which results can be obtained has unfortunately resulted in their use for inappropriate purposes. Specifically, expression of a marker gene, like expression of any other gene, does not necessarily correlate with stable transformation. When developing a transformation system, screening hundreds-to-thousands of putative transgenics is often required to identify that first rare event or to optimize the transformation system. In scenarios such as these, it makes sense to use a rapid high-throughput screen to select the material that deserves subsequent analysis and marker analysis, like PCR, provides a powerful tool for such screening. Problems arise, however, when positive results from such screening are used as verification of transgenic status. As described above, other technologies must be used to obtain this information.

Inheritance of transgenes

The same techniques and approaches described above for determining the transgenic status of transformed material can be applied to analysis of transgene inheritance. Just as transgenes can be stably integrated into a host plant genome without being expressed, giving rise to potential analytical problems, transgenes can be inherited in a predictable manner but not expressed (see Refs 4,42 and refs therein). Noninheritance as well as inheritance inconsistent with mendelian genetics can also occur (see Refs 4,42 and refs therein). Therefore, obtaining a complete picture of transgene inheritance may require use of multiple complementary molecular and biochemical approaches. The same suggestions and caveats presented earlier in this review regarding application of various techniques should also be considered when developing an approach for analysis of inheritance.

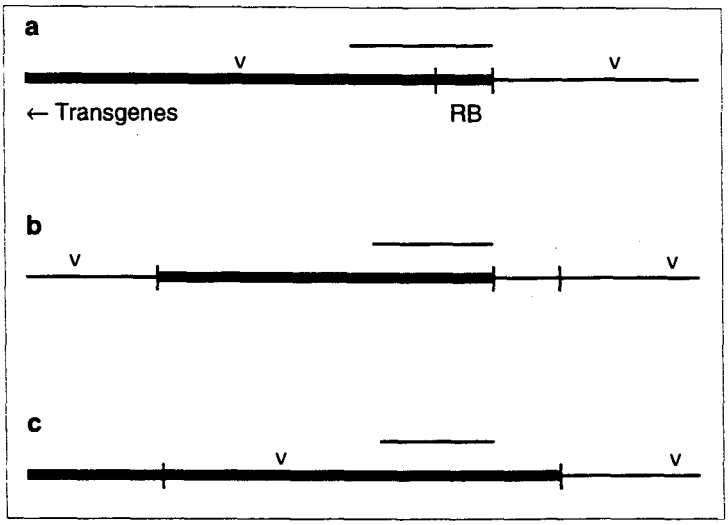


Figure 3

Southern analysis for determining the presence of multiple integration sites. The schematic diagrams represent vector (transgene) DNA (▬) integrated into plant DNA (—). Restriction sites are shown (v) and probe locations are indicated by the short line positioned above each vector DNA. (**a**) An approach used for transgenics produced using *Agrobacterium*, in which a restriction enzyme is used that cuts at a known site within the T-DNA (which may or may not be near the genes of interest) and at an unknown site in the plant DNA. The probe is typically located very near or at the right border (RB). (**b**) An approach used for transgenics produced via direct DNA delivery, in which a restriction enzyme that does not cut within the vector DNA is used. Because the vector–plant DNA junctions are not as well defined as for transformants produced using *Agrobacterium*, it is, typically, useful to employ a transgene coding sequence (particularly the selectable marker gene) as the probe as these sequences are the least likely to have been disrupted during integration. (**c**) An alternative approach, in which a restriction enzyme is used that cuts once within the vector DNA. In this approach it is useful to use an enzyme that does not cut within (or immediately flanking) a PTU and, as for approach (b), employing a probe from a PTU coding sequence is recommended.

Concluding comments

As noted earlier, analysis of transformed plant material has one of two objectives: determining if the material is transgenic and its characterization. Various molecular and biochemical techniques are available for such analysis but, as described in this review, each addresses the objectives differently. Therefore, quality analytical data must always be considered the product of not only technical competence but also an understanding of both the experimental objectives and the characteristics of the material being analyzed. A deficiency in any one of these three areas may result in untenable conclusions.

Acknowledgements

I thank David Hondred and Penny Hunst for reviewing the manuscript, and present and past colleagues at Pioneer Hi-Bred and Garst/Zeneca Seeds for helping shape the thoughts and approaches presented here.

References

1 Czerlinofsky, A. P., Hain, R., Baker, B. and Wirtz, U. (1986) *DNA* 5, 473–482
2 Deroles, S. C. and Gardner, R. C. (1988) *Plant Mol. Biol.* 11, 355–364
3 Bellini, C., Guerche, P., Spielmann, A., Goujaud, J., Lesaint, C. and Caboche, M. (1989) *J. Hered.* 80, 361–367
4 Register, J. C., III *et al.* (1994) *Plant Mol. Biol.* 25, 951–961
5 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual* (2nd edn), Cold Spring Harbor Laboratory Press
6 Hayashi, K. (1994) in *The Polymerase Chain Reaction* (Mullis, K. B., Ferre, F. and Gibbs, R. A., eds), pp. 3–13, Birkhäuser
7 Ferre, F., Marchese, A., Pezzoli, P., Grifin, S., Buxton, E. and Boyer, V. (1994) in *The Polymerase Chain Reaction* (Mullis, K. B., Ferre, F. and Gibbs, R. A., eds), pp. 67–88, Birkhäuser
8 Coen, D. M. (1994) in *The Polymerase Chain Reaction* (Mullis, K. B., Ferre, F. and Gibbs, R. A., eds), pp. 89–96, Birkhäuser
9 Ferre, F. (1992) *PCR Methods Appl.* 2, 1–9
10 Raeymaekers, L. (1995) *Genome Res.* 5, 91–94
11 Dieffenbach, C. W. and Dveksler, G. S. (1993) *PCR Methods Appl.* 3, S2–S7
12 Dragon, E. (1993) *PCR Methods Appl.* 3, S8–S9
13 Hartley, J. L. and Rashtchian, A. (1993) *PCR Methods Appl.* 3, S10–S13
14 Dunsmuir, P., Bond, D., Lee, K., Gidoni, D. and Townsend, J. (1988) in *Plant Molecular Biology Manual* (Gelvin, S. B. and Schilperoort, R. B., eds), pp. C1/1–17, Kluwer Academic Press
15 Bers, G. and Garfin, D. (1985) *BioTechniques* 3, 276–288
16 Rogers, S. O. and Bendich, A. J. (1988) in *Plant Molecular Biology Manual* (Gelvin, S. B. and Schilperoort, R. B., eds), pp. A6/1–11, Kluwer Academic Press
17 Arumuganathan, K. and Earle, E. D. (1991) *Plant Mol. Biol. Rep.* 9, 208–218
18 Sanders, R. A., Sheehy, R. E. and Martineau, B. (1992) *Plant Mol. Biol. Rep.* 10, 164–172
19 Langridge, P., Brettschneider, R., Lazzeri, P. and Lorz, H. (1992) *Plant J.* 2, 631–638
20 Nap, J-P., Keizer, P. and Jansen, R. (1993) *Plant Mol. Biol. Rep.* 11, 156–164
21 Leitch, A. R., Schwarzacher, T., Jackson, D. and Leitch, I. J. (1994) *In Situ Hybridization: A Practical Guide*, Bios Scientific Publishers
22 Jiang G., ed. (1995) *In Situ Polymerase Chain Reaction and Related Technology*, Eaton Publishing
23 Chelly, J. and Kahn, A. (1994) in *The Polymerase Chain Reaction* (Mullis, K. B., Ferre, F. and Gibbs, R. A., eds), pp. 97–109, Birkhäuser
24 Nagy, F., Kay, S. A. and Chua, N-H. (1988) in *Plant Molecular Biology Manual* (Gelvin, S. B. and Schilperoort, R. B., eds), pp. B4/1–29, Kluwer Academic Press
25 Lee, J. J. and Costlow, N. A. (1987) *Methods Enzymol.* 152, 633–648
26 Lissemore, J. and Quail, P. H. (1988) *Mol. Cell. Biol.* 8, 4840–4850
27 Chen, D. and Klebe, R. J. (1993) *PCR Methods Appl.* 3, 127–129
28 Gause, W. C. and Adamovicz, J. C. (1993) *PCR Methods Appl.* 3, S123–S135
29 Moore, R. E., Shepherd, J. W. and Hoskins, J. (1990) *Nucleic Acids Res.* 18, 1921
30 Shuldiner, A. R., Tanner, K., Moore, C. A. and Roth, J. (1991) *BioTechniques* 11, 760–763
31 Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press
32 Bronstein, I., Voyta, J. C., Murphy, O. J., Bresnick, L. and Kricka, L. J. (1992) *BioTechniques* 12, 748–753
33 Bishop, J., Dunstan, F. D. J., Nix, B. J. and Swift, A. (1993) *J. Immunol. Methods* 166, 191–200
34 Gentleman, R. C., Hamada, M. S., Matthews, D. E. and Wilson, A. R. (1994) *J. Am. Stat. Assoc.* 89, 1200–1208
35 Pruslin, F. H., To, S. E., Winston, R. and Rodman, T. C. (1991) *J. Immunol. Methods* 137, 27–35
36 Pampori, N. A., Pampori, M. K. and Shapiro, B. H. (1995) *BioTechniques* 18, 588–590
37 Gallagher, S. R., ed. (1992) *GUS Protocols: Using the GUS Gene as a Reporter of Gene Expression*, Academic Press
38 Bronstein, I. *et al.* (1994) *BioTechniques* 17, 172–177
39 Kost, B., Schnorf, M., Potrykus, I. and Neuhaus, G. (1995) *Plant J.* 8, 155–166
40 Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994) *Science* 263, 802–805
41 Reichel, C. *et al.* (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 5888–5893
42 Finnegan, J. and McElroy, D. (1994) *Biotechnology* 12, 883–888

References

1 Yarranton, G. T. (1992) *Curr. Opin. Biotech.* 3, 506–511
2 Christopherson, K. S., Mark, M. R., Bajaj, V. and Godowski, P. J. (1992) *Proc. Natl Acad. Sci. USA* 89, 6314–6318
3 Yao, T-P., Segraves, W. A., Oro, A. E., Mckeown, M. and Evans, R. M. (1992) *Cell* 71, 63–72
4 Reznikoff, W. S. (1992) *Mol. Microbiol.* 6, 2419–2422
5 Barkley, M. D. and Bourgeois, S. (1980) in *The Operon* (Miller, J. H. and Reznikoff, W. S., eds), pp. 177–220, Cold Spring Harbor Laboratory Press
6 Hu, M. C-T. and Davidson, N. (1987) *Cell* 48, 555–566
7 Brown, M., Figge, J., Hansen, U., Wright, C., Jeang, K. T., Khoury, G., Livingston, D. M. and Roberts, T. M. (1987) *Cell* 49, 603–612
8 Figge, J., Wright, C., Collins, C. J., Roberts, T. M. and Livingston, D. M. (1988) *Cell* 52, 713–722
9 Deuschle, U., Pepperkok, R., Wang, F., Giordano, T. J., McAllister, W. T., Ansorge, W. and Bujard, H. (1989) *Proc. Natl Acad. Sci. USA* 86, 5400–5404
10 Fuerst, T. R., Fernandez, M. P. and Moss, B. (1989) *Proc. Natl Acad. Sci. USA* 86, 2549–2553
11 Zhang, Y., Keck, J. G. and Moss, B. (1992) *J. Virol.* 66, 6470–6479
12 Deuschle, U., Gentz, R. and Bujard, H. (1986) *Proc. Natl Acad. Sci. USA* 83, 4134–4137
13 Deuschle, U., Hipskind, R. A. and Bujard, H. (1990) *Science* 248, 480–483
14 Gossen, M. (1993) PhD thesis, University of Heidelberg
15 Reines, D. and Mote, J. (1993) *Proc. Natl Acad. Sci. USA* 90, 1917–1921
16 Labow, M. A., Baim, S. B., Shenk, T. and Levine, A. J. (1990) *Mol. Cell. Biol.* 10, 3343–3356
17 Zambetti, G. P., Olson, D., Labow, M. and Levine, A. J. (1992) *Proc. Natl Acad. Sci. USA* 89, 3952–3956
18 Baim, S. B., Labow, M. A., Levine, A. J. and Shenk, T. (1991) *Proc. Natl Acad. Sci. USA* 88, 5072–5076
19 Wyborski, D. L. and Short, J. M. (1991) *Nucleic Acids Res.* 19, 4647–4653
20 Hillenand Wissmann, A. (1989) in *Protein–Nucleic Acid Interaction. Topics in Molecular and Structural Biology* (Saenger, W. and Heinemann, U., eds), Vol. 10, pp. 143–162, Macmillan Press
21 Degenkolb, J., Takahashi, M., Ellestad, G. and Hillen, W. (1991) *Antimicrob. Agents Chemother.* 35, 1591–1595
22 Kleinschnidt, C., Tovar, K-H., Hillen, W. and Porschke, D. (1988) *Biochemistry* 27, 1094–1104
23 Gatz, C., Frohberg, C. and Wendenburg, R. (1992) *Plant J.* 2, 397–404
24 Bonin, A. L. (1993) PhD thesis, University of Heidelberg
25 Gossen, M. and Bujard, H. (1992) *Proc. Natl Acad. Sci. USA* 89, 5547–5551
26 Oliva, B., Gordon, G., McNicholas, P., Ellestad, G. and Chopra, I. (1992) *Antimicrob. Agents Chemother.* 36, 913–919
27 Hlavka, J. J. and Boothe, J. H. (eds) (1985) *Handbook of Experimental Pharmacology*, Vol. 78, Springer-Verlag
28 Gill, G. and Ptashne, M. (1988) *Nature* 334, 721–724

Foreign gene expression in transgenic cereals

David McElroy and Richard I. S. Brettell

Recent advances in transformation technology have resulted in the routine production of transgenic plants for an increasing number of cereal species. With a view to improving cereal quality and agronomic performance by genetic engineering, attention is beginning to focus on the characterization of those molecular elements that will be used to regulate foreign gene expression in transgenic cereals.

Cereals comprise a commercially valuable group of plants species that could benefit from the introduction and expression of foreign genes controlling improved grain quality and such agronomically important traits as tolerance to disease and stress. However, while it is over ten years since methods were first developed for the genetic engineering of dicotyledonous (dicot) plants, most cereals have not proven readily amenable to either *Agrobacterium*-mediated gene delivery, or (with the exception of rice) to the routine regeneration of fertile transgenic plants from directly transformed protoplasts[1]. The use of microprojectile-bombardment-mediated transformation of embryogenic tissue culture material, with the subsequent generation of transgenic plants, has overcome the regeneration problems associated with the production of plants from cereal protoplasts. Using this technology, transgenic plants have been obtained from microprojectile-bombarded suspension cultures of rice[2] and maize[3,4], callus cultures of oats[5], sugarcane[6] and wheat[7], and immature zygotic embryos of rice[8], maize[9], wheat[10] and barley[11].

With the development of transformation systems for an increasing number of previously recalcitrant monocotyledonous (monocot) species, attention is

D. McElroy (currently at the USDA/ARS Plant Gene Expression Center, Albany, CA 94710, USA) and R. I. S. Brettell are at the CSIRO Division of Plant Industry, Canberra, ACT 2601, Australia.

Table 1. Reporter genes used in cereal transformation

Properties	β-glucuronidase[a]	Luciferase[b]	Anthocyanin regulators[c]
Source	*E. coli*	Firefly	Maize
Background activity in plants	Low (some cases due to bacterial contaminants)	Low	Low–moderate (depending upon species/tissues)
Nature of assay	Destructive	Nondestructive	Nondestructive
Enzyme stability	High	Low	Low
Sensitivity of assay	High	Moderate	Low
Simplicity of assay	Good	Poor	Good
Quantitative nature of assay	Good	Good	Poor
Versatility of assay (quantitative, histochemical, etc.)	Good	Poor	Poor
Adverse effects on transgenic plant metabolism	Low	Low	Some (at high levels)
Relative cost of assay systems	Moderate	High (requires expensive detection equipment)	Low
Cereal transformed	Rice[15]	Maize[16]	Maize[17]

[a,b,c]Data from Refs 12, 13 and 14, respectively.

now turning towards the characterization of those molecular elements that might be used to regulate the expression of foreign genes in transgenic cereals. Although such studies are not as far advanced as those for transgenic dicots, several steps have been taken towards the development of regulatory systems for use in cereal transformation. These are: (1) the development of novel reporter genes; (2) the characterization of promoter elements; and (3) the elucidation of mechanisms that act to enhance or interfere with foreign gene expression in transgenic cereals. In combination with the characterization of genes controlling agronomically important traits, these developments will provide a basis for improving cereal quality by genetic engineering.

Reporter genes used in cereal transformation

Reporter genes (Table 1) are used in cereal transformation to analyse gene expression. The reporter gene should (ideally) exhibit low background activity in cereals and should not have any detrimental effects on metabolism. The reporter gene product should have moderate stability *in vivo*, so that down-regulation of gene expression as well as gene activation can be detected. Finally, the reporter gene should be able to be assayed by a nondestructive, quantitative, sensitive, simple to perform and inexpensive system. Unfortunately, the reporter genes currently in general use do not fulfil all these criteria.

β-glucuronidase

The β-glucuronidase (*gus*) gene[12], encoded by the *uidA* locus of *E. coli*, is the most popular reporter gene used in cereal transformation. β-glucuronidase catalyses the hydrolysis of a wide range of fluorometric and histochemical β-glucuronide substrates. GUS enzyme activity can be assayed easily and sensitively in plants, the expression of *gus* gene fusions can be quantified by fluorometric assay, and histochemical analysis can be used to localize gene activity in transgenic tissues. However, *gus* gene-expression assays are destructive and the GUS protein shows relatively high stability *in vivo* – monitoring GUS activity is therefore not optimal for detecting gene inactivation. Furthermore, a dependence on the use of bacterial reporter genes to monitor the efficiency of cereal-transformation protocols can often be misleading.

Luciferase

The products of the firefly (*Photinus pyralis*) luciferase gene[13] catalyses the oxidation of D(–)-luciferin in the presence of ATP to generate oxyluciferin and yellow-green light. The activity of luciferase gene fusions can be assayed nondestructively in transformed cereal tissue. Unfortunately, penetration of the luciferin substrate can be limiting in whole-plant material. In addition, the equipment needed to monitor luciferase gene expression is relatively expensive at present.

Anthocyanins

A reporter system that does not require the application of external substrates for its detection utilizes the *C1*, *B* and *R* genes, which code for *trans*-acting factors that regulate the anthocyanin biosynthetic pathway in maize seeds. The introduction of these regulatory genes (under the control of constitutive promoters) into cereal cells by microprojectile bombardment induces cell-autonomous pigmentation in non-seed tissues[14]. However, there is some evidence to suggest that expression of these *trans*-acting factors in transformed cereal tissues can be debilitating (R. Birch, pers. commun.).

Table 2. Summary of constitutive promoters used in transgenic cereals

Promoter	Source	Relative activity in cereal cells	Use in transgenic cereals	Construct[a]	Refs
35S	Cauliflower mosaic virus 35S RNA transcript	Low	Rice	35S–*bar*	2,8
				35S–*hpt*	18
				35S–*als*	19
				35S–*gus*	20
				35S–*nptII*	20
				35S–*dhfr*	21
			Maize	35S–*bar*	4,9
				35S–*nptII*	22
			Fescue	35S–*hpt*	23
35S–*Adh1* intron 1	Cauliflower mosaic virus 35S promoter and first intron of maize alcohol dehydrogenase 1 gene	Low	Maize	35S–*Adh1* intron 1–*bar*	3
			Oats	35S–*Adh1* intron 1–*bar*	5
			Wheat	35S–*Adh1* intron 1–*bar*	7
Emu	Modified maize alcohol dehydrogenase 1 promoter and first intron	Moderate	Sugar cane	Emu–*nptII*	6
			Rice	Emu–*nptII*	24
Act1–*Act1* intron 1	Rice actin 1 gene	Moderate	Rice	*Act1*–*Act1* intron 1–*bar*	2
				Act1–*Act1* intron 1–*gus*	15
Ubi1–*Ubi1* intron 1	Maize ubiquitin 1 gene	High	Wheat	*Ubi1*–*Ubi1* intron 1–*bar*	10
				Ubi1–*Ubi1* intron 1–*gus*	10
			Barley	*Ubi1*–*Ubi1* intron 1–*bar*	11
				Ubi1–*Ubi1* intron 1–*gus*	11
			Rice	*Ubi1*–*Ubi1* intron 1–*bar*	25

[a]Abbreviations: *gus*, β-glucuronidase gene; *nptII*, neomycin phosphotransferase II gene; *hpt*, hygromycin phosphotransferase gene; *bar*, phosphinothricin acetyl transferase gene; *als*, acetolactate synthase gene; *dhfr*, dihydrofolate reductase gene; *Adh1*, alcohol dehydrogenase gene; *Ubi1*, ubiquitin 1 gene.

Promoter elements used to control foreign gene expression in transgenic cereals

For an increasing number of monocot species, the routine production of transgenic plants has facilitated the characterization of those regulatory promoter elements that might eventually be used to control the expression of heterologous genes in transformed cereals.

Constitutive promoters

The use of constitutive promoters in transgenic cereals (Table 2) has generally been restricted to the expression of genes encoding resistance to antibiotics or herbicides[26] (Table 3). These selectable-marker genes are used to discriminate between transgenic and nontransgenic cells in systems that generally display low transformation efficiencies. The usefulness of individual selectable-marker genes is a function both of the properties of the respective resistance proteins that they encode, and of the relative sensitivity of the target tissue to the corresponding selective agent[26]. A number of antibiotics and herbicides have been used as selective agents in cereal transformation. However, resistance to phosphinothricin (PPT)-based herbicides, using the *bar* gene from *Streptomyces hygroscopicus* is fast becoming the method of choice for selecting fertile transgenic cereals[2–5,7,10,11,23,24].

In transient assays of *gus* reporter-gene constructs, the constitutive promoters commonly used in cereal transformation show differences in their relative activity in monocot cells[27]. For example, the promoter of the cauliflower mosaic virus (CaMV) 35S RNA transcript (*35S*), which has been used extensively in dicot transformation, shows relatively low activity in transient assays of monocot cells transformed with *gus* reporter genes[15,27,29]. A number of strategies have been used to increase the effectiveness of constitutive promoters in transformed cereals. The incorporation of an intron into the transcriptional unit of the foreign gene increases the abundance of the mRNA and increases gene expression in transformed monocot cells[17,27,28]. The combination of the first intron of maize alcohol dehydrogenase 1 gene (*Adh1*) and the CaMV *35S* promoter has been used to control expression of the *bar* gene in transformed maize[3], oat[5] and wheat[7]. An alternative strategy is to modify a monocot promoter so that the gene that it controls is expressed constitutively at high levels in cereal cells.

An example of this strategy is the inclusion of a modified maize *Adh1* sequence in the Emu promoter[29], which has been used to control expression of the neomycin phosphotransferase (*nptII*) gene in transformed rice[24] and sugar cane[6]. Finally, efforts have been made to isolate monocot promoters that naturally show high-level constitutive activity. Examples of such sequences include the promoters of the genes for rice actin (*Act1*) and maize ubiquitin (*Ubi1*), which have been used to express the *bar* gene in transgenic wheat[10], barley[11] and rice[2,25]. The eventual applications of these constitutive promoters in transgenic cereals will include the repression of endogenous and/or pathogenic gene expression through antisense technologies[30], the overproduction of biomolecules in transgenic plants[31], and the overexpression and assaying of disease-resistance genes in transgenic cereals, prior to the subsequent employment of nonconstitutive expression strategies[32,33].

Nonconstitutive promoters

Advances in cereal genetic engineering have tended to lag behind corresponding developments both in dicot transformation systems and in the characterization of dicot genes. This trend has had two major effects on the study of foreign gene expression in transgenic cereals. First, the availability of nonconstitutive dicot promoters for use in those monocot species where transformation technologies are available has increased. Second, there has been a tendency to study monocot gene expression either in transgenic dicots, or in transient assays of transformed cereal cells such as protoplasts[15,28], or microprojectile-bombarded tissues[16,17]. However, there is evidence to suggest that promoter elements from monocot species are not always regulated correctly either in transgenic dicot plants, or in monocot transient-assay systems[34].

Recent progress in rice transformation has enabled researchers to begin to study the regulation of gene expression in transgenic cereal tissues using *gus* reporter-gene fusions (Table 4). Whether the promoters used in these studies have originated from monocots (for example, *Adh1* from maize[35,36], *His3* from wheat[37], *LHCP* and *rbcS* from rice[38,39]), dicots (for example, *rbcS* from tomato[39] and *pinII* from potato[40]) or plant pathogens (for example, *rolC* from *Agrobacterium rhizogenes*[41] and rice tungro bacilliform virus major transcript[42]), it has been found that inclusion of the promoter region alone is sufficient to give the expected pattern of reporter-gene expression in transgenic rice. However, a comparison of the rice and tomato *rbcS* promoters[39] in transgenic rice revealed that monocot promoters can show higher activity than their dicot homologues.

Finally, the intron-mediated enhancement of gene expression in monocot cells[27,28], while increasing the activity of the (dicot) *pinII* promoter in transgenic rice, did not alter its pattern of regulation or expression[40]. However, there are very few other examples of such intron-mediated enhancement effects with nonconstitutive promoters and, consequently, a decision as to whether or not a particular cereal expression system will require the inclusion of an intron should be made following an initial investigation using an appropriate reporter gene.

Inactivation of foreign gene expression in transgenic cereals

There are a number of conditions that can interfere with gene expression in transgenic plants. These include poor recognition of promoter elements and differences in transcription efficiency between the source of the promoter and the transformed material[43], inefficient mRNA termination, polyadenylation, processing and/or stability[28,44,45], inefficient translation due to low rates of translation initiation[27] and/or inappropriate codon usage[46], and poor design of expression vectors, which might contain multiple repeated regions (leading to potential gene inactivation by recombination) or multiple genes with opposing orientations leading to transcriptional interference and/or antisense-mediated gene inactivation[47]. Some of these potential problems can be obviated by designing the expression vector appropriately. However, as transgenic cereals make their way from the laboratory to the field, there are a number of other endogenous mechanisms that can interfere with gene expression and lead to the non-mendelian inheritance of introduced traits.

Gene inactivation by methylation and co-suppression

As monocot transformation technologies improve and an increasing number of transgenic cereals are examined under field conditions, there are a growing number of reports describing non-mendelian inheritance of the introduced genes. These reports have, to date, involved the loss of reporter-gene expression under laboratory conditions and/or the loss of selectable-marker-gene expression in a nonselective field environment. In rice, this loss of gene expression would appear to be independent of the nature of the introduced genes[21,48]. While both copy number and the position of the integrated gene(s) can influence gene expression in transgenic plants, other explanations have been invoked to account for the non-mendelian inheritance of introduced genes, and for the observation that some plants do not express all copies of the integrated foreign gene[49].

Alterations in phenotype are commonly observed in plants regenerated from cultured cells and tissues, and there are an increasing number of reports in which both stable and reversible genetic alterations in both rice[50] and maize[51] have been linked to changes in patterns of DNA methylation. The most common DNA modification in plant cells is cytosine methylation at CG dinucleotides and CNG (where N can be any base) trinucleotides. Methylcytosine can interfere with protein–DNA interactions, and methylation-induced DNA modifications are believed to be a normal part of the control of gene expression in plant cells[49]. It has been shown that methylation can influence *35S–gus* reporter-gene expression in transgenic plants of

tobacco[52] and rice[53] (R. I. S. Brettell, unpublished), and that 5-azacytadine-induced demethylation is sufficient to reactivate *gus* gene expression[21,53] (R. I. S. Brettell, unpublished). The recognition and inactivation of introduced genes might be regarded as a defence mechanism that protects plants against the expression of potentially deleterious foreign DNA species[54]. A corollary of these observations is that it could be misleading to depend on the use of non-selectable reporter genes when following the efficiency of cereal transformation protocols.

In cereal transformation, whether using gene delivery by direct transfer to protoplasts or microprojectile bombardment of intact plant cells, there is a tendency towards the integration of multiple copies of the introduced genes[2,4,5,8,38,52]. The term, co-suppression, has been used to describe a situation in which multiple copies of a gene are coordinately suppressed[55]. It has been suggested that methylation of sequences required for active gene expression could account for this phenomenon[49]. In dicot transformation, such co-suppression usually involves the participation of an introduced gene and its endogenous homologue. However, gene inactivation has also been reported following the introduction of different *35S*-fusion constructs into transgenic dicots[56], and it would not be difficult to imagine similar effects occurring between multiple copies of foreign genes introduced into transformed cereal cells. Whatever the mechanism, the inactivation of introduced genes will have important implications for the maintenance of novel phenotypes in transgenic cereal plants.

Table 3.

Selective agent	Mode of action
Antibiotics	
Kanamycin	Binds to the organellar 30S ribosomal subunit to inhibit the initiation of RNA translation
Geneticin (G418)	Binds to the organellar 30S ribosomal subunit to inhibit the initiation of RNA translation
Hygromycin	Interacts with the elongation factor EF-2 to inhibit peptide chain elongation
Herbicides	
Phosphinothricin (PPT)	Glutamate analogue, which inhibits glutamine synthase leading to a cytotoxic accumulation of ammonia
Sulfonylurea	Inhibition of ALS leading to starvation of branched-chain amino acids

Future prospects

A number of issues still need to be addressed before transformation technology can be exploited fully in the agronomic improvement of cereal crops. Significant advances have been made in the application of this technology to a number of previously recalcitrant cereal species; the challenge is to extend this technology to those elite cultivars and advanced breeding lines that are of real agronomic importance.

Reporter genes are extremely useful for the study of gene regulation in transgenic cereals; however, none of the currently available visual markers have all the desired properties of an ideal reporter system, and efforts should be made to develop more, versatile reporter genes for use in cereal transformation. Antibiotic- and herbicide-resistance genes have proven to be invaluable in the development of cereal transformation technologies. However, alternative, environmentally benign, selection systems need to be developed for use in transgenic cereal crops that are to be grown freely in field situations.

The immediate products of current efforts in monocot transformation will be the engineering of herbicide-resistant cereals but, in the medium term, targets in cereal biotechnology will include (amongst many others) the engineering of disease resistance[30,32,33] and the modification of the quality of cereal grain[31]. Such applications will require specific spatial and/or temporal control of gene expression. Results from promoter studies tend to suggest that (assuming no morphological or anatomical limitation) both monocot and dicot promoters maintain their correct pattern of activity in transgenic cereals. However, for reasons that might be related to differences in the biochemistry, physiology and/or morphology of monocots and dicots, as well as to obviate potential problems with the use of non-cereal genetic elements in transgenic cereals, the isolation and utilization of monocot promoters will continue to be a rewarding area of research activity. In this respect, it is worth noting that the first report of a field trial of transgenic maize expressing an agronomically important trait involved the expression of insecticidal *Bacillus thuringiensis* resistance genes, under the control of leaf- and pollen-specific promoters from maize, to target gene expression to those tissues that are most susceptible to attack by the European corn borer[18].

While it is the characterization of constitutive and nonconstitutive promoter elements that has advanced the most in the area of foreign gene regulation in transgenic cereals, it should be recognized that there are other non-promoter elements that also contribute to the control of gene expression in transgenic plants.

Summary of selective agents and resistance genes used in transgenic cereals[a]

Resistance gene	Resistance enzyme	Origin	Mode of resistance	Use in transgenic cereals	Promoter	Refs
nptII (*aphA2*)	NPTII	*E. coli*	Detoxification by phosphorylation	Rice Maize	*CaMV 35S* *CaMV 35S*	20 22
nptII (*aphA2*)	NPTII	*E. coli*	Detoxification by phosphorylation	Sugar cane Rice	*Emu* *CaMV 35S*	6 20
hpt (*aphIV*)	HPT	*E. coli*	Detoxification by phosphorylation	Sugar cane Rice Fescue	*Emu* *CaMV 35S* *CaMV 35S*	6 18 23
bar	PAT	*Streptomyces hygroscopicus*	Detoxification by acetylation	Maize	*CaMV 35S–Adh1* intron 1 *CaMV 35S*	3
				Oats	*CaMV 35S–Adh1* intron 1	5
				Wheat	*CaMV 35S–Adh1* intron 1 *Ubi1–Ubi1* intron 1	7 10
				Rice	*CaMV 35S* *Act1–Act1* intron 1 *Ubi–Ubi1* intron 1	8 2 25
				Barley	*Ubi1–Ubi1* intron 1	11
				Fescue	*CaMV 35S*	23
csr-1	Mutant ALS	*Arabidopsis thaliana*	Tolerance	Rice	*CaMV 35S*	19

[a]Abbreviations: NPTII, neomycin phosphotransferase; CaMV, cauliflower mosaic virus; HPT, hygromycin phosphotransferase; Adh, alcohol dehydrogenase; ALS, acetolactate synthase; PAT, phosphinothricin acetyltransferase; ubi, ubiquitin.

Table 4. Summary of nonconstitutive promoter activities in transgenic rice using *gus* reporter-gene fusion constructs

Source	Gene	Promoter	Pattern of promoter–*gus* fusion gene expression in transgenic rice	Refs
Maize	Alcohol dehydrogenase 1 gene	*Adh1*	Constitutive in root caps, anthers, filaments, pollen, scutellum, endosperm and embryo shoot and root. Anaerobically induced in roots	35,36
Wheat	Histone 3 gene	*His 3*	Dividing root cells	37
Rice	Light-harvesting chlorophyll a/b-binding gene of photosystem II	*LHCP*	Light-inducible in leaves, stems and floral organs	38
Rice, tomato	Ribulose-1,5-bisphosphate carboxylase/ oxygenase small subunit gene	*rbcS*	Light-inducible in mesophyll cells	39
Potato	Wound-inducible II gene	*PinII*	Systemic induction by wounding, methyl jasmonate and abscisic acid	40
Agrobacterium rhizogenes	Open reading frame 12 (ORF12) of the Ri plasmid TL-DNA region	*rolC*	Vascular tissue and embryogenic tissue	41
Rice tungro bacilliform virus	Rice tungro bacilliform virus major transcript gene	*RTBV*	Leaf phloem tissue	42

These components include those that act to regulate the efficiency of transcript termination[44], transcript stability[45], post-transcriptional modification[28] and translation efficiency[27,46]. An increased understanding of these regulatory components in monocot plants will prove profitable for achieving the regulated expression of foreign genes in transformed cereals.

Finally, an understanding of the biological basis of foreign-gene inactivation, coupled to efforts to control the copy number and location of foreign-gene integration in the cereal genome, should prove beneficial for maintaining the fidelity of introduced genes in transgenic cereal crops.

Acknowledgements

The authors wish to thank Mick Graham and Liz Dennis for critical reading of the manuscript and for their many helpful suggestions. DM was supported by a postdoctoral research fellowship from the Australian Grains Research and Development Corporation, Grant CSP15.

References

1 Potrykus, I. (1990) *Physiol. Plantarum* 79, 125–134
2 Cao, J., Duan, X., McElroy, D. and Wu, R. (1992) *Plant Cell Rep.* 11, 586–591
3 Fromm, M. E., Morrish, F., Armstrong, C., Williams, R., Thomas, J. and Klein, T. M. (1990) *Bio/Technology* 8, 833–844
4 Gordon-Kamm, W. J., Spencer, T. M., Mangano, M. L., Adams, T. R., Daines, R. J., Start, W. G., O'Brien, J. V., Chambers, S. A., Adams, W. R., Willetts, N. G., Rice, T. B., Mackey, C. J., Krueger, R. W., Kausch, A. P. and Lemaux, P. G. (1990) *Plant Cell* 2, 603–618
5 Somers, D. A., Rines, H. W., Gu, W., Kaeppler, H. F. and Bushnell, W. R. (1992) *Bio/Technology* 10, 1589–1594
6 Bower, R. and Birch, R. G. (1992) *Plant J.* 2, 409–416
7 Vasil, V., Castillo, A. M., Fromm, M. E. and Vasil, I. K. (1992) *Bio/Technology* 10, 667–674
8 Christou, P., Ford, T. and Kofron, M. (1991) *Bio/Technology* 9, 957–962
9 Koziel, M. G., Beland, G. L., Bowman, C., Carozzi, N. B., Crenshaw, R., Crossland, L., Dawson, J., Desai, N., Hill, M., Kadwell, S., Launis, K., Lewis, K., Maddox, D., McPherson, K., Meghji, M. R., Merlin, E., Rhodes, R., Warren, G. W., Wright, M. and Evola, S. V. (1993) *Bio/Technology* 11, 194–200
10 Weeks, T., Anderson, O. D. and Blechl, A. E. (1993) *Plant Physiol.* 102, 1077–1084
11 Wan, Y. and Lemaux, P. G. (1994) *Plant Physiol.* 104, 37–48
12 Jefferson, R. A. and Wilson, K. J. (1991) in *Plant Molecular Biology Manual* (Gelvin, S., Schilperoot, R. and Verma, D. P., eds), pp. 1–33, Kluwer Academic Publishers
13 de Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R. and Subramani, S. (1987) *Mol. Cell. Biol.* 7, 725–737
14 Goff, S. A., Klein, T. M., Roth, B. A., Fromm, M. E., Cone, K. C., Radicella, J. P. and Chandler, V. L. (1990) *EMBO J.* 9, 2517–2522
15 Zhang, W., McElroy, D. and Wu, R. (1991) *Plant Cell* 3, 1155–1165
16 Klein, T. M., Roth, B. A. and Fromm, M. E. (1989) *Proc. Natl Acad. Sci. USA* 86, 6681–6685
17 Luehrsen, K. R. and Walbot, V. (1991) *Mol. Gen. Genet.* 225, 81–93
18 Shimamoto, K., Terada, R., Izawa, T. and Fujimoto, H. (1989) *Nature* 338, 274–276
19 Li, Z., Hayashimoto, A. and Murai, N. (1992) *Plant Physiol.* 100, 662–668
20 Battraw, M. and Hall, T. C. (1992) *Plant Sci.* 86, 191–202
21 Meijer, E. G. M., Schilperoort, R. A., van Os-Ruygrot, P. E. and Hensgens, L. A. M. (1991) *Plant Mol. Biol.* 16, 807–820
22 Rhodes, C. A., Pierce, D. A., Mettler, I. J., Mascarenhas, D. and Detmer, J. J. (1988) *Science* 240, 204–207
23 Wang, Z-Y., Takamizo, T., Iglesias, V. A., Osusky, M., Nagel, J., Potrykus, I. and Spangenberg, G. (1992) *Bio/Technology* 10, 691–696
24 Chamberlain, D. A., Brettell, R. I. S., Last, D. I., Witrzens, B., McElroy, D., Dolferus, R. and Dennis, E. S. (1994) *Aust. J. Plant. Physiol.* 21, 95–112
25 Toki, S., Takamatsu, S., Nojiri, C., Ooba, S., Anzai, H., Iwata, M., Christensen, A. H., Quail, P. H. and Uchimiya, H. (1992) *Plant Physiol.* 100, 1503–1507
26 Wilmink, A. and Dons, J. J. M. (1993) *Plant Mol. Biol. Rep.* 11, 165–185
27 McElroy, D., Blowers, A. D., Jenes, B. and Wu, R. (1991) *Mol. Gen. Genet.* 231, 150–160
28 Callis, J., Fromm, M. and Walbot, V. (1987) *Genes Dev.* 1, 1183–1200
29 Last, D. I., Brettell, R. I. S., Chamberlain, D. A., Chaudhury, A. M., Larkin, P. J., Marsh, E. L., Peacock, W. J. and Dennis, E. S. (1991) *Theoret. Appl. Genet.* 81, 581–588
30 Bejarano, E. R. and Lichtenstein, C. P. (1992) *Trends Biotechnol.* 10, 383–388
31 Visser, R. G. F. and Jacobsen, E. (1993) *Trends Biotechnol.* 11, 63–68
32 Lamb, C. J., Ryals, J. A., Ward, E. R. and Dixon, R. A. (1992) *Bio/Technology* 10, 1436–1445
33 Wilson, T. M. A. (1993) *Proc. Natl Acad. Sci. USA* 90, 3134–3141
34 Sakamoto, M., Sanada, Y., Tagiri, A., Murakami, T., Ohashi, Y. and Matsuoka, M. (1991) *Plant Cell Physiol.* 32, 385–394
35 Zhang, W. and Wu, R. (1988) *Theoret. Appl. Genet.* 76, 835–840
36 Kyozuka, J., Fujimoto, H., Izawa, T. and Shimamoto, K. (1991) *Mol. Gen. Genet.* 228, 40–48
37 Terada, R., Nakayama, T., Iwabuchi, M. and Shimamoto, K. (1993) *Plant J.* 3, 241–252
38 Tada, Y., Sakamoto, M., Matsuoka, M. and Fujimura, T. (1991) *EMBO J.* 10, 1803–1808
39 Kyozuka, J., McElroy, D., Hayakawa, T., Xie, Y., Wu, R. and Shimamoto, K. (1993) *Plant Physiol.* 102, 991–1000
40 Xu, D., McElroy, D., Thornburg, R. and Wu, R. (1993) *Plant Mol. Biol.* 22, 573–588
41 Matsuki, R., Onodera, H., Yamauchi, T. and Uchimiya, H. (1989) *Mol. Gen. Genet.* 220, 12–16
42 Bhattacharyya-Pakrasi, M., Peng, J., Elmer, J. S., Laco, G., Shen, P., Kaniewska, M. B., Kononowicz, H., Wen, F., Hodges, T. K. and Beachy, R. N. (1993) *Plant J.* 4, 71–79
43 Peterhans, A., Datta, S. K., Datta, K., Goodall, G. J., Potrykus, I. and Paszkowski, J. (1991) *Mol. Gen. Genet.* 222, 361–368
44 Hernandez, G., Cannon, F. and Cannon, M. (1989) *Plant Cell Rep.* 8, 195–198
45 Pierce, D. A., Mettler, I. J., Lachmansingh, A. R., Pomeroy, L. M., Weck, E. A. and Mascarenhas, D. (1987) in *Plant Gene Systems and Their Biology* (McIntosh, L. and Key, J., eds), pp. 301–310, Alan R. Liss
46 Murray, E. E., Lotzer, J. and Eberle, M. (1989) *Nucleic Acids Res.* 17, 477–499
47 Paszty, C. J. R. and Lurquin, P. F. (1990) *Plant Sci.* 772, 69–79
48 Schuh, W., Nelson, M. R., Bigelow, D. M., Orum, T. V., Orth, C. E., Lynch, P. T., Eyles, P. S., Blackhall, N. W., Jones, J., Cocking, E. C. and Davey, M. R. (1993) *Plant Sci.* 89, 69–79
49 Finnegan, E. J., Brettell, R. I. S. and Dennis, E. S. (1993) in *DNA Methylation: Molecular Biology and Biological Significance* (Jost, J. P. and Saluz, H. P., eds), pp. 218–261, Birkhäuser Verlag
50 Brown, P. T. H., Kyozuka, J., Sukekiko, Y., Kimura, Y., Shimamoto, K. and Lörz, H. (1990) *Mol. Gen. Genet.* 223, 324–328
51 Brettell, R. I. S. and Dennis, E. S. (1991) *Mol. Gen. Genet.* 229, 365–372
52 Weber, H. and Graessmann, A. (1989) *FEBS Lett.* 253, 163–166
53 Weber, H., Zeichmann, C. and Graessmann, A. (1990) *EMBO J.* 9, 4409–4415
54 Doerfler, W. (1991) *Biol. Chem.* 372, 557–564
55 van den Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. M. and Stuitje, A. R. (1990) *Plant Cell* 2, 291–299
56 Matzke, M. A., Primig, M., Trnovsky, J. and Matzke, A. J. M. (1989) *EMBO J.* 8, 643–649

Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat

Ann E. Blechl* and Olin D. Anderson

Agricultural Research Service, US Department of Agriculture, Western Regional Research Center, 800 Buchanan St., Albany, CA 94710-1105.
**Corresponding author (email: ablechl@pw.usda.gov).*

Received 5 February 1996; accepted 15 May 1996.

High-molecular-weight glutenin subunits (HMW-GS), one class of seed storage proteins, are important determinants of the bread-making quality of wheat flour. To change the amount and composition of these proteins via genetic engineering, a gene encoding a novel hybrid subunit under the control of native HMW-GS regulatory sequences was inserted into wheat. Of 26 independent transgenic lines identified by bialaphos selection, 18 expressed the cotransformed hybrid HMW-GS gene in their seed. The hybrid subunit accumulated to levels comparable to those of the native HMW-GS. These results show that a native HMW-GS gene promoter can be used to obtain high levels of expression of seed storage and, potentially, other proteins in transgenic wheat endosperm. Transgene expression was stable for at least three seed generations in the majority of lines. These experiments demonstrate the feasibility of constructing wheat plants with novel seed protein compositions.

Keywords: glutenin promoter, genetic engineering, seed storage protein modification

The high-molecular-weight glutenin subunits (HMW-GS), members of a family of seed storage proteins synthesized in developing wheat endosperm, are important determinants of the processing characteristics of wheat flours[1,2]. Although the molecular basis of the role of the HMW-GS in gluten functionality is not understood, it is assumed to reside in two prominent structural features: a central repetitive domain composed of short amino acid motifs that compose up to 85% of the protein sequence, and nonrepetitive terminal domains that contain the majority of the cysteine residues[2]. These cysteines mediate the intermolecular disulfide bonds among HMW-GS and with low-molecular-weight glutenin subunits (LMW-GS) to build protein polymers with a modal size of 1,000,000 Daltons[1]. The larger polymers are insoluble in sodium dodecylsulfate (SDS) buffers unless disrupted either mechanically by sonication or chemically by reducing agents[1]. The amount and molecular weight distribution of this unextractable component is correlated with flour functional properties[3].

Genes encoding the HMW-GS are inherited as tightly linked pairs, each pair encoding an x-type and a y-type subunit from the *Glu-1* homoeologous loci on the group 1 chromosomes of the A, B, and D genomes of hexaploid bread wheats[4]. By convention, individual HMW-GS are numbered and designated by genome origin and type. The x-type and y-type subunits are highly homologous in sequence and structure, differing chiefly in the lengths of their repeat and unique N-terminal domains[2]. Genetic studies have shown that differences in the quantity and/or in specific alleles of HMW-GS can account for significant differences in the bread-making quality among various cultivars[4]. Two subunits, Dx5 and Dy10, encoded by paired alleles of the D genome, have been associated with increased dough strength[4].

Although HMW-GS constitutes only 5% to 10% of total seed protein, their central role in determining the elasticity of wheat doughs makes members of this gene family important candidates for genetic engineering[5]. The development of reliable and reproducible systems of wheat transformation[6-8] now makes it possible to modify HMW-GS composition in order to investigate structural/functional relationships. Several fundamental questions to be addressed are: (1) Can transgenic HMW-GS genes be expressed at levels high enough to affect dough properties? (2) Will expression be stable over generations? And (3) Can the levels of HMW-GS relative to other seed proteins be increased, or is there a threshold that limits HMW-GS accumulation?

Results

Construction of a hybrid HMW-GS gene. In order to analyze modified wheat seed protein composition, the products of the introduced transgenes must be distinguishable from those of the endogenous genes. The cultivar "Bobwhite" was used for genetic transformation because of its high regeneration frequency[8]. We had available a set of fusions of various HMW-GS coding regions, constructed by Shani et al.[9], in the *Escherichia coli* expression vector, pET-3a (ref. 10). The hybrid subunits exhibited mobilities in SDS-PAGE distinct from those of the native HMW-GS[9]. One such construct, pET-3a-10/5, is a fusion between the Dy10 and Dx5 genes at a conserved HindIII site at the junction of the N-terminal and repetitive coding domains. The encoded protein consists of amino acids 1 to 124 of the mature Dy10 polypeptide followed by amino acids 130 to 848 of the mature Dx5 polypeptide (Fig. 1A). The central region of repeating amino acid motifs and C-terminal nonrepetitive domains are identical to those of Dx5. Compared to the native Dx5 subunit, the Dy10:Dx5 chimera h[a]
15 more amino acids in its N-terminal nonrepetitive r[e]
including five, instead of four, cysteines[2].

The hybrid coding region of pET-3a-10/5 was [c]

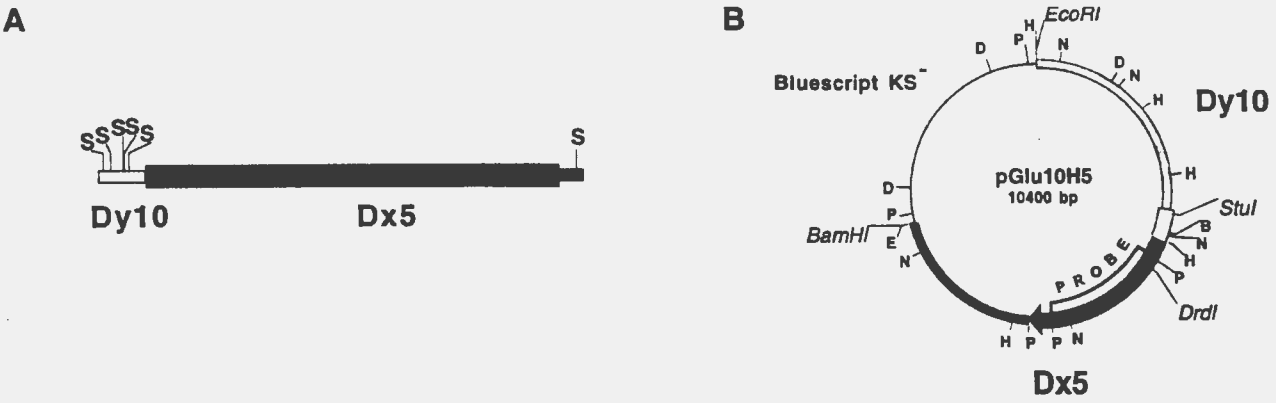


Figure 1. (A) Diagram of mature Dy10:Dx5 hybrid polypeptide encoded by pET-3a-10/5[9] and pGlu10H5. Amino acids from the Dy10 subunit are shown in gray and those from the Dx5 subunit in black. The N- and C-terminal unique regions are shown as thin boxes and the repeat region as a thicker box. The locations of the cysteine residues that form inter- or intramolecular disulfide bonds are labelled S. (B) Diagram of pGlu10H5. Segments of wheat DNA are shown as boxes, the gray boxes from the *Glu-D1-2b* gene[12] encoding Dy10 and the black boxes from the *Glu-D1-1d* gene[12] encoding Dx5. The coding region of the hybrid HMW-GS is shown as thicker boxes, the Bluescript KS⁻ vector by a thin line. The arrow indicates the direction of transcription and translation in wheat cells. The PvuII fragment used as the probe for the Dx5 coding region in Southern blot analyses is shown by the bracket. Restriction sites used in assembling the plasmid are designated by their full names, other sites by single letter abbreviations: B, BamHI; D, DrdI; E, EcoRI; H, HindIII; N, NcoI; and P, PvuII.

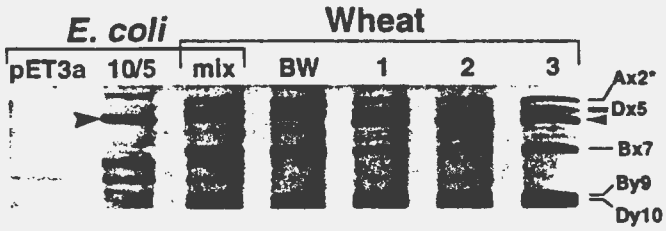


Figure 2. SDS-PAGE of *E. coli* and wheat endosperm protein extracts. Protein was extracted from either immature endosperm dissected from seeds of cultivar Bobwhite (BW) or transgenic plants (1, 2, and 3), or from *E. coli* containing the expression plasmids pET-3a[10] (pET3a) or pET-3a-10/5[9] (10/5). The "mix" lane contains a mixture of the Bobwhite endosperm and pET-3a-10/5 *E. coli* extracts to show the separation of the hybrid subunit (arrows) from endogenous wheat proteins. The positions of the native HMW-GS are labelled to the right.

flanking regions from the native Dy10 and Dx5 genes that contain the regulatory sequences needed for expression in wheat, resulting in the plasmid pGlu10H5 (Fig. 1B). The native Dy10 gene provides the promoter, transcription start site, the 5′ transcribed untranslated region, and the first 145 codons including those for the 21 amino acid signal peptide. The native Dx5 gene provides the remainder of the coding region, as well as the transcription termination and poly(A) addition signals.

Cotransformation of the hybrid HMW-GS gene and UBI:BAR. The pGlu10H5 plasmid was cotransformed with UBI:BAR[11] into immature embryos of cultivar Bobwhite by microprojectile bombardment[8]. Twenty-six independent lines were selected based on the ability of T_1 progeny embryos to germinate in the presence of 3 mg/L bialaphos 3 to 4 weeks after anthesis[8]. Proteins were extracted from the immature endosperm tissue corresponding to each germinated embryo and screened by SDS-PAGE for the presence of the hybrid glutenin (Fig. 2). The control lanes show that the migration of the hybrid HMW-GS is distinguishable from that of the native HMW-GS under these electrophoretic conditions. The hybrid HMW-GS is present in *E. coli* containing pET-3a-10/5 (lane 10/5) and absent from bacterial cells containing the pET-3a vector (lane pET3a) or wheat extracts from the immature endosperm of the untransformed parental cultivar Bobwhite (lane BW). The "mix" lane contains a mixture of 10/5 and BW extracts; the hybrid subunit can be seen to migrate ahead of the native Dx5 subunit and slightly behind two minor bands in the wheat seed extract. Protein extracts from endosperm of the three transgenic plants all contain the hybrid glutenin polypeptide. Of the 25 independent plants resistant to bialaphos, 18 (69%) showed expression of the hybrid subunit in T_1 generation seeds.

Coexpression stability. At each successive generation, 24 embryos were tested for their ability to germinate on bialaphos. Only resistant plants were carried forward to the next generation. Plants were tentatively identified as homozygous for a single UBI:BAR transgene locus if all their progeny exhibited resistance. The expression of the unselected pGlu10H5 transgenes also was assessed in successive generations. In 11 of 15 lines analyzed thus far, expression cosegregates with bialaphos-resistance, as shown for three generations of two different lines (Fig. 3). Analyses of other transgenic lines showed that the majority exhibit expression levels in successive generations about equal to or higher than those of their original T_1 seeds. Increases in hybrid HMW-GS accumulation, such as are evident between the T_1 and T_2 generations of line 10 (Fig. 3), are most likely due to the increase in transgene dosage of the presumptive homozygous progeny of T_1 plants, compared with the heterozygous progeny of T_0 plants.

Crosslinking ability of the hybrid subunit. To test whether or not the hybrid HMW-GS could crosslink by disulfide bonds and form high-molecular-weight polymers, the solubility of HMW-GS in SDS buffers with and without reducing agents was determined (Fig. 4A). In this test, both the hybrid and native subunits are only partially extracted by SDS buffer (lane S). Additional hybrid and native subunits are solubilized when the pellet is extracted in SDS buffer containing a reducing agent (lane I). In the transgenic sample, the proportion of hybrid HMW-GS present in the insoluble fraction is similar to the proportion of native subunits. This shows that a fusion of a y-type subunit N-terminus containing five cysteines with an x-type repeat region and C-terminus is capable of forming intermolecular disulfide bonds.

Relative expression levels. A comparison of the levels of hybrid HMW-GS accumulation in mature seeds of several transgenic plants is shown in Figure 4B. The hybrid HMW-GS is present at levels comparable to those of the individual endogenous HMW-GS in all the lines except 2 and 3. Extract 3 contains an extra polypeptide of unknown origin that migrates just ahead of the hybrid HMW-GS. Extract 2 is unusual in that accumulation of the native HMW-GS is low compared to the hybrid HMW-GS and to other

proteins, chiefly seed storage proteins (LMW-GS and gliadins), whose levels are about equal to those seen in the other lanes. In the other extracts, no decreases in expression of the native proteins are apparent.

In order to obtain a preliminary assessment of the levels of the hybrid HMW-GS relative to those of the native HMW-GS and other seed proteins, gels were scanned by densitometry. Although proteins of molecular weights <30 kD are lost under the conditions used to separate the hybrid HMW-GS from the nearby Dx5 subunit band, the other storage proteins that constitute the bulk of the endosperm remain on the gels and were used to normalize expression of the HMW-GS within a given lane (Table 1). The hybrid HMW-GS band constitutes 15% to 41% of the total HMW-GS in each extract except number 2. Even a single copy of the hybrid HMW-GS transgene (lines 1 and 3) can support significant expression levels, comparable to those of the native HMW-GS. In extract 2, overall levels of HMW-GS accumulation are reduced to about 70% of those of the Bobwhite control. In the other extracts, the accumulation of the hybrid HMW-GS does not appear to be at the expense of the endogenous HMW-GS. The apparent result, at least for these extracts, is an increase in total HMW-GS compared to the proportion of native subunits in the parental cultivar.

DNA analysis. Genomic Southern blot analysis is complicated by the presence of cross-reacting bands derived from the endogenous HMW-GS genes (Fig. 5 control lanes). The enzymes EcoRI (E) and NcoI (N) were chosen for DNA digestion because fragments derived from pGlu10H5 are predicted to differ in size from those of the endogenous genes known from previous genomic Southern blot analyses[2] or DNA sequence analyses. The Dx5-derived probe hybridizes to EcoRI fragments from the Bx7 gene at 19.2 kbp, from Dx5 at 8.7 kbp, from Ax2* at 6.8 kbp, and from the By9 gene at 15.8 kbp. EcoRI fragments derived from the other two y-type genes comigrate with the x-type gene fragments, the silent *Glu1-ay* (Ay null) at 8.2 kbp, and the Dy10 at 6.2 kbp. The identity of only two of the NcoI fragments in the control DNA could be assigned from sequence data: the most intensely hybridizing 2.1-kbp fragment is from the Dx5 gene[12] and the 3.3-kbp fragment is from the Ax2* gene (O. Anderson, unpublished result).

Several additional bands are evident in the lanes containing DNA from the transgenic plants (Fig. 5). Line 1 has the simpler pattern. The EcoRI lane shows one additional band (arrow) between the Dx5 and Ax2* bands of the size expected for an intact EcoRI fragment from plasmid pGlu10H5 (7.4 kbp). The NcoI lane contains one additional hybridizing band of 1.85 kbp (arrow), the size of the intact NcoI fragment from pGlu10H5 that would hybridize to the probe. DNA from line 4 contains several copies of these two fragments as well as others presumably derived from integrated copies of pGlu10H5 that have been rearranged or truncated such that the spacings of the relevant NcoI and EcoRI sites have been disrupted.

As expected, at least one intact copy of the NcoI fragment from pGlu10H5 homologous to the Dx5 coding region is present in each transgenic line expressing the hybrid HMW-GS. Among the eight transgenic lines shown in Figure 4B, lines 1 and 3 contain a single copy insert while the others contain 2 to 6 copies (Table 1). Expression levels are not strictly correlated with gene copy number and exhibit a narrower range of values.

Discussion

The experiments reported here show that changes in the composition and relative levels of HMW-GS can be effected by the addition of gene copies to the wheat genome. In this case, we used an unique coding region created by the fusion of two native genes in order to distinguish the products of the transgene from the native

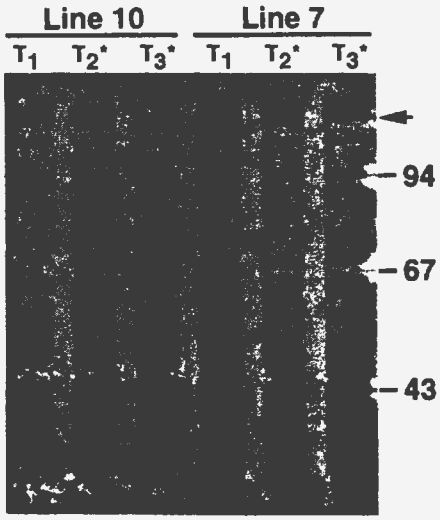


Figure 3. SDS-PAGE of proteins accumulated in seeds of three successive generations of transgenic wheat lines. Extracts were prepared from a sample of eight mature T_1, T_2, or T_3 seeds of plants from two different lines (number 7 is the same line shown in Fig. 4B and Table 1). The arrow marks the location of the hybrid HMW-GS. The positions of protein standards are indicated to the right by their molecular weights in kD.

endosperm proteins. In the analyses performed so far, the behavior of the hybrid subunit is similar to that of the native HMW-GS: the protein accumulates in endosperm tissue over the course of seed development, and at least a portion of it participates via disulfide bonds in the formation of SDS-insoluble polymers. Characterization of the assembly properties of this protein is in progress.

Expression of the hybrid HMW-GS is under the control of the native Dy10 5′ and Dx5 3′ flanking sequences. Even a single gene copy supports accumulation to levels comparable to those of the native HMW-GS genes (Fig. 4B, extract 1). Thus, the wheat HMW-GS genes can be expected to serve as sources of effective transcriptional control sequences suitable for expression of other proteins in the endosperm of transgenic wheat and other cereals.

A preliminary quantitative assessment by scanning densito-

Table 1. Densitometer analysis of relative HMW-GS transgene expression and copy number in individual wheat lines.

Line	Hybrid/total HMW-GS* (%)	HMW-GS levels transgenic/control†	Transgene coding region copies‡
1	20	1.4	1
2	55	0.7	5–6
3	15	1.5	1
4	41	1.5	4–5
5	24	1.4	2
6	26	1.3	2
7	24	1.3	3–4
8	20	1.4	3–4

*Average from densitometry scans of two gels (e.g. Fig. 4B). Each background-corrected value for the hybrid HMW-GS peak was divided by the background-corrected value for the sum of all six HMW-GS peaks in the same lane.
†The sum of all six HMW-GS peaks was divided by the background-corrected total scan of the same lane. The values obtained for each transgenic line were then divided by the value of the control.
‡Calculated from densitometry scans of NcoI fragments in autoradiographs of genomic Southern blot analyses (Fig. 5).

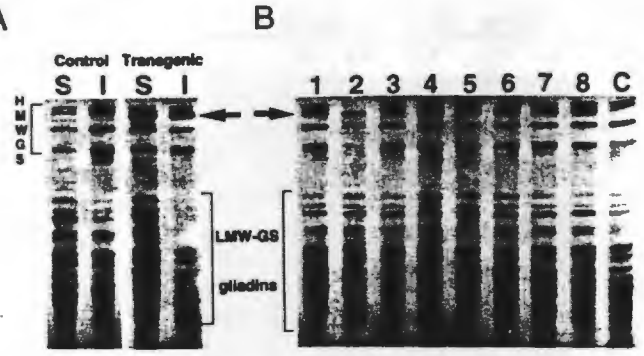


Figure 4. SDS-PAGE of proteins accumulated in mature seeds of the parental cultivar Bobwhite (control and C) and transgenic wheat lines (numbered lanes). Arrows mark the locations of the hybrid HMW-GS. The regions of the gels containing the abundant seed storage proteins—LMW-GS and gliadins—are indicated by brackets. Proteins smaller than 30 kD are not seen. (A) Proteins extracted before (S) and after (I) addition of reducing agent. Seeds from a control plant and a T_3 progeny plant of transgenic line 4 were extracted with SDS-PAGE loading buffer lacking reducing agent. After centrifugation at 16,000 g for 5 min, the proteins in the supernatant were reduced by bringing the concentration of β-mercaptoethanol to 2%. These extracts were loaded into the lanes marked S, SDS-soluble protein. The proteins in the pellet were solubilized with loading buffer containing 2% β-mercaptoethanol and a volume equivalent to the same percentage of extract loaded into the lanes marked I, SDS-insoluble proteins. The region of the gel containing the HMW-GS is indicated to the left by the bracket. (B) Extracts were prepared from eight mature seeds of single T_2 (Lines 3, 4, 6, and 8) or T_3 plants classified as homozygous because all of their progeny were resistant to bialaphos.

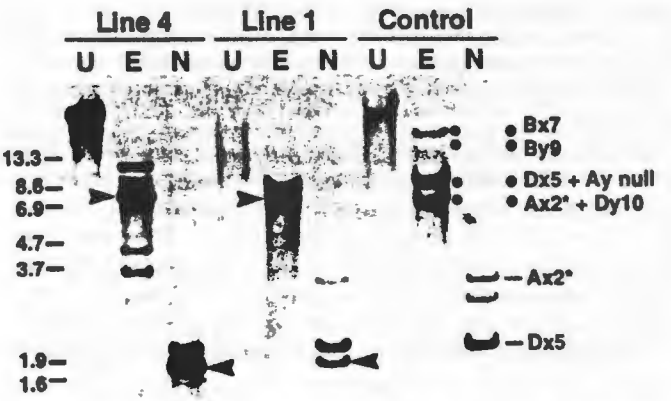


Figure 5. Genomic Southern blot analysis of transgenic wheat DNA. Genomic DNAs from transgenic T_2 (line 4) or T_3 plants (line 1 and control) were hybridized to a 1578 bp PvuII fragment from the coding region of the Dx5 gene (PROBE in Fig. 1B). The control plant is from a transgenic line that is resistant to bialaphos and contains copies of UBI:BAR, but none of pGlu10H5. Lanes marked U contain undigested DNA. Lanes marked E and N contain EcoRI and NcoI digests, respectively. The positions of bands derived from the endogenous HMW-GS genes that hybridize to this probe are indicated, where known, to the right as solid circles for the EcoRI fragments and as dashes for the NcoI fragments. The positions of DNA fragments of the sizes expected for the NcoI and EcoRI digests of pGlu10H5 are marked with arrows. The locations of molecular weight standards run on the same gel are indicated by their sizes in kbp to the left.

metry (Table 1) suggests that addition of HMW-GS gene copies raises the levels of HMW-GS accumulation relative to the other proteins (chiefly the other storage proteins) in mature seeds. The numbers in Table 1 are calculated from scans of one sample of eight seeds from a single greenhouse-grown plant of each line. True quantitation will require analysis of multiple replicate samples when a larger supply of homozygous seed is available and when the plants are grown in a number of environments in field tests. Other methods used to quantitate the levels of individual proteins in flour, such as RP-HPLC and capillary electrophoresis[13], may also be applied in order to assign more precise numerical values to the apparent increases.

Thus far, at least eight lines still exhibit expression at or near their original levels through their fourth greenhouse generation (data not shown). However, in four of the 15 wheat lines analyzed, fixation of the resistance trait in the T_2 or T_3 generations coincided with a loss or decline in expression levels of the unselected hybrid HMW-GS relative to those exhibited by the T_1 endosperm (data not shown). These lines might contain one or more integrated copies of pGlu10H5 that are not linked to the selected UBI:BAR transgene(s).

In line 2, synthesis of the hybrid subunit is associated with a decline in the levels of the endogenous HMW-GS. Interestingly, this line has the highest transgene copy number. Decreases in expression of homologous endogenous genes are often seen in transgenic plants when related genes are introduced into the genome[14]. Further experiments are in progress to test whether or not the decline in endogenous HMW-GS accumulation in this unusual line is due to some type of transgene-mediated suppression. If this is the case, it means that additional gene copies can result in either an increase or a decrease in total HMW-GS content. The latter might be desirable in wheat varieties in which reduced dough elasticity would be useful.

We have created a set of independent lines of transgenic wheat that synthesize substantial amounts of a hybrid seed storage protein under the control of native HMW-GS promoter and other flanking sequences. These results demonstrate the feasibility of manipulating the composition of wheat kernels by genetic engineering. Changes in both the levels and the types of seed storage proteins of wheat are now possible. Such modifications will enhance and extend the utilization properties of wheat flour.

Experimental protocol

Construction of the hybrid HMW-GS expression plasmid. The following three DNA fragments were isolated and combined in a single ligation reaction with the Bluescript KS⁻ (Stratagene, La Jolla, CA) vector cut with EcoRI and BamHI: (1) the 2800-bp fragment from a clone of the native Dy10 gene[12] beginning at the EcoRI site 5′ to the gene and ending at the StuI site 74 bp after the A in the start codon, (2) the 715-bp StuI/DrdI fragment that contains the junction region in pET-3a-10/5 (ref. 9); and (3) the 3800-bp fragment from a clone of the native Dx5 gene[12] that starts at the DrdI site 744 bp after the A in the start codon and ends at a BamHI site in the vector just outside the EcoRI site in the 3′ flanking region. A plasmid, pGlu10H5, with the correct structure was identified by restriction analysis (Fig. 1B). UBI:BAR[11] and pGlu10H5 plasmid DNAs were prepared by the alkaline lysis method using a Qiagen kit (Chatsworth, CA) for the final purification.

Wheat transformation. 12.5 μg of each DNA were cotransformed into wheat using microprojectile bombardment as described by Weeks et al.[4], except that the immature embryos were transferred to callus induction media containing 0.4 M mannitol for 4 h before and 20 h after bombardment, a treatment adapted by T. Weeks (unpublished results) for wheat from the results of Vain et al[15]. Forty-five independent callus pieces were selected for bialaphos resistance from 4980 embryos bombarded in 10 different experiments; each yielded one to five plants. From these, 26 independent bialaphos-resistant T_1 plants were obtained for an overall transformation frequency of 26 per 4980 embryos (0.5%).

Protein analysis. Protein extracts from *E. coli* and immature wheat endosperm were prepared by grinding the cells or tissue in SDS-PAGE sample buffer [66 mM Tris, pH 6.8, 3% (w/v) SDS, 3% (v/v) glycerol, 0.05% (w/v) bromphenol blue, 2% (v/v) β-mercaptoethanol]. Protein extracts from mature dry wheat seeds were prepared by pulverizing mixtures of eight whole seeds with a WIG-L-BUG™ apparatus (Crescent Dental Manufacturing Co., Lyons, IL) and dissolving the resultant powder in sample buffer. SDS-PAGE was performed according to the methods of Laemmli[16]

in separation gels composed of 10% (w/v) acrylamide, 0.05% (w/v) bis-acrylamide, and run until the dye front reached the bottom of the 15-cm gel. Gels were stained in a colloidal suspension of Coomassie brilliant blue G-250 (ref. 17). The lanes of stained gels were scanned using an Alpha Innotech (San Leandro, CA) IS-1000 Digital Imaging System. Two adjacent interlane spaces were also scanned as background for each lane. The two background readings were separately subtracted from the peak values in a lane (in the case of HMW-GS) or from the total lane scan, and then averaged.

DNA analysis. DNA was isolated from the leaves of transgenic plants of each line and assayed for the presence of transgenes by genomic Southern blot analysis as previously reported[4], except that approximately 10 μg of DNA were loaded into each lane. The probe consisted of the 1578 bp PvuII fragment from the coding region of Dx5 (bracketed in Fig. 1B). Autoradiographs of the blots were scanned using the Alpha Innotech IS-1000 Digital Imaging System. The NcoI bands 2093 bp and 1854 bp in length, derived from the native Dx5 and pGlu10H5 genes, respectively, are both completely homologous to the probe. The ratios of the densities of the 1854- to 2093-bp NcoI fragments were calculated to give the estimates of the copy numbers of intact hybrid HMW-GS coding regions in Table 1.

Acknowledgments

We thank Dafna Elrad, Hung Le, Joe Kuhl, and Eric Schlossberg for excellent technical assistance and Troy Weeks for advice on wheat transformation. We thank Susan Altenbach, Bob Graybosch, and Allan Zipf for critical reading of the manuscript. Bialaphos was kindly provided by Meiji Seika Kaisha, Ltd. (Yokohama, Japan). This work was supported by the Agricultural Research Service, CRIS # 5325-21430-001-00D. References to a company and/or product by the USDA are only for the purposes of information and do not imply approval or recommendation of the product to the exclusion of others that may be suitable.

1. MacRitchie, F. 1992. Physiochemical properties of wheat proteins in relation to functionality. *Adv. Food and Nutrition Res.* **36**:1–87.
2. Shewry, P.R., Halford, N.G., and Tatham, A.S. 1989. The high molecular weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role in wheat gluten structure and functionality, pp. 163–219 in *Oxford surveys of plant molecular and cell biology*. Miflin, B.J. (ed.). Oxford University Press, Oxford, UK.
3. Weegels, P.L., Hamer, R.J., and Schofield, J.D. 1996. Functional properties of wheat glutenin. *J. Cer. Sci.* **23**:1–18.
4. Payne, P.I. 1987. Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. *Annu. Rev. Plant Physiol.* **38**:141–153.
5. Shewry, P.R., Tatham, A.S., Barro, F., Barcelo, P., and Lazzeri, P. 1995. Biotechnology of breadmaking: Unraveling and manipulating the multi-protein gluten complex. *Bio/Technology* **13**:1185–1190.
6. Vasil, V., Castillo, A.M., Fromm, M.E., and Vasil, I.K. 1992. Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology* **10**:667–674.
7. Vasil, V., Srivastava, V., Castillo, A.M., Fromm, M.E., and Vasil, I.K. 1993. Rapid production of transgenic wheat plants by direct bombardment of cultured immature embryos. *Bio/Technology* **11**:1553–1558.
8. Weeks, J.T., Anderson, O.D., and Blechl, A.E. 1993. Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). *Plant Physiol.* **102**:1077–1084.
9. Shani, N., Steffen-Campbell, J.D., Anderson, O.D., Greene, F.C., and Galili, G. 1992. Role of amino- and carboxy-terminal regions in the folding and oligomerization of wheat high molecular weight glutenin subunits. *Plant Physiol.* **98**:433–441.
10. Rosenberg, A.H., Lade, B.N., Chui, D., Lin, S., Dunn, J.J., and Studier, F.W. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**:125–135.
11. Cornejo, M.-J., Luth, D., Blankenship, K.M., Anderson, O.D., and Blechl, A.E. 1993. Activity of a maize ubiquitin promoter in transgenic rice. *Plant Mol. Biol.* **23**:567–581.
12. Anderson, O.D., Greene, F.C., Yip, R.E., Halford, N.G., Shewry, P.R., and Malpica-Romero, J.-M. 1989. Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. *Nucleic Acids Res.* **17**:3156–3157.
13. Bietz, J.A. and Simpson, D.G. 1992. Electrophoresis and chromatography of wheat proteins: available methods, and procedures for statistical evaluation of the data. *J. Chromatography.* **624**:53–80.
14. Matzke, M.A. and Matzke, J.M. 1995. How and why do plants inactivate homologous (trans)genes? *Plant Physiol.* **107**:679–685.
15. Vain, P., McMullen, M.D., and Finer, J.J. 1993. Osmotic treatment enhances particle bombardment-mediated transient and stable transformation of maize. *Plant Cell Rep.* **12**:84–88.
16. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
17. Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **9**:255–262.

High efficiency transformation of maize (*Zea mays* L.) mediated by *Agrobacterium tumefaciens*

Yuji Ishida*, Hideaki Saito, Shozo Ohta, Yukoh Hiei, Toshihiko Komari, and Takashi Kumashiro†

*Plant Breeding and Genetics Research Laboratory, Japan Tobacco Inc., 700 Higashibara, Iwata, Shizuoka 438, Japan. †Present address: Agribusiness Division, Japan Tobacco Inc., 2-1 Toranomon 2-chome, Minato-ku, Tokyo 105, Japan. *Corresponding author (e-mail: KYR03634@niftyserve.or.jp).*

Received 29 December 1995; accepted 28 February 1996.

Transformants of maize inbred A188 were efficiently produced from immature embryos cocultivated with *Agrobacterium tumefaciens* that carried "super-binary" vectors. Frequencies of transformation (independent transgenic plants/embryos) were between 5% and 30%. Almost all transformants were normal in morphology, and more than 70% were fertile. Stable integration, expression, and inheritance of the transgenes were confirmed by molecular and genetic analysis. Between one and three copies of the transgenes were integrated with little rearrangement, and the boundaries of T-DNA were similar to those in transgenic dicotyledons and rice. F1 hybrids between A188 and five other inbreds were transformed at low frequencies.

Keywords: transformation, maize, *Agrobacterium tumefaciens*

Application of *Agrobacterium*-mediated method of gene transfer has until now been limited to dicotyledonous plants, although this method of gene delivery to higher plants has advantages, such as the transfer of relatively large segments of DNA with little rearrangement, and integration of low numbers of gene copies into plant chromosomes. Although monocotyledons are not the natural hosts of *Agrobacterium tumefaciens*[1], infection of maize and other cereals with *A. tumefaciens* has been attempted in various laboratories.

Competence of *A. tumefaciens* in infection of maize was first indicated in the studies of "agro-infection" by Grimsley et al.[2], in which cDNA of maize streak virus was delivered to maize plants by *A. tumefaciens* and the plants became systemically infected. Gould et al.[3] inoculated shoot apices of maize with *A. tumefaciens* and obtained a few transgenic plants, and Shen et al.[4] observed expression of a β-glucuronidase (GUS) gene delivered to maize shoots by *A. tumefaciens*. The studies of the *Agrobacterium* methods in other cereals also provided indications of successful transformation[5–7]. For example, Mooney et al.[7] produced transformed cells from wheat embryos cocultivated with *A. tumefaciens*, and Chan et al. obtained a few transgenic rice plants by inoculating immature embryos with *A. tumefaciens*[8,9]. However, the transformation frequency in these methods was rather low, and some of the studies did not provide sufficient molecular and genetic evidence of production of transgenic plants. Consequently, these methods have not been widely adapted.

Recently, Hiei et al.[8] reported a method to efficiently produce transgenic plants from rice calli cocultivated with *A. tumefaciens*. They claimed the choice of starting materials, tissue culture conditions, bacterial strains, and vectors were essential in efficient gene transfer. Here we describe a method to efficiently transform maize by cocultivation of immature embryos with *A. tumefaciens*. We produced a large number of transformants of A188 and demonstrated stable integration, expression, and inheritance of transgenes.

Results

Infection and selection. Immature embryos of maize inbred line A188 were cocultivated with 1.0×10^9 cfu/ml of *A. tumefaciens* LBA4404(pSB131) (Fig. 1) in LS-AS medium. The embryos were between 1.0 and 1.2 mm in length, and between 80% and 100% of the immature embryos expressed GUS after cocultivation (Fig. 2A). The immature embryos were transferred to a selection medium, LSD1.5, containing phosphinothricin (PPT). PPT-resistant calli emerged from between 38% and 90% of the immature embryos (Table 1, Fig. 2B) and expressed GUS uniformly (Fig. 2C). These calli were typical type I calli, which are compact clusters of relatively organized cells[9]. A large number of shoots were regenerated from the PPT-resistant calli that were transferred to a regeneration medium containing PPT (Fig. 2D), and showed strong expression of GUS in the leaves (Fig. 2E). PPT-resistant, GUS-positive plants were obtained from >11.8% of the immature embryos initially cocultivated (Table 1). Efficiency of transformation was remarkably consistent from experiment to experiment, and as many as 44 independent transgenic plants were obtained in a single trial (experiment no. 6) in the best case. In the following experiments, all of the parameters were identical to those above unless otherwise indicated.

Other tissue types. Type 1 calli[9] produced from immature embryos of A188 were infected with LBA4404(pSB131). Although most of the calli expressed GUS after cocultivation, few PPT-resistant cells were obtained during selection. Shoot tips of A188 and cells in a suspension culture from cultivar 'Black Mexican Sweet' (BMS) were also infected with LBA4404(pSB131), but GUS expression was only observed in a few tissue pieces. Therefore, infection of these tissues was not studied further.

Stage of immature embryos. Immature embryos of A188 between 1.5 and 2.0 mm and between 2.0 and 2.5 mm in length were also tested. GUS expression was detected in most of the immature embryos after cocultivation, and the PPT-resistant calli derived from such immature embryos showed uniform expression of GUS. Stable transformants, however, were obtained at a very low frequency (Table 1).

Concentration of inoculum. Infection with $>5.0\times10^9$ cfu/ml

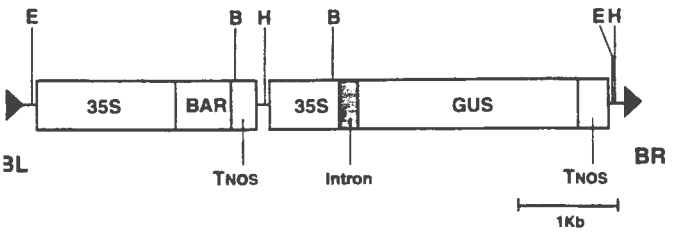


Figure 1. T-DNA of pSB131. Abbreviations: BR, right border; BL, left border; GUS, β-glucuronidase; BAR, phosphinothricin acetyltransferase; 35S, 35S promoter; TNOS, 3′ signal of nopaline synthase; B, BamHI; E, EcoRI; H, HindIII

of LBA4404(pSB131) resulted in large clusters of GUS-expressing cells in immature embryos after cocultivation, but no transformants were obtained. When the bacterial concentration was decreased to 1.0×10^8 cfu/ml, PPT-resistant cells were obtained from <10% of the immature embryos. PPT-resistant cells were not obtained when the bacterial concentration was $<1.0\times10^7$ cfu/ml.

Media based on the N6 medium. Many media for maize tissue culture are derived from the N6 medium[9–14]. Immature embryos of A188 were cocultivated with LBA4404(pSB131) in N6-AS medium and transferred to a selection medium, N6D1.5, containing PPT. Although GUS expression was observed in 80% to 100% of the immature embryos after cocultivation, PPT-resistant calli were not obtained. Therefore, media containing LS salts[15] are superior to N6-based media, and were used in all further experiments.

Hygromycin resistance. Another selective-marker, a hygromycin-resistance gene, was tested. When immature embryos of inbred A188 were infected with LBA4404(pTOK233)[8], most of the immature embryos expressed GUS after cocultivation. The embryos were then cultured on a hygromycin-containing medium, and resistant calli were obtained 8 weeks after cocultivation. Plants were regenerated from the calli on hygromycin-containing medium and showed strong expression of GUS in the leaves. The frequencies of transformation by hygromycin selection were consistent, and transformants were obtained from 5% to 10% of the immature embryos (Table 1.).

Other strains of *A. tumefaciens*. Immature embryos of A188 were infected with LBA4404(pIG121Hm)[8] and other strains of *A. tumefaciens*. Expression of GUS after the cocultivation was found at high frequencies, but the level of expression was considerably lower than in those immature embryos infected with LBA4404(pSB131) or LBA4404(pTOK233). Thus, use of other strains was not studied further.

Other genotypes of maize. Five inbred lines (W117, W59E, A554, W153R, and H99) and five F1 hybrids (W117 x A188, W59E x A188, A554 x A188, W153R x A188, and H99 x A188) were examined. Immature embryos of these genotypes were cocultivated with LBA4404(pSB131). GUS expression was detected in most of the immature embryos after cocultivation, and transformed plants were obtained from all F1 hybrids. The frequencies of transformation of F1 hybrids varied from 0.4% to 5.3% (Table 1), whereas no transformants were obtained from the inbred lines.

Characterization of the plants in the R0 generation. A total of 120 independent, PPT-resistant, GUS-positive plants of A188 from immature embryos infected with LBA4404(pSB131) were grown in a greenhouse. Almost all of the plants were normal in morphology (Figs. 2F, 2G, and 2H) and the majority (about 70%) of them produced as many seeds as seed-derived control plants by self-pollination (Fig. 2I).

Thirty-three of the transformed plants were analyzed by Southern hybridization. Isolated DNA was digested with BamHI or EcoRI and allowed to hybridize with *bar* and *gus* probes. Both genes were detected in all of the R0 plants analyzed, whereas no hybridization signal was detected in the nontransformed plants

Table 1. Efficiency of maize transformation.

Variety	Vector	Size of immature embryos (mm)	Experiment no.	Number of immature embryos: Inoculated (A)	Produced antibiotic-resistant callus	Produced antibiotic-resistant plants	Produced antibiotic-resistant, GUS+ plants (B)	Frequency (B/A, %)
A188	pSB131	1.0–1.2	1	44	28	9	6	13.6
			2	52	33	10	7	13.5
			3	51	46	13	7	13.7
			4	70	56	26	14	20.0
			5	76	30	12	9	11.8
			6	369	200	71	44	11.9
			7	121	46	33	20	16.5
			8	27	15	8	5	18.5
			9	36	26	18	11	30.6
			10	77	38	32	16	20.8
		1.5–2.0	1	57	11	0	0	0.0
		2.0–2.5	1	156	33	3	2	1.3
W117 x A188	pSB131	1.0–1.2	1	112	36	8	4	3.6
			2	114	26	10	6	5.3
W59E x A188			1	104	44	1	1	1.0
A554 x A188			1	247	46	7	5	2.0
W153R x A188			1	284	69	2	1	0.4
H99 x A188			1	219	18	4	3	1.4
A188	pTOK233	1.0–1.2	1	72	15	8	7	9.7
			2	22	5	2	2	9.1
			3	22	1	1	1	4.5
			4	19	2	2	1	5.3

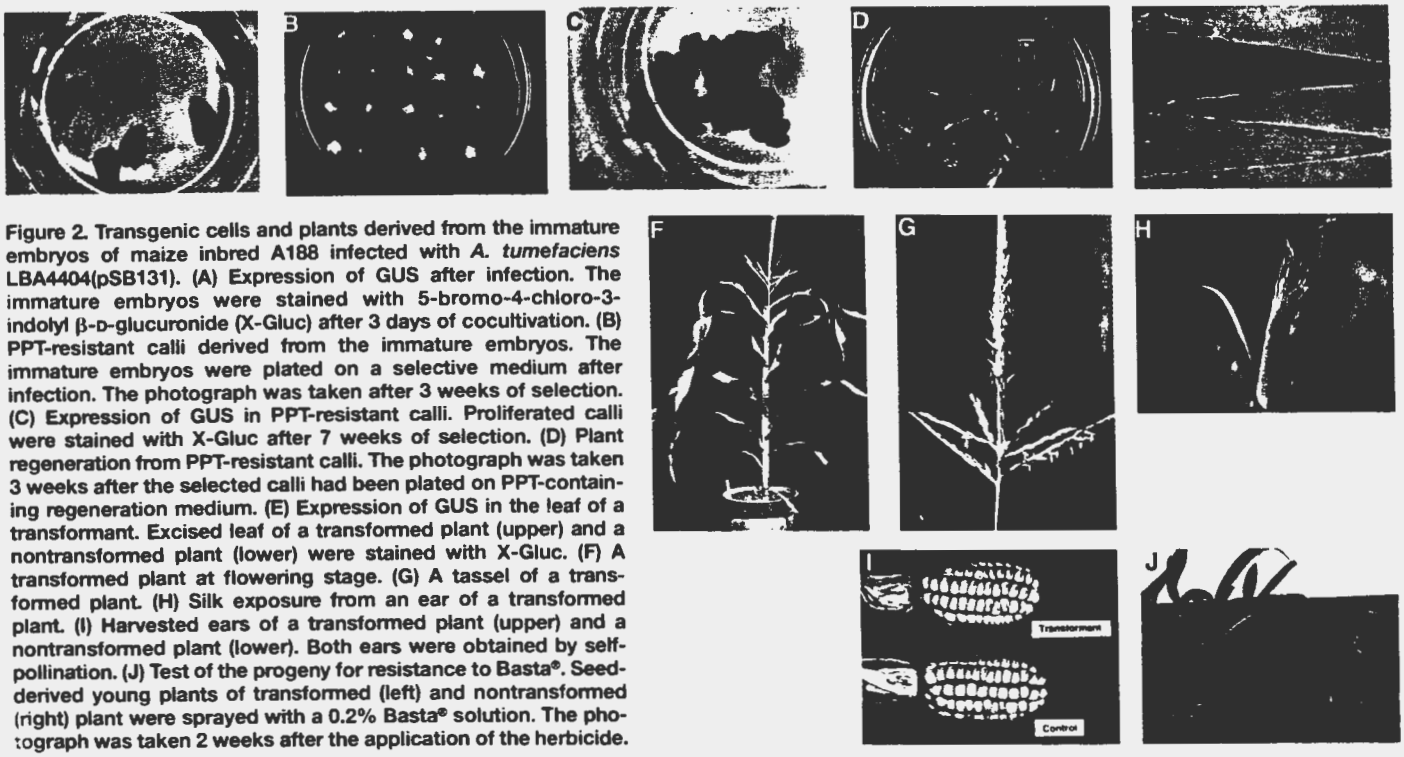


Figure 2. Transgenic cells and plants derived from the immature embryos of maize inbred A188 infected with *A. tumefaciens* LBA4404(pSB131). (A) Expression of GUS after infection. The immature embryos were stained with 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc) after 3 days of cocultivation. (B) PPT-resistant calli derived from the immature embryos. The immature embryos were plated on a selective medium after infection. The photograph was taken after 3 weeks of selection. (C) Expression of GUS in PPT-resistant calli. Proliferated calli were stained with X-Gluc after 7 weeks of selection. (D) Plant regeneration from PPT-resistant calli. The photograph was taken 3 weeks after the selected calli had been plated on PPT-containing regeneration medium. (E) Expression of GUS in the leaf of a transformant. Excised leaf of a transformed plant (upper) and a nontransformed plant (lower) were stained with X-Gluc. (F) A transformed plant at flowering stage. (G) A tassel of a transformed plant. (H) Silk exposure from an ear of a transformed plant. (I) Harvested ears of a transformed plant (upper) and a nontransformed plant (lower). Both ears were obtained by self-pollination. (J) Test of the progeny for resistance to Basta®. Seed-derived young plants of transformed (left) and nontransformed (right) plant were sprayed with a 0.2% Basta® solution. The photograph was taken 2 weeks after the application of the herbicide.

(Figs. 3A, 3B, 3C, and 4). As expected from the T-DNA map of pSB131 (Fig. 1), digestion of the DNA with BamHI yielded various band sizes longer than 1.9 kb that hybridized to the *bar* probe, and various band sizes longer than 2.3 kb that hybridized to the *gus* probe (Figs. 3A and 3B) . The number of hybridizing bands reflected the copy number of the transgenes in the plant genome, which varied from one to three (Figs. 3A, 3B, and Table 2). Nineteen of the 33 plants contained a single copy of the *bar* gene, and 23 plants contained a single copy of the *gus* gene. Because the EcoRI sites are located very close to either border of the T-DNA of pSB131, detection of a 5.4-kb EcoRI fragment in this analysis strongly indicated integration of an intact copy of the T-DNA. (Fig. 3C). Thirty-one of the 33 plants contained a 5.4-kb EcoRI fragment hybridizing to the *bar* probe (Table 2). Therefore, it is likely that approxi-mately 40% of the transformants carried a singly copy of the intact T-DNA.

Analysis of the T-DNA boundaries. The junction regions of introduced DNA and plant genome were cloned from several plants that contained a single copy of the 5.4-kb EcoRI fragment by the inverse PCR method[16]. Sequence analysis revealed that the junctions were located in or near the 25-bp repeats (Fig. 5). This observation is similar to the results from the analysis of the junctions in dicotyledons[17,18] and rice[8]. However, only four of the 10 right junctions sequenced were at the site found in tobacco transformants, and it is not clear why the right boundaries in maize appeared less precise than those in dicotyledons and rice.

Inheritance of marker genes. Selfed progeny of 40 of the 120 transformants of A188 grown in the greenhouse were examined for PPT resistance and GUS expression (Table 3). Resistant and sensitive seedlings were distinguishable 6 days after the application of Basta®. The sensitive plants died within 2 weeks after the treatment, while the resistant plants were as healthy as nontreated plants (Fig. 2J). PPT resistance and GUS expression were strongly linked and a segregation ratio of 3:1 for both traits (resistant:sensitive and positive:negative) was observed for 28 of the 40 lines. A few lines showed strange segregation ratios of 1:1 and 1:3 (Table 3).

DNA was extracted from the R1 progeny of transformants 131, 238, 248, and 249, shown in Tables 2 and 3, and analyzed by Southern hybridization. The *bar* gene and the *GUS* gene were present in the PPT-resistant, GUS-positive progeny and absent from the sensitive, negative progeny (Fig. 5. The data for *bar* are not presented.). The bands that were identical in size to the bands detected in the R1 plants were also present in their respective parents.

Discussion

The method of maize transformation reported here is efficient and reproducible. Only 10 weeks were needed to obtain transformed plants from infected immature embryos, and several lines of evidence show that the transgenes are stably incorporated into maize genome. Clear Mendelian transmission of the T-DNA to the progeny was demonstrated by genetic analysis. Drug-resistance and GUS expression were tightly linked, and segregation of the T-DNA was confirmed by Southern hybridization. Sequence analysis of the junctions between T-DNA and plant DNA in the maize transformants revealed that T-DNA boundaries in maize were similar to those in dicotyledons and rice.

In *Agrobacterium*-mediated gene transfer, expression of DNA segments in *Agrobacterium* attached to inoculated tissues or in other contaminating microorganisms needs to be carefully distinguished from expression of integrated foreign DNA. Here we take advantage of a *GUS* gene that contains in the coding region an intron that is not expressed in bacterial cells[19]. Thus, the strong expression of GUS we observed in the immature embryos after the cocultivation with *A. tumefaciens* and in PPT- or hygromycin-resistant maize plants was not due to bacterial contamination. A large number of transformants were analyzed by Southern hybridization, and the size of BamHI fragments hybridized to the probes differed from plant to plant, indicating random insertion of the transgenes into maize chromosomes.

Various transformation techniques have often been associated with aberrations in morphology, fertility, and other agronomically important characteristics[20,21]. In this study, almost all of the 120 independent transgenic maize plants characterized in detail were

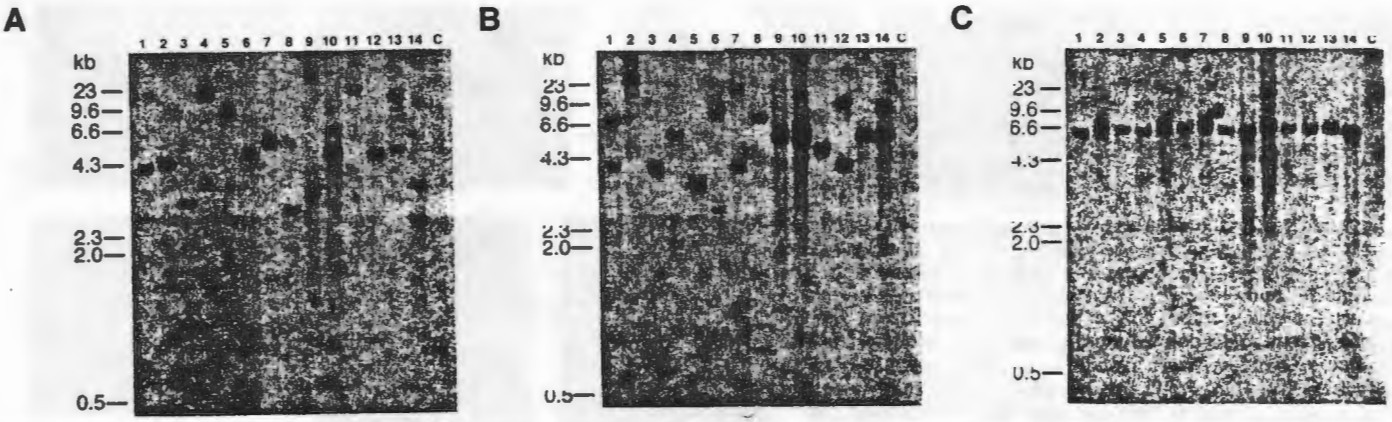


Figure 3. Southern blot analysis of transformed plants (R0 generation). DNA extracted from PPT-resistant and GUS-positive plants was digested with BamHI (A & B) or EcoRI (C), and allowed to hybridize to the *bar* (A, C) or *gus* (B) probe. Lane C, non-transformed control plant; lanes 1–14, transformed plants (No. 176, 185, 187, 191, 194, 197, 198, 238, 239, 241, 244, 245, 248, and 249 shown in Table 2) regenerated from PPT-resistant calli, which were derived from independent immature embryos infected with LBA4404(pSB131).

Table 2. Copy number of transgenes in maize transformants produced by LBA4404(pSB131).

	Copy number		
Transformant	*bar*	GUS	5.4kb EcoRI*
2	2	1	+
3	2	2	+
7	2	1	+
23	2	2	+
25	3	1	+
33	2	1	+
42	1	1	+
47	1	1	+
131	2	2	+
176	1	2	+
185	1	1	+
187	1	1	+
191	1	1	+
194	1	1	+
197	1	2	+
198	1	2	+
238	1	1	+
239	2	1	+
241	3	3	+
244	1	1	+
245	1	2	+
248	2	1	+
249	2	2	+
252	1	1	−
253	1	1	−
258	1	1	+
263	2	1	+
282	1	1	+
289	1	1	+
293	1	1	+
295	1	1	+
291	2	2	+
294	2	1	+

*Detection of 5.4-kb EcoRI fragments, which contained most of the T-DNA region, in the Southern blot analysis.

Table 3. Genetic analysis of independent transformants produced by LBA4404(pSB131).

	PPT resistance				GUS expression				
	Number of plants in R1 generation		Ratio		Number of plants in R1 generation			Ratio	
Transformant (R0)	R	S	R:S	χ^2	+	C	−	+:−	χ^2
1	37	9	3:1	0.72	23		23		
5	39	13	3:1	0.00	37		15	3:1	0.41
6	46	5	15:1	1.10	42		9	3:1	1.47
24	36	15	3:1	0.53	33		18	3:1	2.88
53	15	41			14		42		
55	39	17	3:1	0.86	40		16	3:1	0.38
68	38	16	3:1	0.62	38		16	3:1	0.62
75	45	10	3:1	1:36	3	33	19		
76	34	21			28	6	21		
77	48	7			48		7		
79	34	22			34		22		
80	44	12	3:1	0.38	44		12	3:1	0.38
83	42	14	3:1	0.00	41		15	3:1	0.10
91	38	15	3:1	0.31	38		15	3:1	0.31
93	46	18	3:1	0.33	46		18	3:1	0.33
102	19	5	3:1	0.22	19		5	3:1	0.22
105	39	15	3:1	0.22	36		16	3:1	0.92
115	43	13	3:1	0.10	43		13	3:1	0.10
123	42	12	3:1	0.22	42		12	3:1	0.22
124	35	14	3:1	0.33	35		14	3:1	0.33
125	37	9	3:1	0.72	37		9	3:1	0.72
126	37	12	3:1	0.01	37		12	3:1	0.01
131	49	1	15:1	1.54	49		1	15:1	1.54
133	41	13	3:1	0.02	41		13	3:1	0.02
134	29	9	3:1	0.04	29		9	3:1	0.04
136	31	19			9	22	19		
139	25	15	3:1	3.33	25		15	3:1	3.33
140	38	12	3:1	0.03	38		12	3:1	0.03
144	39	17	3:1	0.86	39		17	3:1	0.86
145	51	1	15:1	1.66	49		2	15:1	0.47
154	37	14	3:1	0.16	37		14	3:1	0.16
156	28	8	3:1	0.15	27		8	3:1	0.09
191	14	8	3:1	1.52	14		8	3:1	1.52
197	17	8	3:1	0.65	17		8	3:1	0.65
238	20	10	3:1	1.11	20		10	3:1	1.11
241	22	5	3:1	0.60	22		5	3:1	0.60
244	24	6	3:1	0.40	24		6	3:1	0.40
245	26	6	3:1	0.67	26		6	3:1	0.67
248	24	5	3:1	0.93	24		5	3:1	0.93
249	27	3	15:1	0.72	27		3	15:1	0.72

R, resistant; S, sensitive; C, chimeric expression. Ratios that give the smallest χ^2 values are shown together with the χ^2 values.

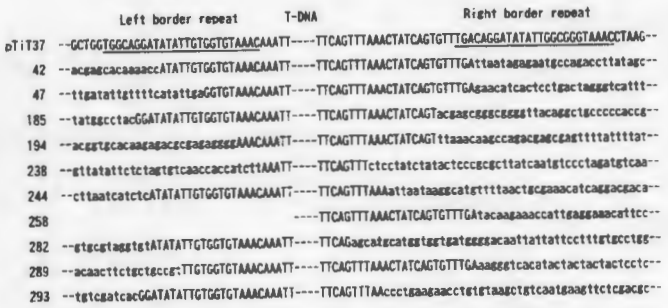


Figure 4. Sequence analysis of T-DNA/plant DNA junctions. Sequences of the junctions found in selected maize transformants are shown below the sequences of the T-DNA borders of pTiT37. Sequences presumably originated from maize genomic DNA are shown in lowercase letters. As the Southern blot analysis indicated these plants each contained a single copy of the T-DNA; the left and right border sequences in a plant possibly corresponded to the ends of a single T-DNA segment and are presented here in such a way. The left junction in transformant 258 was not determined.

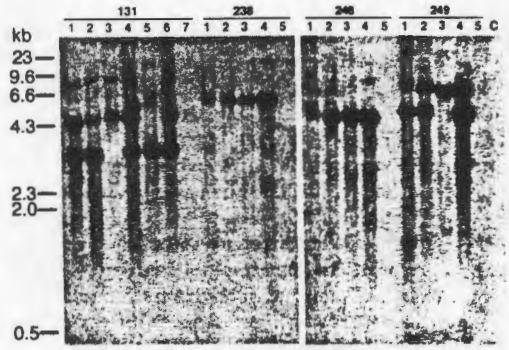


Figure 5. Southern blot analysis of the R1 progeny of transformed plants 131, 238, 248, and 249 shown in Table 2 and 3. DNA extracted from R0 plants (lane 1); PPT-resistant, GUS-positive R1 progeny (lanes 2–6 for transformant 131, lanes 2–4 for transformants 238, 248, and 249); PPT-sensitive, GUS-negative R1 progeny (lane 7 for transformant 131, lane 5 for transformants 238, 248, and 249); and nontransformed control plant (lane C) was digested with BamHI, fractioned by electrophoresis, transferred to a nylon membrane, and allowed to hybridize to the *gus* probe.

normal in morphology, and 70% were fully fertile. Furthermore, Southern hybridization and sequence analysis of T-DNA boundaries revealed that a majority (around 70%) of the maize transformants contained a single copy of the T-DNA with no notable rearrangements, and no transformants contained more than three copies of the T-DNA.

Factors that affect the efficiency of transformation include the types and stages of maize tissues infected, the concentration of *A. tumefaciens*, composition of the media for tissue culture, selection marker genes, kinds of vectors, and the maize genotype. This multiplicity of factors is probably the reason that transformation methods mediated by *A. tumefaciens* have not been readily developed. It was relatively easy to find conditions for good GUS expression after the cocultivation, but drug-resistant cells were selected in only a few instances. Therefore, GUS expression data did not necessarily help adjust infection parameters. It is possible that the main hurdle in transformation was not in delivery of DNA fragments into plant cells, but in recovery of cells that acquired the T-DNA in their chromosomes.

pTOK233 and pSB131 belong to a class of vectors called super-binary vectors. These vectors carry the *virB*, *virC*, and *virG* of A281, a strain highly efficient in transformation of higher plants[24]. It is evident that super-binary vectors are very useful in maize transformation. Although pSB131 was apparently higher than pTOK233 in the efficiency of transformation, this may be because various parameters were first optimized for PPT selection.

The present study and the previous study of rice transformation by *Agrobacterium*[8] have provided strong support for the hypothesis that T-DNA is transferred from *Agrobacterium* to dicotyledons and monocotyledons by an identical molecular mechanism. Therefore, being monocotyledonous is no longer a reason to restrict the application of *Agrobacterium*-mediated gene transfer techniques to other important cereal crops.

able 4. Media for culture of maize tissues.

Medium	Composition
LS-inf	LS major salts and LS minor salts, 0.5 mg/L nicotinic acid, 0.5 mg/L pyridoxine HCl, 1.0 mg/L thiamine HCl, 100 mg/L myo-inositol, 1.0 g/L casamino acid, 1.5 mg/L 2,4-D, 68.5 g/L sucrose, 36.0 g/L glucose, pH 5.2
LS-AS	LS major salts, LS minor salts, 0.5 mg/L nicotinic acid, 0.5 mg/L pyridoxine HCl, 1.0 mg/L thiamine HCl, 100 mg/L myo-inositol, 700 mg/L L-proline, 1.5 mg/L 2,4-D, 20 g/L sucrose, 10 g/L glucose, 500 mg/L MES, 100 μM acetosyringone, 8 g/L agar, pH 5.8
LSD1.5	LS-AS medium without glucose and acetosyringone, plus 250 mg/L cefotaxime, pH 5.8
LSZ	LSD1.5 medium without 2,4-D, plus 5.0 mg/L zeatin, pH 5.8
' '2LSF	Half-strength LS major salts of LSD1.5 medium without 2,4-D and L-proline, pH 5.8
-inf	N6 major salts, N6 minor salts and N6 vitamins, 1.0 g/L casamino acid, 1.5 mg/L 2,4-D, 68.5 g/L sucrose, 36.0 g/L glucose, pH 5.2
N6-AS	N6 major salts, N6 minor salts, N6 vitamins, 700 mg/L L-proline, 1.5 mg/L 2,4-D, 20 g/L sucrose, 10 g/L glucose, 500 mg/L MES, 100 μM acetosyringone, 8 g/L agar, pH 5.8
N6D1.5	N6-AS medium without glucose and acetosyringone, plus 250 mg/L cefotaxime, pH 5.8

Experimental protocol

Plant materials. Maize inbred lines A188, W117, W59E, A554, W153R, H99, and cultiver BMS were supplied from the National Institute of Agribiological Resources of Japan. F1 hybrids were obtained by cross-pollination in a greenhouse. Immature embryos of 1.0–1.2 mm in length were aseptically excised from kernels of plants grown in a greenhouse. Such immature embryos were generally obtained between 9 and 14 days after pollination (DAP), depending on environmental factors. For the study of optimal stages, immature embryos of 1.5–2.0 mm (11–16 DAP) and of 2.0–2.5 mm (13–18 DAP) were prepared. Type I calli of A188 and a suspension culture of BMS cells were prepared according to the procedure previously described[21].

Bacterial strains and plasmids. *A. tumefaciens* strain LBA4404(pTOK233) has been previously described[9]. The T-DNA of pTOK233 contained a hygromycin-resistance gene (*hpt*), a kanamycin-resistance gene, and a gene for GUS, which has an intron in the N-terminal region of the coding sequence and is connected to the 35S promoter of cauliflower mosaic virus[19]. This *intron-gus* gene expresses GUS activity in plant cells but not in the cells of *A. tumefaciens*[19]. pSB131 was constructed as follows: A PPT-resistance gene (*bar*) connected to the 35S promoter was excised as a 2.2-kb HindIII-EcoRI fragment from pDE110 (ref. 25) and inserted between the HindIII and EcoRI sites of pTOK246 (manuscript in preparation), which consisted of the origin of replication of pBR322, a spectinomycin-resistance gene, and the border fragments of T-DNA, to generate pSB25. The *intron-gus* was transferred as a 3.1-kb HindIII fragment from pGL2-IG[9] to pSB25 to give pSB31. pSB31 was then introduced to *A. tumefaciens* strain LBA4404(pSB1) (manuscript in preparation) by bacterial mat-

gs[26]. Bacteria carrying the cointegrate from pSB1 and pSB31 are designated B131. pSB1 is a wide host range plasmid that contained a region of homol-y to pSB31 and a 15.2-kb KpnI fragment from the virulence region of TiBo542. Thus, pSB131 contained the *bar* and *intron-gus* in the T-DNA.

Infection. LBA4404(pSB131) and LBA4404(pTOK233) were grown for days on YP medium (5 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl, 15 g/L gar, pH 6.8) supplemented with 50 mg/L spectinomycin (for pSB131) or 0 mg/L hygromycin (for pTOK233). The bacteria were collected with a platinum loop and suspended at a density of 1.0×10^9 cfu/ml in LS-inf medium or N6-inf medium. When necessary, bacterial suspensions of different densities were prepared. The immature embryos were washed once with LS-inf or N6-inf media. The immature embryos were immersed in the bacterial suspension, stirred for 30 sec with a vortex mixer (Vortex Genie 2, Scientific Industries) for thorough immersion, and allowed to stand for 5 min. No apparent disruption or wounding of the immature embryos was observed. The immature embryos were cultured on LS-AS medium or N6-AS medium in the dark at 25°C for 3 days (Table 4). During the incubation, the embryo axes were in contact with the medium and the scutella were exposed to air.

Selection and regeneration of transformants. After the cocultivation, the immature embryos were transferred to LSD1.5 or N6D1.5 media supplemented with 5 mg/L PPT (for infection with pSB131) or 10 mg/L hygromycin (for infection with pTOK233) for selection of transformed cells. After 2 weeks of incubation at 25°C in the dark, the immature embryos were subcultured on LSD1.5 or N6D1.5 media supplemented with 10 mg/L PPT or 30 mg/l hygromycin at 25°C in the dark for 3 weeks. Clusters of cells that proliferated from the immature embryos and showed the characteristics of the type I calli[9] were excised with scalpel and cultured on media of the same composition at 25°C in the dark for 3 weeks. Calli proliferated from the culture were excised again and cultured on LSZ medium supplemented with 5 mg/L PPT or 30 mg/L hygromycin at 25°C under continuous illumination (about 50 μmol m^{-2} sec^{-1}). All media for selection and regeneration contained 250 mg/L cefotaxime for elimination of *A. tumefaciens*. Regenerated plants were transferred to 1/2LSF medium, and incubated under the same condition for 2 weeks (Table 4). The plants were transferred to soil in pots and grown in a greenhouse.

Assay for GUS activity. Expression of GUS in maize calli and plants was examined by a calorimetric assay using substrate 5-bromo-4-chloro-3-indolyl β-D-glucuronide (X-Gluc) described by Hiei et al.[8].

Test of the progeny for resistance to PPT. The selfed progeny (R1 generation) of transformed plants were grown for 8 days in a greenhouse, and 0.2% Basta® (Hoechst, Frankfurt, Germany) solution was applied to the leaves with a writing brush or sprayed to the leaves. Basta® is a commercial formulation of glufosinate, which is the ammonium salt of PPT. Resistance was scored 6 days after the treatment.

Isolation of DNA and Southern hybridization. DNA was extracted from leaf tissues of R0 and R1 plants by the procedure described by Komari et al.[27]. Ten mg of DNA were digested with BamHI or EcoRI and fractionated on a 0.8% agarose gel by electrophoresis at 1.5 V/cm for 15 h. Southern hybridization was carried out as described by Sambrook et al.[28]. The *bar* probe was prepared by PCR from pSB25 using primers 5′-ATGGACCCA-GAACGACGCCCG-3′ and 5′-TCAGATCTCGGTGACGGGCAG-3′. The *gus* probe was prepared by PCR from pBI221 (ref. 29) using primers 5′-ATGT-TACGTCCTGTAGAAAC-3′ and 5′-ATGGTGCGCCAGGAGAGTTG-3′. The reaction mixture (50 µl) for PCR consisted of 1 ng of template DNA, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.2 mM each of dGTP, dATP, dTTP and dCTP, 1 unit of Taq DNA polymerase and 10 pmol each of primers. Thermal cycling for 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C was performed for 35 cycles.

Sequencing of border regions of the inserted T-DNA. Junction regions of the introduced T-DNA and maize genomic DNA were analyzed using an inverse PCR method[18]. Genomic DNA was digested with either one of BamHI, SalI, XbaI, and SacI, circularized by self-ligation, relinearized by digestion with EcoRI and HindIII, and used as a template. PCR was performed as described above except for the use of 250 ng of template DNA. Primers for analysis of the right boundaries were I-1: 5′-CGTTGCG-GTTCTGTCAGTTCCA-3′, GUS: 5′-TCACGGGTTGGGGTTTCTAC-3′, nos: 5′-ATCATCGCAAGACCGGCAAC-3′, and primers for analysis of the left boundaries were LS1: 5′-TCAGTACATTAAAAACGTCCGCA-3′, bar: 5′-CAGCTGGACTTCAGCCTGCC-3′, nos-F: 5′-GGTGTCATCTATGTTAC-TAG-3′. The amplified fragments were subcloned into pCRII (Invitrogen, San Diego, CA) and sequenced by Applied Biosystems (Foster City, CA) 373A DNA sequencer.

Acknowledgments

The authors thank Jan Leemans for providing us with pDE110, Tomoaki Kubo and Sumio Iwai for helpful discussions and advice.

1. De Cleene, M. and Deley, J. 1976. The host range of crown gall. *Bot. Rev.* **42**:389–466.
2. Grimsley, N., Hohn, T., Davies, J.W., and Hohn, B. 1987. *Agrobacterium*-mediated delivery of infectious maize streak virus into maize plants. *Nature* **325**:177–179.
3. Gould, J., Devey, M., Hasegawa, O., Ulian, E.C., Peterson, G., and Smith, R.H. 1991. Transformation of *Zea mays* L. using *Agrobacterium tumefaciens* and the shoot apex. *Plant Physiol.* **95**:426–434.
4. Shen, W.-H., Escudero, J., Schläppi, M., Ramos, C., Hohn, B., and Koukolíková-Nicola, Z. 1993. T-DNA transfer to maize cells: Histochemical investigation of β-glucuronidase activity in maize tissues. *Proc. Natl. Acad. Sci. USA* **90**:1488–1492.
5. Chan, M.-T., Lee, T.-M., and Chang, H.-H. 1992. Transformation of indica rice (*Oryza sativa* L.) mediated by *Agrobacterium tumefaciens*. *Plant Cell Physiol.* **33**:577–583.
6. Chan, M.-T., Chang, H.-H., Ho, S.-L., Tong, W.-F., and Yu, S.-M. 1993. *Agrobacterium*-mediated production of transgenic rice plants expessing a chimeric α-amylase promoter/β-glucuronidase gene. *Plant Mol. Biol.* **22**:491–506.
7. Mooney, P.A., Goodwin, P.B., Dennis, E.S., and Llewellyn, D.J. 1991. *Agrobacterium tumafiens*-gene transfer into wheat tissues. *Plant Cell Tissue Organ Cult.* **25**:209–218.
8. Hiei, Y., Ohta, S., Komari, T., and Kumasho, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal* **6**:271–282.
9. Armstrong, C.L. and Green, C.E. 1985. Established and maintenance of friable, embrygenic maize callus and the involvement of L-proline. *Planta* **164**:207–214.
10. Koziel, M.G. et al. 1993. Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Bio/Technology* **11**:194–200.
11. Wan, Y., Widholdm, J.M., and Lemaux, P.G. 1995. Type I callus as a bombardment target for generating fertile transgenic maize (*Zea mays* L.) *Planta* **196**:7–14.
12. Frame, B.R., Drayton, P.R., Bagnall, S.V., Lewnau, C.J., Bullock, W.P., Wilson, H.M., Dunwell, J.M., Thompson, J.A., and Wang, K. 1994. Production of fertile transgenic maize plants by silicon carbide whisker-mediated transformation. *The Plant Journal* **6**:941–948.
13. D'Halluin, K., Bonne, E., Bossut, M., Beuckeleer, M.D., and Leemans, J. 1992. Transgenic maize plants by tissue electroporation. *The Plant Cell* **4**:1495–1505.
14. Chu, C.-C. 1978. The N6 medium and its applications to anther culture of cereal crops, pp. 43–50 in *Proc. Symp. Plant Tissue Culture*. Science Press, Peking.
15. Linsmaier, E. and Skoog, F. 1965. Organic growth factor requirements of tobacco tissue culture. *Physiol. Plant.* **18**:100–127.
16. Does, M.P., Dekker, B.M.M., de Groot, M.J.A., and Offringa, R. 1991. A quick method to estimate the T-DNA copy number in transgenic plants at an early stage after transformation, using inverse PCR. *Plant Mol. Biol.* **17**:151–153.
17. Yadav, N.S., Vanderleyden, J., Bennett, D.R., Barnes, W.M., and Chilton, M.D. 1982. Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. USA* **79**:6322–6326.
18. Zambryski, P., Depicker, A., Kruger, K., and Goodman, H.M. 1982. Tumor induction by Agrobacterium tumefaciens: analysis of the boundaries of T-DNA. *J. Mol. Appl. Genet.* **1**:361–370.
19. Ohta, S., Mita, S., Hattori, T., and Nakamura, K. 1990. Construction and expression in tobacco of a β-glucuronidase (GUS) reporter gene containing an intron within the coding sequence *Plant Cell Physiol.* **31**:805–813.
20. Rhodes, C.A., Pierce, D.A., Mettler, I.J., Mascarenhas, D., and Detmer, J.J. 1988. Genetically transformed maize plants from protoplasts. *Science* **240**:204–207.
21. Gordon-Kamm, W.J. et al. 1990. Transformation of maize cells and regeneration of fertile transgenic plants. *The Plant Cell* **2**:603–618.
22. Fromm, M.E., Morrish, F., Armstrong, C., Williams, R., Thomas, J., and Klein, T.M. 1990. Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Bio/Technology* **8**:833–839.
23. Walters, D.A., Vetsch, C.S., Potts, D.E., and Lundquist, R.C. 1992. Transformation and inhentance of a hygromycin phosphotansferase gene in maize plants. *Plant Mol. Biol.* **18**:189–200.
24. Komari, T. 1990. Transformation of cultured cells of *Chenopodium quinoa* by binary vectors that carry a fragment of DNA from the virulence region of pTiBo542. *Plant Cell Rep.* **9**:303–306.
25. Denecke, J., Goselé, J., Botterman, J., and Cornelissen, M. 1989. Quantitative analysis of transiently expressed genes in plant cells. *Methods in Molecular and Cellular Biology* **1**:19–27.
26. Ditta, G., Stanfield, S., Corbin, D., and Helinski, D.R. 1980. Broad host range DNA cloning system for Gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* **77**:7347–7351.
27. Komari, T., Saito, Y., Nakakido, F., and Kumashiro, T. 1989. Efficient selection of somatic hybrids in *Nicotiana tabacum* L. using a combination of drug-resistance markers introduced by transformation. *Theor. Appl. Genet.* **77**:547–552.
28. Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Jefferson, R.A. 1987. Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol. Biol. Rep.* **5**:387–405.

Maydica 40 (1995): 275-281

REGENERATION POTENTIAL OF TROPICAL, SUBTROPICAL, MIDALTITUDE, AND HIGHLAND MAIZE INBREDS

N.E. Bohorova*, B. Luna, R.M. Brito, L.D. Huerta, D.A. Hoisington

CIMMYT, Int., Lisboa 27, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

Received June 11, 1995

ABSTRACT - Plant regeneration potential of 44 tropical and subtropical, 23 midaltitude, and eight highland maize inbred lines was evaluated during the first phase of a project directed at transforming CIMMYT maize (*Zea mays* L.) germplasm. Plants were grown at CIMMYT's experiment stations in El Batan and Tlaltizapan, Mexico. Immature embryos (1.0-1.5 mm) from selfed ears were harvested for establishing callus cultures. Basal N6 medium supplemented with 2 mg/l Dicamba (N6C1) proved better than other media tested for callus initiation and maintenance. When N6C1 medium was supplemented with silver nitrate, an increase in both embryogenic callus formation and regeneration potential was observed. Regenerable embryogenic calli were obtained from 50% of tropical and subtropical lines, 87% of midaltitude lines, and 75% of highland lines tested. Plant regeneration from embryogenic calli was achieved within the first 10 days after calli were transferred to N6 medium without auxins or to MS medium supplemented with IAA and 6-BAP. Complete plantlets were formed within 20-30 days and then transferred to the greenhouse for evaluation and seed production. Compared to the original plants, regenerated plants showed no major alterations in morphology or chromosome number. Successful regeneration of these maize inbreds should provide the means to directly transform elite CIMMYT maize inbred germplasms.

KEY WORDS: Maize; Culture media; Immature embryos; Embryogenic callus; Regeneration.

INTRODUCTION

Maize (*Zea mays* L.) is grown on some 56 million hectares in subtropical, midaltitude and highland ecologies of the developing world. An ongoing concern of the CIMMYT Maize Program is to continue introducing new germplasm from sources within CIMMYT and elsewhere and to encourage the development of inbred lines (Hess and Wedderburn, 1994). Genetic engineering provides a means by which new genetic material can be introduced, often from entirely unrelated species.

In recent years, embryogenic maize tissue cultures have been routinely used as targets for producing transgenic plants via the biolistics process (Fromm *et al.*, 1990; Gordon-Kamm *et al.*, 1990; Genovesi *et al.*, 1992; Koziel *et al.*, 1993). Production of genetically transformed plants depends both on the ability to integrate foreign genes into target cells and the efficiency with which plants are regenerated from genetically transformed cells. Numerous maize inbreds and hybrids have been regenerated from embryo derived calli (Green and Phillips, 1975; Lu *et al.*, 1983; Lowe *et al.*, 1985; Hodges *et al.*, 1986; Lupotto and Lusardi, 1988; Vasil and Vasil, 1986), but efficient plant regeneration has been obtained from relatively few of them. The inbred line, A188, has little agronomic value but is superior to most other maize inbreds in its capacity to regenerate plants from embryogenic callus.

Most studies on maize regeneration have utilized genotypes adapted to temperate zones (Vasil *et al.*, 1984; Duncan *et al.*, 1985), and little attention has been focused on the regeneration potential of maize germplasm adapted to tropical and subtropical regions (Prioli and Silva, 1989; Bohorova and Hoisington, 1992). In tropical maize, the genes controlling plant regeneration under culture conditions appear to be present in higher frequencies within the Cateto race than within the Tuxpeño race. All eight inbreds tested by Prioli and Silva (1989), derived solely from Cateto germplasm, had regeneration capacity, compared to only 2 of 14 inbreds derived from Tuxpeño germplasm.

Our continuing effort is to screen elite CIMMYT inbred maize lines for embryogenic callus formation and plant regeneration, thereby identifying lines which can be used in the direct introduction of potentially useful agronomic traits by genetic transfor-

* For correspondence (fax +52 5 726-7559).

mation methods. This report summarizes the results of the evaluation of 44 tropical and subtropical, 23 midaltitude, and eight highland maize inbred lines for the potential to be regenerated from callus cultures and therefore, useful for transformation protocols.

MATERIALS AND METHODS

Plant Material

Forty-four tropical and subtropical, 23 midaltitude, and eight highland maize inbred lines from CIMMYT's Maize Program were planted at CIMMYT's experiment stations in El Batan, State of Mexico (19°31´N, 98°50´W, 2249 masl), and Tlaltizapan (18°41´N, 99°9´W, 940 masl), State of Morelos, Mexico. Self-pollinated plants grown in 1991-1993 were the source of immature embryos used for starting *in vitro* cultures. The tropical and subtropical CIMMYT maize lines (CML) were derived from different gene pools—Tuxpeño, ETO, Antigua, Suwan 1—as well as various maize populations and hybrids from national programs and seed companies in subtropical and midaltitude areas of the world. Midaltitude inbred lines used in these experiments were developed at CIMMYT's station in Harare, Zimbabwe, in areas at roughly 900-1800 masl (H. CORDOVA, personal communication). The tropical highland maize inbreds were developed at CIMMYT's experimental station in El Batan, Mexico (LOTHROP, 1994).

Callus Initiation and Maintenance

Immature embryos were excised from seeds harvested 18-30 days after pollination at El Batan and 15-17 days after pollination at Tlaltizapan (depending on genotype). Whole ears were surface-sterilized with 70% ethanol for 1 min, followed by 20% Clorox containing 10 drops/liter of polyoxyethylene sorbitan monooleate (Tween-80) for 30 min and rinsed three times with sterile de-ionized water. Immature embryos, 0.5-2.0 mm in size, were aseptically removed from the kernels and placed flat-side down, scutellum up, on the initiation medium. Fifty immature embryos collected from three to six plants of each genotype were tested. For callus initiation, the cultures were incubated in the dark at 26°C.

Three media were evaluated for all genotypes tested. N6C1 consisted of modified N6 basal medium (N6) (CHU *et al.*, 1975) supplemented with 200 mg/l casein hydrolysate, 2.302 mg/l L-proline, 30 g/l sucrose, 2 mg/l 3,6-dichloro-o-anisic acid (Dicamba). N6C1SN medium consisted of N6C1 supplemented with 15.3 mg/l silver nitrate. N6C3 was composed of N6 basal medium plus 2 mg/l 2,4-dichlorophenoxyacetic acid (2,4D). For midaltitude and highland inbreds, an additional medium tested was MSE3, which was composed of modified MURASHIGE and SKOOG (1962) medium supplemented with 150 mg/l L-asparagine, 6% sucrose, and 2.5 mg/l 2,4D. The pH of all media was adjusted with NaOH to 5.7, and 8 g/l agar (Bacto) added before sterilization.

The yellow, irregularly shaped, compact tissue usually obtained on initiation media was transferred to the maintenance medium , which was the same as the initiation medium. Somatic embryo formation was evaluated two weeks after immature embryo initiation using a stereomicroscope. Tissue producing more than three somatic embryos was considered embryogenic. Embryogenic tissue was subcultured every 21 days.

Plant Regeneration

Plants were regenerated from embryogenic calli by transferring tissue to glass vials containing modified N6 hormone-free medium containing 20 g/l sucrose (N6R) or basal MS medium with 20 g/l sucrose, 0.5 mg/l indol-3-acetic acid (IAA), 1 mg/l 6-benzylamino purine (6-BAP) (MSR). The pH of all media was adjusted with NaOH to 5.7, and 8 g/l agar (Bacto) added before sterilization. Calli were incubated in a growth chamber at 26°C with a 16:8 light:dark photoperiod. The percentage of plant regeneration was calculated based on the number of embryos regenerating at least three plants out of the total number of embryos plated on the callus initiation medium. To establish good root systems, plantlets were transferred to modified MS medium containing 1 mg/l naphthaleneacetic acid (NAA), and later transplanted to Jiffy pots and kept in a growth chamber for one week, before being moved to the greenhouse.

Cytological Investigation

Mitotic metaphase cells from the root meristem of 120 plants regenerated from CML67 were analyzed in squash preparations, and chromosomes were stained using the technique of JEWELL and ISLAM-FARIDI (1993).

RESULTS

Callus Formation

Four factors were found to influence the development of embryogenic calli: 1) the size of immature embryos, 2) the genotype, 3) the callus media, and 4) the environment where the initial plants were grown.

The optimal embryo length for callus formation was determined to be 1.0-1.5 mm. Embryos longer than 2 mm produced compact calli or germinated without any callus formation, while those under 0.5 mm in length showed no response. Excised embryos placed on initiation medium displayed visible changes in the scutellum during the first week of *in vitro* culture. The primary growth region was in the scutellum near the coleorhizal end of the embryo, where cells proliferated and grew rapidly. Within two weeks, the scutellar surface became irregular and was transformed into callus tissue. Seventy seven percent of all lines tested formed compact and nodular Type I calli, with many scutellum-like bodies (ARMSTRONG and GREEN, 1985) (Figure 1). A lower number of the lines tested (23%) formed Type II calli. The friable Type II callus (ARMSTRONG and GREEN, 1985) was yellowish in color and contained organized somatic embryos at the coleorhizal end of the scutellum (Figure 2). FRANSZ and SCHEL (1991) concluded that most friable calli result from less differentiated tissue, such as young embryos or less differentiated regions of Type I compact calli.

On medium containing 2 mg/l Dicamba (N6C1), the calli formed were friable, yellowish, and fast growing. Dicamba is one of the best growth regulators for producing and growing regenerable maize calli (CON-

FIGURE 1 - Compact, nodular Type I callus from immature embryos of subtropical maize inbred CML131 after 7 days on N6C1SN medium.

FIGURE 2 - Friable Type II calli from immature embryos of tropical maize inbred CML67 with organized somatic embryos.

FIGURE 3 - Regenerated plantlets at optimum size for transplanting into pots.

GER *et al.*, 1982; DUNCAN *et al.*, 1985), as is Dicamba combined with L-proline (ARMSTRONG and GREEN, 1985; LUPOTTO and LUSARDI, 1988). When N6C1 medium was supplemented with silver nitrate, a greater number of immature embryos formed friable calli with small somatic embryos on the surface. Generally calli exhibited a slower growth rate on the medium with silver nitrate. Silver nitrate is known to be a potent inhibitor of ethylene action, which influences cell division and cytodifferentiation (SONGSTAD *et al.*, 1988).

Of the 44 tropical and subtropical genotypes tested, 31 of those grown at El Batan and 38 of the ones from Tlaltizapan produced embryogenic calli on at least one of the media tested. Of those genotypes forming embryogenic calli, CML121, CML128, CML135, and CML137 exhibited few differences across the three media tested, while CML139, Ki14, and CML71 expressed definite media preferences (N6C1). Inbred lines CML130 and Mp496, harvested at El Batan, as well as CML63, CML64, CML65, CML70, CML124, and CML138, harvested at El Batan and Tlaltizapan, did not form embryogenic calli on any media tested.

The experiments with midaltitude inbreds were based on results of embryogenic callus formation and plant regeneration from the experiments with tropical and subtropical maize. Most of the genotypes evaluated produced calli on at least one of the four media

TABLE 1 - *Plant regeneration expressed in percent from immature embryos of tropical and subtropical maize inbreds on MSR medium. Plants were grown at CIMMYT's field stations at El Batan or Tlaltizapan, Mexico and embryos were initially plated on either N6C1 or N6C1SN media.*

	CALLUS INITIATION MEDIA		
	N6C1		N6C1SN
Inbred	El Batan	Tlaltizapan	Tlaltizapan
TROPICAL			
CML24	14.1	0.0	0.0
CML38	0.0	–	–
CML60	16.6	0.0	0.0
CML61	–	0.0	0.0
CML62	–	17.0	17.1
CML63	0.0	0.0	0.0
CML64	0.0	0.0	0.0
CML65	0.0	0.0	0.0
CML66	14.0	8.8	13.8
CML67R	–	77.0	88.1
CML67G	–	56.0	75.0
CML68	0.0	0.0	15.5
CML69	–	1.5	3.0
CML70	–	0.0	0.0
CML71	0.0	16.7	24.0
CML72	0.0	50.2	65.4
CML73	0.0	0.0	0.0
CML74	–	5.7	44.2
Ki3	12.1	8.1	28.4
Ki14	0.0	45.8	64.6
Mp704	0.0	0.0	0.0
Tz18	0.0	0.0	0.0
Hi27	0.0	0.0	12.2
Hi34	11.1	18.4	18.6
ICAL27	0.0	–	–
Mp496	0.0	–	–
SUBTROPICAL			
CML117	0.0	0.0	0.0
CML119	–	0.0	0.0
CML121	14.0	0.0	19.1
CML122	–	6.8	16.2
CML123	–	6.7	37.5
CML124	0.0	0.0	0.0
CML125	15.0	–	–
CML126	0.0	–	–
CML127	0.0	0.0	0.0
CML128	27.5	41.2	46.7
CML129	0.0	–	–
CML130	0.0	26.5	27.3
CML131	–	30.0	48.0
CML134	0.0	0.0	0.0
CML135	13.4	22.0	22.0
CML137	13.7	22.2	52.6
CML138	0.0	0.0	0.0
CML139	0.0	52.5	87.1

tested; however, differences were observed in the number of embryos producing calli. On N6C1SN medium, callus formation started on the immature embryo explants 3-4 days after culture, and friable, nodular calli were visible on the tenth day. Fifteen from 19 genotypes (79%) formed embryogenic calli on N6C1SN medium, 15 from 20 (75%) on N6C3, 15 from 21 (71%) on N6C1, and 11 from 19 (58%) on MSE3. On MSE3, calli were friable, and organized somatic embryos were visible under the stereo microscope. Embryos of CML213, CML214, CML215, CML216, and CML217 produced the most vigorous embryogenic calli across the four media tested.

While the number of highland maize inbreds tested were limited, we found that N6C1SN medium initiated scutellar callus and supported callus growth from most cultivars tested. Immature embryos of CML241 had the highest callus induction frequency, and 56 calli from 77 embryos (73%) on N6C1SN medium were classified as embryogenic.

Plant Regeneration

The capacity to regenerate plantlets was correlated to the ability to form embryogenic calli, however not all embryos with embryogenic callus regenerated plants. The classification of the callus as embryogenic did not necessarily imply regenerability. Regeneration was induced by transferring calli to the same medium as for callus formation but with low sucrose concentration and either no auxins (N6R) or IAA and 6-BAP (MSR). Green shoot formation was evident within a week and plantlet regeneration occurred in 20 days on both media. The composition of the medium used for callus formation was found to be an important factor in determining regeneration. We found that MSR medium was more efficient than N6R medium for plant regeneration, and only the data on MSR will be presented.

High rates of shoot differentiation and plantlet regeneration were observed during the first subculture period and maintained over 12 subcultures (5-15 plants per embryo at each subculture). The highest frequency of plants regenerated from embryogenic calli of tropical inbred lines grown at El Batan was achieved with CML128, which had a regeneration frequency of 28% (Table 1). Nine genotypes - CML24, CML60, CML66, CML121, CML125, CML135, CML137, Ki3, and Hi34 - had 11-17% regeneration on at least one of the callus initiation media. Ki3 showed low callus initiation and therefore a low overall regeneration (28%), but plant regeneration from all initiated calli was high - 15 plants per embryo at each subculture.

Distinct differences in regenerable callus production were observed between Tlaltizapan-grown tropical materials and those grown at El Batan (Table 1). Of the 38 Tlaltizapan-grown inbreds tested, 22 (58%) produced highly regenerable calli compared to 10 from 33 (30%) of El Batan-grown lines. Tlaltizapan is hotter and dryer than El Batan, and all tropical and subtropical lines tested were developed under and for those conditions. Differences in the embryogenic response of inbred lines grown at El Batan and Tlaltizapan indicate that the environment experienced by donor plants may be important in determining rates of callus initiation and plant regeneration. In some inbreds, such as CML67, CML72, and CML139, shoots grew from almost all embryos tested- 88%, 65%, and 87%, respectively - with an average of 5-15 plants per embryo. CML128, CML131, CML137, and Ki14 showed plant regeneration rates from 47% to 65%. The main effect of media used for callus formation was on the frequency of shoot formation in the embryos tested. For example, the regeneration response of CML74 embryos cultured on N6C1 medium was 6%, while on N6C1SN it was 44%. Ten inbreds from El Batan and Tlaltizapan showed no regeneration.

Within 20 days, regenerated calli formed plantlets with axial roots and 2-3 leaves and were transferred to Jiffy pots and later transplanted to soil (Figure 3). Within two weeks, plants from tropical materials were taken to the greenhouse for further growth . Most plants (80%) were morphologically normal, similar to the original lines, reached full maturity, and were male and female fertile. All evaluated plants had 20 chromosomes. A few plants (5-10%) were classified as tassels-seeds, with seeds present at the base of the tassel. This type of abnormality was observed to occur when plantlets were kept on regeneration media for a long period (about 30 days). The abnormality was eliminated by transferring plantlets with 2-3 leaves within 15 days to rooting media and then to Jiffy pots.

In midaltitude maize, high rates of shoot formation were observed when calli were transferred to MSR medium (Table 2). Of the 23 inbreds tested, 20 (87%) regenerated at least some plantlets on MSR medium, and of those 10 (39%) regenerated plantlets with a frequency of more than 50%. Genotypic differences were detected in the media's ability to initiate shoots and plantlets. The most responsive inbreds were CML214, CML215, CML216, and CML217, which had a regeneration rate of 70-100%. Regeneration rates registered for CML195, CML198, CML200, CML209, and CML210 were 8-100% on at least one of the callus initiation media used. Inbreds CML204, CML206, and CML211 were the least responsive, and CML197, CML205, CML212 were not regenerated.

Plantlets were transferred to soil and two weeks later taken to the greenhouse for further growth. Nearly 80% of the established plants were fertile and set seeds.

TABLE 2 - *Plant regeneration (%) from immature embryos of midaltitude inbreds on MSR medium. Plants were grown at CIMMYT's field station in Tlaltizapan, Mexico and embryos were initially plated on one of the four media indicated.*

	CALLUS INITIATION MEDIA			
Inbred	N6C1	N6C1SN	N6C3	MSE3
CML195	26.3	63.4	14.0	17.6
CML196	12.0	33.3	22.0	4.5
CML197	0.0	0.0	–	–
CML198	60.5	50.0	61.3	8.2
CML199	0.0	–	–	36.0
CML200	–	54.5	–	10.0
CML201	46.7	–	12.5	0.0
CML202	4.3	25.0	45.4	0.0
CML203	0.0	27.8	0.0	–
CML204	0.0	28.6	0.0	0.0
CML205	0.0	0.0	0.0	0.0
CML206	0.0	0.0	15.1	22.1
CML207	21.9	18.5	28.5	–
CML208	18.5	0.0	0.0	0.0
CML209	28.6	–	100.0	0.0
CML210	22.7	26.7	23.8	51.8
CML211	20.0	11.8	5.5	0.0
CML212	–	–	0.0	0.0
CML213	51.8	72.1	58.8	31.0
CML214	85.0	96.7	40.0	95.6
CML215	100.0	100.0	85.7	–
CML216	71.1	100.0	100.0	61.1
CML217	40.6	96.1	80.7	26.7

TABLE 3 - *Regeneration (%) of highland maize inbreds on MSR medium using callus initiation media. Plants were grown at CIMMYT's field station in El Batan, Mexico and embryos were initially plated on one of the four media indicated.*

	CALLUS INITIATION MEDIA			
Inbred	N6C1	N6C1SN	N6C3	MSE3
CML239	16.6	36.6	3.2	8.3
CML240	4.0	0.0	4.0	4.0
CML241	59.7	72.7	46.1	65.7
CML242	19.4	43.8	15.6	51.0
CML243	20.8	28.0	7.3	5.4
CML244	0.0	0.0	0.0	0.0
CML245	0.0	0.0	0.0	0.0
CML246	2.7	40.4	0.0	0.0

We have also observed differences in regeneration capacity among highland maize genotypes on different callus initiation media (Table 3). Regeneration frequencies for CML241 and CML242 were 73% and 44%, respectively, on N6C1SN medium, the best responses recorded for highland inbreds. In terms of donor genetic material, highland CML241 is closely related to CML242 and CML243 but different from the other lines tested. The response of cultured embryos ranged from 1-4% for CML240, 8-37% for CML239, CML243, and 40% for CML246. Six out of eight (75%) highland genotypes tested, regenerated plantlets that grew to maturity.

DISCUSSION

We have shown that, under appropriate culture conditions, embryogenic calli can be produced from immature embryos of several maize genotypes. The scutellar region closest to the coleorhizal area of immature embryos obviously contains cells capable of producing somatic embryos under certain experimental conditions. Factors influencing the expression of totipotency in cell culture are genotype, composition of plant culture medium, growth regulator, and embryo size. Callus initiation and maintenance with Dicamba were found to be important for induction and maintenance of embryogenic calli.

The composition of macro- and micro-elements in modified N6 media may also have an effect on embryogenic callus formation and plant regeneration. Nitrogen and the form of nitrogen must have a very crucial role in callus formation and regeneration. Basic N6 medium which contains a lower nitrogen concentration than MS medium proved excellent for embryogenic callus initiation and maintenance with most lines tested. HODGES *et al.* (1986) demonstrated that differences in the effects of media on somatic embryogenesis and plant production indicate the importance of different genes at various stages of regeneration.

The genotype also has a great effect on plant regeneration. BINGHAM *et al.* (1975) provided the first evidence that shoot regeneration in callus cultures was genetically controlled and could be manipulated using conventional breeding techniques. Different regeneration potentials among genotypes have been reported by DUNCAN *et al.* (1985), TOMES and SMITH (1985), and HODGES *et al.* (1986). The fact that some inbreds produced embryogenic calli and regenerated plants while others were completely incapable of regeneration shows that the genotype is important for regeneration and that there must be genes that are important in controlling regeneration (HODGES *et al.*, 1986). WILLMAN *et al.* (1989) suggest that at least one gene - or block of genes - controls the expression of somatic embryogenesis in maize tissue cultures. Recent studies by ARMSTRONG *et al.* (1992) also support the opinion that there is a major gene (or genes) on the long arm of chromosome 9 in A188 that promotes embryogenic callus initiation and plant regeneration in B73, Mo17, and many other recalcitrant maize inbreds.

In tropical maize, investigated by PRIOLI and SILVA (1989), genes controlling the response for plant regeneration under the described culture conditions were reported to be present in higher frequencies in the Cateto race than in the Tuxpeño race. In our experiments, plants were regenerated from 7 out of 15 maize inbreds derived from the Tuxpeño germplasm, although only 2 had a regeneration rate greater than 40%. The two tropical maize inbreds (CML67 and CML72) showing the highest regeneration potential in our experiments, originated from Antigua germplasm. The Antigua lines, derived from Antigua Group 2 or from crosses with other Antigua bank accessions, are highly resistant to corn borers (D. JEWELL, personal communication). The formation of friable embryogenic calli with high regeneration potential from maize inbreds of Antigua origin may indicate that Antigua germplasm also possesses a gene or genes responsible for plant regeneration. Studies are underway to determine the genetic basis of regeneration potential in these Antigua-based inbreds.

Thus plant regeneration from embryogenic calli of CML67 and CML72 is quick, efficient, and can be used to produce high populations of phenotypically normal plants. The embryogenic/regeneration system - embryogenic callus initiation on N6C1SN medium, maintenance of calli on N6C1, and plant regeneration on MSR medium is being used in our transformation experiments.

To the best of our knowledge, this is the first report of successful plant regeneration from immature embryos of midaltitude and highland inbreds. The high regeneration potential demonstrated by tropical and subtropical maize, midaltitude, and highland materials promises to allow direct transformation of tropical germplasm.

ACKNOWLEDGMENTS - The authors would like to thank all Maize Program breeders for providing the lines used in this study; D. Jewell, J. Deutsch, M. Bjarnason, and M. Willcox for producing the immature embryos; and G. Vergara for his excellent technical assistance. They also thank Alma McNab for help in preparing this manuscript and M. Willcox, M. Khairallah and S. Fennell for their useful comments.

REFERENCES

ARMSTRONG C.L., C.E. GREEN, 1985 Establishment and maintenance of friable, embryogenic maize callus and the involvement of L-proline. Planta **164:** 207-214.

ARMSTRONG C.L., J. ROMERO-SEVERSON, T.K. HODGES, 1992 Improved tissue culture response of an elite maize inbred through backcross breeding, and identification of chromosomal regions important for regeneration by RFLP analysis. Theor. Appl. Genet. **84:** 755-762.

BINGHAM E.T., L.V. HURLEY, D.M. KAATZ, W. SAUNDERS, 1975 Breeding alfalfa regenerates from callus tissue in culture. Crop Sci. **15:** 719-721.

BOHOROVA N.E., D.A. HOISINGTON, 1992 Plant regeneration from immature embryos of tropical maize inbreds. pp. 20-21. *In:* Resumenes del XIV Congreso Nacional de Fitogenetica, editor. Soc. Mex. de Fitogenetica, A.C., Chapingo, Mexico.

CHU C.C., C.C. WANG, C.S. SUN, C. HSU, K.C. YIN, C.V. CHU, C.Y. BI, 1975 Establishment of an efficient medium for anther culture of rice through comparative experiments on nitrogen sources. Sci. Sinica (Peking) **18:** 659-668.

CONGER B.V., L.L. HILENSKI, K.W. LOWE, J.V. CARABIA, 1982 Influence of different auxins at varying concentrations on callus induction and growth from embryo and leaf-tip explants in Gramineae. Environ. Exper. Botany. **22:** 39-48.

DUNCAN D.R., M.E. WILLIAMS, B.E. ZEHR, J.M. WIDHOLM, 1985 The production of callus capable of plant regeneration from immature embryos of numerous *Zea mays* genotypes. Planta **165:** 322-332.

FRANSZ P.F., J.H.N. SCHEL, 1991 An ultrastructural study on the early development of *Zea mays* somatic embryos. Can. J. Bot. **69:** 858-865.

FROMM M.E., F. MORRISH, C. ARMSTRONG, R. WILLIAMS, J. THOMAS, T. KLEIN, 1990 Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. Bio/technology **8:** 833-839.

GORDON-KAMM W.J., T.M. SPENCER, M.L. MANGANO, T.R. ADAMS, R.J. DAINES, W.G. START, J.V. O'BRIEN, S.A. CHAMBERS, W.R. ADAMS, N.G. WILLETTS, T.B. RICE, C.J. MACKEY, R.W. KRUEGER, A.P. KAUSCH, P.G. LEMAUX, 1990 Transformation of maize cells and regeneration of fertile transgenic plants. The Plant Cell **2:** 603-618.

GENOVESI D., N. WILLETTS, S. ZACHWIEJA, M. MANN, T. SPENCER, C. FICK, W. GORDON-KAMM, 1992 Transformation of an elite maize inbred through microprojectile bombardment of regenerable embryogenic callus. In Vitro Cell. Dev. Biol. **28:** 124A.

GREEN C.E., R.L. PHILLIPS, 1975 Plant regeneration from tissue culture of maize. Crop Sci. **15:** 417-421.

HESS D.C., R.N. WEDDERBURN, 1994 Current Status and Future Direction of the Maize Program. pp. iv-vi. *In:* The Subtropical, Midaltitude and Highland Maize Subprogram, Maize Program Special Report, editor. M.Bjarnason, CIMMYT, Mexico, D.F.

HODGES T.K., K.K. KAMO, C.W. IMBRIE, M.R. BEEWAR, 1986 Genotype specificity of somatic embryogenesis and regeneration in maize. Bio Technology **4:** 219-223.

JEWELL D., N. ISLAM-FARIDI, 1993 A technique for somatic chromosome preparation and C-banding of maize. pp. 484-492. *In:* The Maize Handbook, editors. M. Freeling and V. Walbot, Springer-Verlag.

KOZIEL M., G.L. BELAND, C. BOWMAN, N.B. CAROZZI, R. CRENSHAW, L. CROSSLAND, J. DAWSON, N. DESAI, M. HILL, S. KADWELL, K. LAUNIS, K. LEWIS, D. MADDOX, K. MCPHERSON, M. MEGHJI, E. MERLIN, R. RHODES, G.W. WARREN, M. WRIGHT, S. EVOLA, 1993 Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. Bio/technology **11:** 194-200.

LOTHROP J.E, 1994 Research on maize for highland regions. pp. 19-43. *In:* The Subtropical, Midaltitude and Highland Maize Subprogram, Maize Program Special Report, editor. M. Bjarnason, CIMMYT, Mexico, D.F.

LOWE K., D.B. TAYLOR, P. RYAN, K. PATERSON, 1985 Plant regeneration via organogenesis and embryogenesis in the maize inbred line B73. Plant Science **41:** 125-132.

LU C., V. VASIL, I.K. VASIL, 1983 Improved efficiency of somatic embryogenesis and plant regeneration in tissue cultures of maize (*Zea mays* L.). Theor. Appl. Genet. **66:** 285-289.

LUPOTTO E., M.C. LUSARDI, 1988 Secondary somatic embryogenesis from regenerating plantlets of the inbred line B79 of maize (*Zea mays* L.) Switch from type I to type 2 callus and effect on the regenerative potential. Maydica **33:** 163-177.

MURASHIGE T., F. SKOOG, 1962 A revised medium for rapid growth and bioassays with tobacco tissue cultures. Plant Physiol. **15:** 473-497.

PRIOLI L.M., W.J. SILVA, 1989 Somatic embryogenesis and plant regeneration capacity in tropical maize inbreds. Rev. Brazil. Genet. **12:** 553-566.

SONGSTAD D.D., D.R. DUNCAN, J.M. WIDHOLM, 1988 Effect of L-aminocyclopropane-L-carboxylic acid, silver nitrate, and norbornadiene on plant regeneration from maize callus cultures. Plant Cell Reports. **7:** 262-265.

TOMES D.T., O.S. SMITH, 1985 The effect of parental genotype on initiation of embryogenic callus from elite inbred maize (*Zea mays* L.) germplasm. Theor. Appl. Genet. **70:** 505-509.

VASIL V., I.K. VASIL, 1986 Plant regeneration from friable embryogenic callus and cell suspension cultures of *Zea mays* L. Plant Physiol. **124:** 399-408.

VASIL V., I.K. VASIL, C. LU, 1984 Somatic embryogenesis in long-term callus cultures of *Zea mays* L. (Gramineae). Am. J. Bot. **71:** 158-161.

WILLMAN M.R., S.M. SCHROLL, T.K. HODGES, 1989 Inheritance of somatic embryogenesis and plantlet regeneration from primary (Type I) callus in maize. Vitro Plant Cell Dev. Biol. **25:** 95-100.

Susceptibility of Four Tropical Lepidopteran Maize Pests to *Bacillus thuringiensis* CryI-Type Insecticidal Toxins

N. BOHOROVA,[1] M. CABRERA,[2] C. ABARCA,[2] R. QUINTERO,[2] A. M. MACIEL,[1] R. M. BRITO,[1] D. HOISINGTON,[1] AND A. BRAVO[2]

J. Econ. Entomol. 90(2): 412–415 (1997)

ABSTRACT The relative susceptibility of 4 tropical maize pests, *Spodoptera frugiperda* (J. E. Smith), *Diatraea grandiosella* Dyar, *D. saccharalis* (F.), and *Helicoverpa zea* (Boddie), to the lepidopteran-specific CryI-type proteins produced by *B. thuringiensis* is presented. The toxin with the highest potency against *H. zea* larvae was the CryIAc toxin. *S. frugiperda* larvae were susceptible to CryID and CryIF toxins. The CryIB toxin showed to be highly toxic against *D. grandiosella* and *D. saccharalis*. This information will establish a basis for selecting *B. thuringiensis* strains producing the appropriate CryI proteins to be used for the biological control of these tropical pests.

KEY WORDS *Spodoptera frugiperda, Diatraea grandiosella, Diatraea saccharalis, Helicoverpa zea, Bacillus thuringiensis,* δ-endotoxin

Bacillus thuringiensis (BERLINER) is a gram-positive bacterium that produces crystallic protein inclusions during sporulation. This inclusions are formed by proteins called δ-endotoxins or insecticidal crystal proteins (ICP), which are toxic to insect larvae, many of which are disease vectors or major crop pests. The use of *B. thuringiensis* as a microbial insecticide has several advantages over the use of chemical control agents; *B. thuringiensis* strains are highly specific for certain hosts and are not toxic to other insects, and plants and vertebrates and are completely biodegradable so no residual toxic products accumulate in the environment. Preparations of *B. thuringiensis*-based insecticides applied to the foliage are washed off by rain and may be inactivated by sunlight; therefore, only relatively short-term protection from pest populations is obtained from single applications.

It has been estimated that up to 15% of crops worldwide are lost because of insect damage alone (Boulter et al. 1989). Insecticidal crystal proteins from *B. thuringiensis* have been used as biopesticides during the past 30 yr and a wide array of commercial products are now available and used in the field. Furthermore, some ICP genes have been introduced into the plant genome, with a high level of protection from insect attack (Vaeck et al. 1987, Perlak et al. 1990).

Bacillus thuringiensis strains can be isolated from soil samples from all over the world on plant surfaces, dead insects, and stored grains. Currently, 45 different serotypes and 8 nonflagellated biotypes have been cataloged (Lecadet et al. 1994). Numerous ICP genes have been cloned, sequenced, and classified, based on their homology and specificity (Höfte and Whiteley 1989). To date, 11 subgroups of *cryI* genes specific to lepidopteran pests have been characterized and identified. Each CryI toxin has a narrow range of toxicity against different lepidopteran insects (Feitelson et al. 1992). Five *cryIII* genes specific against coleopteran pests have been described. Both insect orders contain some of the most devastating agricultural pests.

The search for novel *B. thuringiensis* ICPs is a priority worldwide. This toxin will provide different alternatives for insect control and for coping with the problem of insect resistance. In this sense, it is important to have a methodology that allows for accurate identification of the most active toxins against selected pests. The analysis of toxic activity through bioassay allows the identification of the powerful toxins against a particular insect.

Maize (*Zea mays* L.) is one of the principal crops in Mexico and other developing countries, where it is used for human and animal nutrition. The crop is a target of many different pests including 4 lepidoptera species: *S. frugiperda, D. grandiosella, D. saccharalis,* and *H. zea*. The objective of the current study was to determine the relative susceptibility of these 4 tropical pests to the lepidopteran specific proteins produced by *B. thuringiensis*.

Materials and Methods

Purification of Insecticidal Crystal Proteins. Insecticidal crystals containing CryIAa, CryIAb,

[1]International Maize and Wheat Improvement Center, CIMMYT, Ap. Postal 6-641, 06600 México, D.F., México.
[2]Instituto de Biotecnología-UNAM, Ap. Postal 510-3, Cuernavaca 62250, Morelos, México.

0022-0493/97/0412–0415$02.00/0

and CryIAc protoxins were purified from *B. thuringiensis* HD1 strain. CryIAc protoxin was purified from *B. thuringiensis* subsp. *kurstaki* HD73 strain. CryIE toxin was purified from *B. thuringiensis* HD125 strain and CryIF from *B. thuringiensis* 1B181 strain. CryIAb, CryIC, and CryID protoxins were obtained as recombinant proteins expressed in *Escherichia coli.* The *cryIAb* gene was cloned from *B. thuringiensis* subsp. *berliner* 1715 (Höfte et al. 1986). The *cryID* gene was cloned from *B. thuringiensis* subsp. *aizawai* HD68 (Höfte et al. 1990). The *cryIC* gene was cloned from *B. thuringiensis* subsp. *entomocidus* 60.5 (Honée et al. 1988). *E. coli* transformant strains were kindly supplied by M. Peferoen (Plant Genetic Systems, Gent, Belgium). CryIB protoxin was purified from the Mexican isolated *B. thuringiensis* subsp. *thuringiensis* IB43 strain. This strain was characterized by enzyme-linked immunosorbent assay (ELISA) using a specific monoclonal antibody against CryIB toxin and polymerase chain reaction (PCR) analysis using specific primers for the *cryIB* gene described in Cerón et al. (1994).

Crystalline inclusions produced in *B. thuringiensis* strains were purified from spores and cell debris by centrifugation in discontinuous sucrose gradients as described by Thomas and Ellar (1983). Cells were grown in nutritive medium (Difco, Detroit MI) until complete sporulation was achieved. Spores and crystals were centrifuged at 6,000 × *g* for 10 min at 4°C. The pellet was washed 4 times with deionized water. The final pellet was resuspended in 50 mM Tris·HCl (pH 7.5). Spores and crystals were layered on top of a 30-ml discontinuous sucrose gradient, composed of 6 ml each of 67, 72, 79, 84, and 90% (wt:vol) sucrose in 50 mM Tris·HCl, (pH 7.5). Centrifugation was carried out at 80,000 × *g* for 14 h at 4°C. Crystals formed a major band at the interface between 72 and 79% sucrose or between 79 and 84% sucrose, while the spores formed a discrete pellet at the bottom of the tube. The crystal band was removed and washed 3 times in 50 mM Tris·HCl (pH 7.5) and centrifuged at 15,000 × *g* for 5 min at 4°C to remove the sucrose. The final pellet was resuspended in deionized water. Purity of the crystal preparation was monitored by phase contrast microscopy. If necessary, a 2nd centrifugation in the discontinuous sucrose gradients was performed. Crystal inclusion bodies were solubilized in 0.1 M $NaHCO_3$, 10 mM β-mercaptoethanol, pH 9.5 for 4 h at 37°C. After centrifugation at 15,000 × *g* for 10 min, protein concentrations were determined by using the Bradford procedure (Bradford 1976) with bovine serum albumin as a protein standard and sodium dodecyl sulphate/polyacrylamide gel electrophoresis (SDS-PAGE) was conducted as described by Laemmli (1970)

The insecticidal proteins produced in the recombinant *E. coli* strains were purified from 500 ml of saturated culture in LB medium (Sambrook et al. 1989) supplemented with the corresponding antibiotic. Cells were suspended in 50 mM Tris·HCl (ph 7.5), 50 mM EDTA, 15% sucrose, treated with lysozime (100 mg/ml) for 30 min at 37°C and sonicated on ice 4 times for 1 min. The cell debris was washed twice with 200 ml of 50 mM Tris·HCl, 150 mM NaCl (ph 7.5), containing 2% triton X-100, incubated for 30 min in the same buffer at 0°C, centrifuged at 15,000 × *g* for 10 min, and washed twice with 50 mM Tris·HCl, 150 mM NaCl, (ph 7.5). *B. thuringiensis* protoxins present in the pellet were solubilized in 0.1 M $NaHCO_3$, 10 mM β-mercaptoethanol pH 9.5 for 4 h at 37°C and centrifuged at 15,000 × *g* for 10 min. Solubilized protein was loaded on a Sephacryl S-300 column (100 × 2.5 cm) and eluted with a flow rate of 40 ml/h.

Toxicity Assays. Larvae of *S. frugiperda, D. grandiosella, D. saccharalis,* and *H. zea* were obtained from colonies established in the Entomological Rearing Facility at CIMMYT (Mexico) where the bioassay were done.

Artificial diets, prepared as described by Mihm (1982, 1983a, b) were used for all tests. Each toxin was diluted in water, mixed with a vortex mixer for 1 min, sonicated, and added to the diet at the rate of 10 or 100 mg/g diet. The mixture was poured into each well of a 32-well microtiter plates (Cell Wells, Corning Glass Works, Corning, New York). The diet was allowed to solidify and a single larvae (1-d-old larvae of *S. frugiperda, D. grandiosella,* and *D. saccharalis* and 2-d-old larvae of *H. zea*) was added to each well. Microtiter plates were covered with polyester film lidding material and incubated in a growth chamber at 27°C, 65 ± 5% RH, and a photoperiod of 16:8 (L:D) h. Thirty-two insects were evaluated per toxin and each treatment was replicated at least 3 times. Mortality was assessed after 7 d.

Mortality of the control larvae reared on a toxin-free diet and under the same conditions was recorded and used to correct the mortality test with Abbott formula (Abbott 1925) The LC_{50} values and confidence limits were obtained by probit analysis (Finney 1971).

Results and Discussion

Bioassays with neonate lepidopteran larvae were performed with the CryI type ICPs. It has been reported that these proteins are highly specific against lepidoteran insects. The percentage of mortality obtained with 10 mg of each toxin applied per gram of artificial diet is presented in Fig. 1.

Helicoverpa zea is a serious pest of highland maize crops. We found that the toxin with the highest potency against *H. zea* larvae was the CryIAc toxin (Fig. 1; Table 1). This insect was also susceptible to the CryIAb toxin (Fig. 1). These data are in agreement with previous bioassays that showed that CryIAc has a LC_{50} values of 464 ng/cm^2 on *H. zea* larvae (Garczynski et al. 1991). However, there is noinformation regarding the sus-

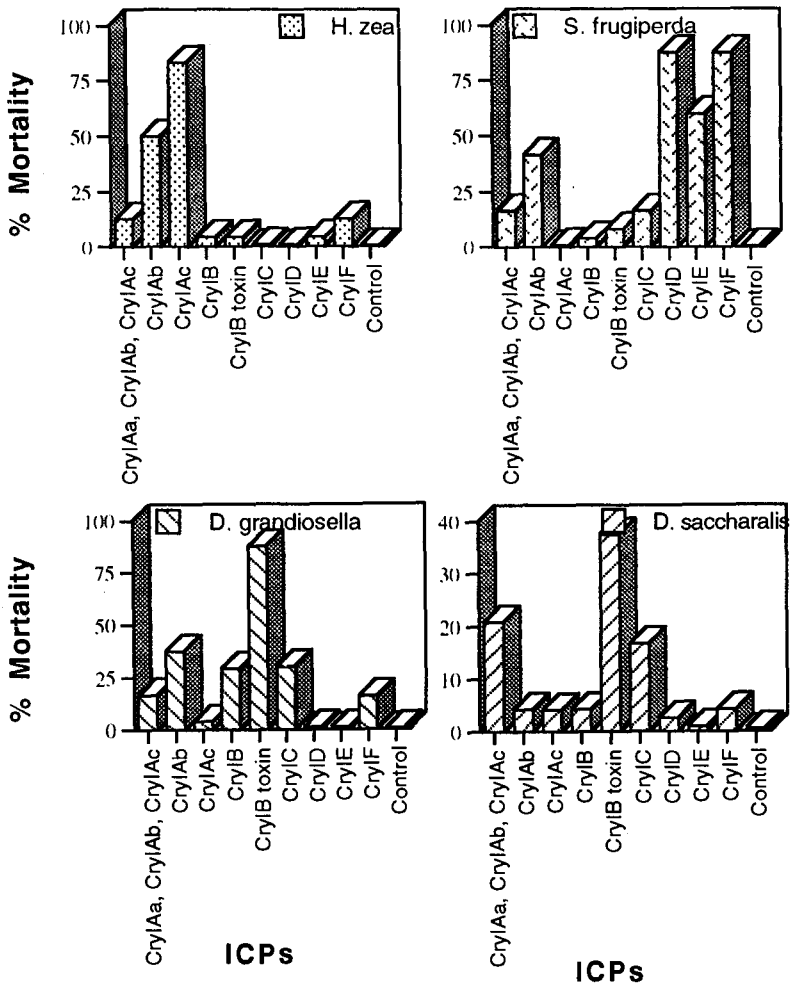


Fig. 1. Response of 4 different lepidopteran insects to CryI insecticidal crystal proteins from *B. thuringiensis*. Percentage mortality obtained with 10 mg of each toxin applied per gram of artificial diet.

ceptibility of this insect toward other Cry toxins. Other *Heliothis* species are also highly susceptible to CryIAc. For example, *H. virescens*, which is a cotton crop pest that was shown to be highly susceptible to CryIAc and CryIAb toxins, presented LC_{50} of 2 and 7 ng, respectively, of toxin applied over the surface of 1 cm^2 of artificial diet (Van Rie et al. 1989).

Table 1. Toxicity of Cry toxins on lepidopteran insects

Insect	ICP	LC_{50}	CI_{95}	*n*
D. grandiosella	CryIB	5.2	3.6–5.5	7
D. saccharalis	CryIB	113.6	45.8–318.9	3
S. frugiperda	CryID	1.54	0.4–2.9	4
H. zea	CryIAc	8.2	7.2–12.9	3

Data are expressed in micrograms of toxin applied per gram of artificial diet. 95% CI were calculated from probit analysis. *n*, Number of replications.

Information regarding the susceptibility of *S. frugiperda* to the Cry protein family is limited. The larval stage of this insect is an important pest of maize, cotton, alfalfa, clover, peanuts, and many garden crops. The CryIC toxin has been reported to be toxic against *S. exigua* (Visser et al. 1990) and *S. littoralis* (Van Rie et al. 1990) with LC_{50} values of 68 and 93 ng/cm^2, respectively. CryID and CryIE have been reported to be toxic to *S. littoralis* with LC_{50} of 423 and 88 ng/cm^2, respectively (Van Rie et al. 1990). In this study, we analyzed the susceptibility of *S. frugiperda* larvae to the family of CryI toxins and we found that the toxins with the highest activity were CryID and CryIF toxins (Fig. 1).

There is little information regarding the control of *D. grandiosella* or *D. saccharalis* by *B. thuringiensis* ICPs. We found that both types of larvae are highly susceptible to CryIB toxin (Fig. 1). Table 1 shows the LC_{50} and the confidence intervals

for the *D. grandiosella* and *D. saccharalis* bioassays. The CryIB protein was highly toxic to *D. grandiosella*, LC_{50} of 5.1 mg/g artificial diet was obtained. CryIB toxin was found to be ≈20 times less active against *D. saccharalis*.

Our findings indicate that no single ICP is active against all pests species. However, these results will be important in selecting *B. thuringiensis* strains producing the appropriate CryI protein to be used for the biological control of each of these tropical pests or the different ICPs combinations needed to control maize pests, where mixed insect populations attack the crop. Also, the knowledge of which toxin has the higher activity against a selected pest will set the basis for selecting which genes will be most appropriate in transgenic plants generation.

Acknowledgments

We thank M. Peferoen for providing recombinant *E. coli* strains and Myriam Ortíz for strain maintenance. This work was supported in part by UNDP; DGAPA IN-214294 and CONACyT 400344-5-4311N.

References Cited

Abbott, W. S. 1925. A method of computing the effectiveness of an insecticide. J. Econ. Entomol. 18: 265–267

Boulter, D., A.M.R. Gatehouse, and V. Hilder. 1989. Use of cowpea trypsin inhibitor (CPTI) to protect plants against insect predation. Biotech. Adv. 7: 489–497.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72: 248–254.

Cerón, J., L. Covarrubias, R. Quintero, A. Ortiz, M. Ortiz, E. Aranda, L. Lina, and A. Bravo. 1994. PCR analysis of the *cryI* insecticidal crystal family genes from *Bacillus thuringiensis*. Appl. Environ. Microbiol. 60: 353–356.

Feitelson, J. S., J. Payne, and L. Kim. 1992. *Bacillus thuringiensis*: insects and beyond. Biotechnology 10: 271–275.

Finney, D. J. 1971. Probit analysis. Cambridge University Press, Cambridge.

Garczynski, S. F., J. W. Crim, and M. J. Adang. 1991. Identification of putative insect brush border membrane-binding molecules specific to *Bacillus thuringiensis* d-endotoxin by protein blot analysis. Appl. Environ. Microbiol. 57: 2816–2820.

Höfte, H., H. Greve, J. Seurink, S. Jansens, J. Mahillon, C. Ampe, J. Van dekerckhove, H. Van derbruggen, M. Van Montagu, M. Zabeau, and M. Vaeck. 1986. Structural and functional analysis of a cloned delta-endotoxin of *Bacillus thuringiensis* berliner 1715. Eur. J. Biochem. 161: 272–280.

Höfte, H., P. Soetaert, S. Jansens, and M. Peferoen. 1990. Nucleotide sequence and deduced amino acid sequence of a new Lepidopteran-specific crystal protein gene from *Bacillus thuringiensis*. Nucleic Acids Res. 18: 5545.

Höfte, H., and H. R. Whiteley. 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol. Rev. 53: 242–255

Honée, G., T. van der Salm, and B. Visser. 1988. Nucleotide sequence of crystal protein gene isolated from *B. thuringiensis* subspecies *entomocidus* 60.5 coding for a toxin highly active against *Spodoptera* species. Nucleic Acids Res. 16: 6240.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (Lond.) 227: 680–685.

Lecadet, M. M., E. Frachon, V. Cusmao-Dumanoir, and H. De Barjac. 1994. An update version of the *Bacillus thuringiensis* strains classification according to H-serotypes, pp. 345. Second International Conference on *Bacillus thuringiensis*., Montpellier France, 28 August–2 September. Société Viganaise d'Imprimerie, Ganges, France.

Mihm, J. 1982. Techniques for efficient mass rearing and infestation in screening for host plant resistance to corn earworm *Heliothis zea*. CIMMYT, Mexico.

1983a. Techniques for efficient mass rearing and infestation of fall armyworm, *Spodoptera frugiperda* (J. E. Smith), for host plant resistance studies. CIMMYT, Mexico.

1983b. Techniques for efficient mass rearing and infestation in screening for host plant resistance to *Diatraea* sp. maize stem borers. CIMMYT, Mexico.

Perlak, F. J., R. W. Deaton, T. A. Armstrong, R. L. Fuchs, S. R. Sims, J. T. Greenplate, and D. A. Fihshoff. 1990. Insect resistant cotton plants. Biotechnology 8: 939–943.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual 3. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Thomas, W. E., and D. J. Ellar. 1983. *Bacillus thuringiensis* var *israelensis* crystal d-endotoxin: effect in insect and mammalian cells in vitro. J. Cell. Sci. 60: 181–197.

Vaeck, M., A. Reynaerts, H. Höfte, S. Jansens, M. De Beuckeller, C. Dean, M. Zabeau, M. Van Montagu, and J. Leemans. 1987. Transgenic plants protected from insect attack. Nature (Lond.) 328: 33–37.

Van Rie, J., S. Jansens, H. Höfte, D. Degheele, and H. Van Mellaert. 1989. Specificity of *Bacillus thuringiensis* delta-endotoxins: importance of specific receptors on the brush border membrane of the midgut of target insects. Eur. J. Biochem. 186: 239–247.

Van Rie, J., S. Jansens, H. Höfte, D. Degheele, and H. Van Mellaert. 1990. Receptors on the brush border membrane of the insect midgut as determinants of the specificity of *Bacillus thuringiensis* delta-endotoxins. Appl. Environ. Microbiol. 56: 1378–1385.

Visser, B., E. Munsterman, A. Strokes, and W. G. Dirlese. 1990. A novel *Bacillus thuringiensis* gene encoding a *Spodoptera exigua* specific crystal protein. J. Bacteriol. 172: 6783–6788.

Received for publication 1 March 1996; accepted 9 December 1996.

Theor Appl Genet (1996) 92:163–169

S. Fennell · N. Bohorova · M. van Ginkel
J. Crossa · D. Hoisington

Plant regeneration from immature embryos of 48 elite CIMMYT bread wheats

Received: 10 March 1995 / Accepted: 7 July 1995

Abstract Forty-eight bread wheat (*Triticum aestivum* L.) released cultivars and elite advanced lines were evaluated for their ability to produce embryogenic callus using three different media. Basal N6 medium supplemented with dicamba (E1), MS medium containing 2,4-D (E3) or MS medium containing 2,4-D plus different amino acids (E5) were used for callus initiation and maintenance. Plant regeneration was achieved on basal MS medium with indole-3-acetic acid (IAA) and 6-benzylamino purine (BAP) and rooting on MS with 1-naphthaleneacetic acid (NAA). Percentage regeneration varied widely with both genotype and initiation medium, with values ranging from 2% to 94%. The number of plantlets produced per embryo ranged from 6 to 42. Thirteen genotypes showed at least 50% regeneration after culture on E5 medium; 3 genotypes after culture on E3 initiation medium and 1 after initiation on E1. After four subcultures, over a 16-week period, 41 genotypes (85%) lost their ability to regenerate plants while the remaining 7 lines (15%) retained plant regeneration potential but at reduced levels. E3 medium was found to be the best for maintaining regeneration potential after four subcultures.

Key words Bread wheat · *Triticum aestivum* · Culture medium · Embryogenic callus · Plant regeneration

Introduction

The genetic modification of plants in vitro via transformation techniques is largely dependent on the ability of the tissue to regenerate whole plants. Wheat immature embryos are the most efficient tissue sources for regenerating whole plants in large numbers (Shimada 1978; Shimada and Yamada 1979; Gosh-Wakerle et al. 1979; Sears and Deckard 1982; Ozias-Akins and Vasil 1982, 1983). Callus induction, maintenance and regeneration from immature embryos have been shown to be genotype-dependent (Sears and Deckard 1982; Maddock et al. 1983; Lazar et al. 1983; Mathias and Simpson 1986) and in addition are strongly influenced by the components of the media used (Elena and Ginzo 1988; Hunsinger and Schauz 1987; He et al. 1989; Ozias-Akins and Vasil 1982; Carman et al. 1988). Mathias and Simpson (1986) assessed the relative contributions made by media additives and genotype in vitro and suggested that the genotype may be more significant than the medium in affecting behaviour in tissue culture.

When the success of conventional breeding techniques is combined with the potential offered by genetic transformation, it is possible that significant contributions can be made to CIMMYT bread wheat germplasm already widely grown in developing countries. The first steps towards the production of stable transgenic plants are the isolation of target genotypes and development of an efficient regeneration system for the host tissue that is to be transformed. For this reason a large number of elite bread wheats were screened for their potential to regenerate plants from immature embryos. From our study of 48 genotypes, optimal genotypes and media were selected that will be used for the initiation of transformation experiments.

Materials and methods

Forty-eight purified CIMMYT bread wheat cultivars and advanced lines were evaluated for their in vitro culture response (Table 1). The genotypes were purified by controlled selfing for two generations with selections made for uniform progeny. Plants were grown under screen-house conditions at CIMMYT's El Batan experimental station in the central highlands of Mexico (19°31′N, 98°50′ W, 2249 masl) in 1994. Immature embryos were excised from seeds harvested 13–17 days after anthesis.

Communicated by Y. Gleba

S. Fennell · N. Bohorova (✉) · M. van Ginkel · J. Crossa · D. Hoisington
CIMMYT, International Maize and Wheat Improvement Centre, Lisboa 27, Apdo. Postal 6-641, 06600, Mexico D.F., Mexico

Table 1 Callus formation and percentage regeneration for 48 elite bread wheats on MSR medium using three callus initiation media

Elite genotype	Medium E1		Medium E3		Medium E5		Mean plant regeneration(%)[a]
	Callus formation(%)	Plant regeneration(%)	Callus formation(%)	Plant regeneration(%)	Callus formation(%)	Plant regeneration(%)	
Luan (X2)[b]	100	46	98	63	100	94	67.7
Turaco (X)	100	49	100	28	100	88	55.0
Milan (X2)	100	30	96	33	100	93	52.0
Attila (X1)	100	26	100	58	98	69	51.0
Angostura (X3)	84	77	91	21	81	49	49.0
Pavon (X1)	100	45	98	30	88	57	44.0
Tia.2 (X)	98	45	100	16	98	62	41.0
Kauz (X3)	88	48	81	5	88	68	40.3
Munia (X1)	67	33	78	43	63	27	34.3
Tinamou (X)	81	11	93	4	100	86	33.7
Junco (x1)	100	22	96	38	100	36	32.0
Pastor (X)	80	2	90	22	94	71	31.7
Munia (X2)	100	9	100	65	96	20	31.3
Don Ernesto (X2)	94	4	100	12	95	77	31.0
Cettia (X2)	30	15	82	2	30	71	29.3
Kauz (X2)	87	17	85	35	100	27	26.3
Sitia (X2)	100	16	95	6	100	43	21.7
Bobwhite (X1)	100	0	96	14	100	50	21.3
Tia.3 (X)	100	7	100	38	93	18	21.0
Catbird (X1)	100	5	–	–	73	57	20.7
Weaver (X)	57	13	97	2	100	44	19.7
Seri M 82 (X2)	100	10	100	2	100	46	19.3
Dharwar Dry (X)	100	13	100	36	98	4	17.7
Ciano 79 (X)	95	2	97	3	95	40	15.0
HD2281 (X)	–	–	88	15	–	–	15.0
Loxia (X)	71	6	88	24	80	14	14.7
Filin (X)	82	7	79	19	95	17	14.3
SW87.2347 (X)	30	0	32	33	33	6	13.0
Catbird (X)	96	0	92	4	86	32	12.0
SW89.1862 (X)	91	0	96	22	85	9	10.3
Chuan Mai21 (X)	54	7	90	11	65	8	8.7
Sitella (X2)	94	0	100	6	96	19	8.3
Irena (X2)	74	0	2	0	89	24	8.0
Sibia (X)	100	6	92	7	–	–	6.5
Baviacora (X)	100	4	100	2	100	13	6.3
NL 623 (X)	89	4	61	0	83	15	6.3
Shanghai 4 (X)	100	2	98	2	100	15	6.3
Opata M 85 (X)	70	0	66	9	67	0	3.0
Star (X2)	86	0	93	0	93	9	3.0
Nanjing 7840 (X)	85	0	100	8	92	0	2.7
PRLII/CM 65531 (X)	100	0	90	0	100	8	2.7
Picus (X)	73	0	45	0	65	5	1.7
Nesser (X1)	55	0	22	0	73	2	0.7
Gamtoos (RWA)	0	0	0	0	0	0	0.0
Redwing (X)	27	0	42	0	31	0	0.0
Ducula (X2)	44	0	34	0	39	0	0.0
Nestor (X)	2	0	4	0	0	0	0.0
Bow*2/PRL (X)	0	0	8	0	5	0	0.0

[a] Minimum significant difference given by Waller-Duncan K-ratio t-test = 30.0; least significant difference at the 0.05 probability level = 27.2

[b] X indicates that the genotype was purified by two seasons of controlled selfing of "bagged" spikes. The number indicates a particular plant, if more than one was initially purified

Seeds were surface-sterilized first in 70% ethanol for 1 min, then in 20% Chlorox containing 10 drops of Tween-80/l for 20 min and rinsed three times with sterile de-ionized water. Immature embryos 1–1.5 mm in size were asceptically excised from the seeds and placed with the scutellum down on the initiation medium. Six embryos were placed in each disposable plastic petri-dish (60 × 15 mm) and approximately 50 embryos were plated per treatment with three media being used for each genotype. As there were not always sufficient embryos available for callus initiation on all three media, data from 47 genotypes are available for media E1 and E3 and from 46 genotpes for medium E5.

The media used were based on N6 (Chu et al. 1975) and MS media (Murashige and Skoog 1962) with modifications as described by Bohorova et al. (1995). The pH of all the media was adjusted to 5.7 with NaOH, and 0.8% agar (Bacto) was added before sterilization. The cultures were incubated in the dark at 26 °C. After 3 weeks on initiation medium (E1, E3 or E5) each piece of callus was cut in two; one-half being placed on regeneration medium and the other half placed on the maintenance medium (the same as that used for callus initiation). Callus cultures were subcultured to fresh media every 3–4 weeks.

Plantlets were regenerated by transferring calli to basal MS medium supplemented with 0.5 mg/l indole-3-acetic acid (IAA),

1 mg/l 6-benzylamino purine (BAP), 40 mg/l thiamine, 150 mg/l L-asparagine and 2% sucrose (MSR medium). When the plantlets had 4–6 leaves they were transferred to a rooting medium comprised of half-strength MS medium supplemented with 1 mg/l l-naphthaleneacetic acid (NAA) and 2% sucrose (MSE medium). The pH of the medium was adjusted to 5.7 with NaOH before sterilization and the addition of 0.8% agar (Bacto). Calli were incubated in a culture growth room at 26°C using a 16:8-h light:dark photoperiod. The plantlets with good root formation were transferred to Jiffy pots and kept in a growth chamber for a week before being transplanted to larger pots and grown to maturity in the greenhouse.

The percentage callus formation was defined as the number of the embryos forming callus divided by the total number of embryos plated on the media. Percentage plant regeneration was calculated as the number of calli producing at least 6 plantlets, divided by the total number of calli produced. One or 2 plants produced per half a callus could have been due to development only from meristematric tissue and production of three or more plants per callus half (6 plantlets per callus) was thought to represent regeneration via the tissue culture process. Regeneration potential was assessed after 3 weeks on MSR medium following the initial 3-week period of embryogenic callus culture and again after four subcultures (approximately 16 weeks) on maintenance medium.

Statistical analysis

Analysis of variance, using the size of the sample as a covariate, and the Waller-Duncan K-ratio *t*-test (Waller and Duncan 1969) were used to determine significant differences in regeneration potential between the 48 genotypes tested and the three initiation media used.

Results

With the exception of the line 'Gamtoos' (RWA), all bread wheats produced callus from immature embryos on at least one of the media. The first observed stage in the production of callus tissue was enlargement of the scutellar surface, which took place approximately 1 week after culture initiation, followed by the production of embryogenic callus after 2–3 weeks (Fig. 1). On N6 medium supplemented with dicamba (E1), embryos tended to form friable, white embryogenic calli, although sometimes more watery, less embryogenic calli were observed with some genotypes. E3 and E5 media produced calli very similar to each other; either pale-yellow or yellow nodular embryogenic calli were observed. Not all of the calli produced were embryogenic, which accounts for the differences observed between the proportion of embryos producing calli and the proportion of calli producing regenerated plants (as observed with the genotype Seri M82) (Table 1). With some genotypes, the callus tissue either produced roots prematurely or became necrotic; this most frequently occurred on medium E5. In terms of the number of embryos forming callus, there were no significant differences between the three media across all genotypes ($P < 0.001$).

The embryogenic calli were transferred to the regeneration medium (MSR) where green parts appeared on a proportion of the calli after 4–7 days (Fig. 2), which developed into plantlets after approximately 3 weeks (Fig. 3).

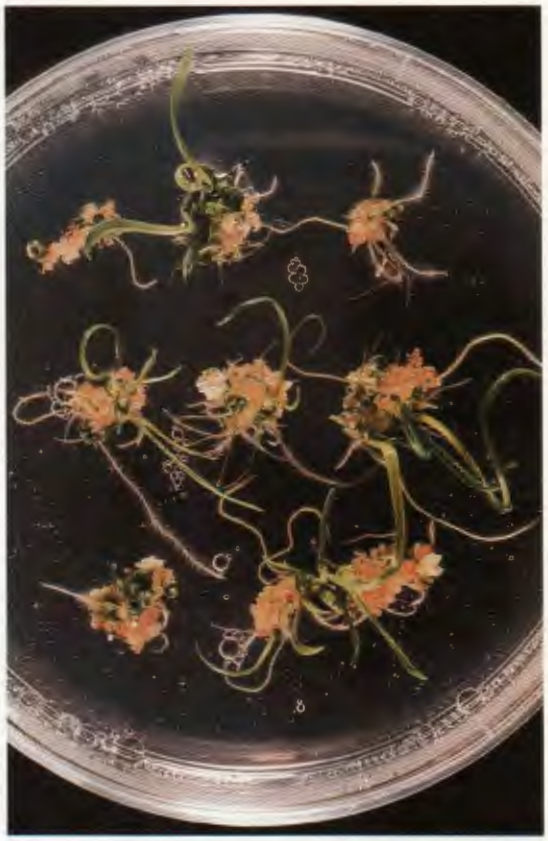
The composition of the callus initiation medium appeared to be an important factor for subsequent plantlet regeneration on MSR medium. Taking the means of the regeneration potential for each genotype across all initiation media and for each media across all genotypes we found that there were statistically significant differences between regeneration percentages for both genotypes and media (both at the 1% level). The covariate analysis (data not shown) indicated that sample size (the number of embryos plated on each medium for each genotype) accounted for a sizeable portion of total variability with regards to percentage regeneration and was statistically significant (at the 5% level).

Percentage regeneration varied widely with respect to both genotype and initiation medium with values ranging from 2% to 94% (Table 1) and number of plantlets produced per embryo ranging from 6 to 42. Sixteen genotypes gave a minimum of 50% regeneration on at least one medium. A value of 50% for regeneration potential was chosen as a minimum since genotypes with values less than this are likely to produce low numbers of plants after transformation. 'Angostura' was the only genotype producing more than 50% regeneration on medium E1. Three genotypes showed more than 50%regeneration on medium E3 and 13 on E5 medium. Only 'Attila' and 'Luan' showed greater than 50% regeneration on two media (E3 and E5) with 'Luan' giving the highest regeneration potential (94% on medium E5) of all the genotypes tested. The genotypes with the highest regeneration efficiencies (producing at least 20 plants per embryo) were 'Nanjing 780', 'Luan', 'Tia.2', 'Picus', 'Tinamou', 'Milan' and 'Catbird'. For each medium there were genotypes that did not produce embryogenic calli (or only formed calli producing fewer than 6 plantlets per embryo). For medium E1, 17 genotypes from 47 (36%) showed no regeneration; for E3, 11 from 47 (23%) and for E5, 7 from 46 (15%).

Table 1 shows the mean regeneration potentials for all 48 genotypes across the three media (taking into account sample size). 'Luan', 'Turaco', 'Milan' and 'Attila' were the best genotypes, with a mean regeneration potential exceeding 50% across the three media tested, while 'Angostura', 'Pavon', 'Tia.2' and 'Kauz' (X3) showed between 40–50% regeneration. Other genotypes, for example 'Tinamou', had a low mean regeneration potential (20.7%) although on one of the media, in this case E5, the percentage regeneration was much higher (86%) (Table 1). Similar results were observed for 'Cettia', 'Don Ernesto', 'Pastor, Munia' (X2), 'Catbird' and 'Bobwhite' where mean regeneration potentials were fairly low, but regeneration on one of the media was high. In contrast, the ninth highest performing variety was 'Munia' (X1), which responded similarly on all media but gave regeneration potentials of less than 50% on each medium. For the purpose of choosing genotypes and media for future transformation work it is desirable to select a genotype that gives high plant regeneration even if it does so on only one medium.

The Waller-Duncan test was also used to compare mean regeneration potentials on the three initiation media across all genotypes. For E5 medium, the mean

1



2

3

Fig. 1 Embyogenic callus formation from an immature embryo of 'Attila' on E3 medium **Fig. 2** Embryogenic calli from immature embryos of 'Attila' possessing regeneration potential on MSR medium **Fig. 3** Regenerated plantlets from 'Attila'

regeneration potential following callus initiation was 32.5%, which was significantly higher than that for E3 (15.4%) and E1 (12.4%) (minimum significant difference = 6.3%). Therefore, across all genotypes, medium E5 produced the highest proportion of embryogenic calli able to give rise to whole plants. However, after four subcultures over a 16-week period, only 7 lines possessed the ability to regenerate plantlets (Table 2), and most levels were substantially reduced. 'Filin', maintained on medium E3, was the genotype least affected by subculturing, which caused only a 10% reduction in regeneration potentials compared with that of 'Tinamou', whose regeneration potential was reduced by 96% after four subcultures. E3 medium was significantly better than E5 for sustained maintenance of regeneration potential. Regeneration on MSR medium, after initiation and 16 weeks of subculturing on E1, failed to produce plantlets in any of genotypes.

Discussion

Previous work by Bohorova et al. (1995) showed that media E1, E3 and E5 were able to produce embryogenic calli in 6 elite bread wheat varieties. In the current study, in which 48 elite bread wheats were evaluated, the findings of Bohorova et al. (1995) were confirmed, with each medium producing embryogenic callus for most lines. Although media E1, E3 and E5 responded equally well as initiation media for the production of calli, the proportion of this calli which was embryogenic and gave rise to regenerated plantlets varied with each medium. The calli produced on E1 medium were often non-compact and watery with few, if any, embryonic structures on the surface. E1 medium is basically N6 medium (Chu et al. 1975) supplemented with dicamba. Some workers have found dicamba to be of value in the induction of embryogenesis in wheat (Maddock et al. 1983; Hunsinger and Schauz 1987). However, in our study, E1 medium seemed to be the least likely of the three to induce and maintain embryogenesis.

Both E3 and E5 media are based on MS medium (Murashige and Skoog 1962) and are supplemented with 2,4-D, a commonly used growth regulator in wheat tissue culture (O'Hara and Street 1978; Ahloowalia 1982; Carman et al. 1988; Elena and Ginzo 1988). The two media differ in several ways, including the type of supplementary amino acids. In addition, E5 contains kinetin and GA_3, which E3 lacks. Though E5 medium was the best medium for induction of embryogenesis, E3 appeared to contain essential components that supported the maintenance of embryogenesis for 16 weeks (at least for 6 genotypes). Somatic embryogenesis is a complex morphogenic process, and its occurrence in tissue culture is due to a combination of several factors. Because the three initiation media used in this study are significantly different from each other in various respects it is not possible to attribute successful embryogenesis to any particular media component. In addition, at the stage of callus initiation from immature embryos, the most important factor regulating the percentage callus production seems to be the genotype rather than the media. Subsequently, in addition to the genotype, the callus initiation medium was also shown to greatly affect the ability of the embryogenic calli to regenerate plantlets (both initially and after four sub-cultures).

The fact that on the same medium some genotypes produced embryogenic calli and regenerated plants while others were completely incapable of regeneration suggests that there are genetic components controlling this trait. Several researchers have attempted to discover genes or regions within the wheat genome responsible for favourable tissue culture responses such as high plant regeneration potential (Galiba et al. 1986; Mathias and Fukui 1986; Felsenburg et al. 1987; Higgins and Mathias 1987; Kaleikau et al. 1989a, b; Langridge et al. 1991; Ben Amer et al. 1992). Chromosomes 7B, 7D and 1D (Galiba et al. 1986), 1RS (Langridge et al. 1991), 1BS, 2BS, 6BL (Felsenburg et al. 1987), 4BL (Mathias and Fukui 1986; Higgins and Mathias 1987) and 2DL, 2AL, 2BS and 2BL (Kaleikau et al. 1989a, b; Ben Amer et al. 1992) have all been identified as possessing regions responsible for these desirable tissue culture response characteristics. Mathias and Atkinson (1988) suggested that allelic variation in wheat at the *Rht/Gai* gene (reduced height/gibberellic acid insensitivity) may

Table 2 Elite genotypes possessing regeneration potential on MSR medium after four subcultures over a 16-week period

Genotype	Callus initiation medium	Plant regeneration(%) after 3 weeks	Plant regeneration(%) after 16 weeks
Tinamou (X)[a]	E5	86	3
Luan(X2)	E3	63	4
Attila(X1)	E3	58	17
Milan(X2)	E3	33	3
Angostura(X3)	E3	21	2
Filin(X)	E3	19	17
Chuan Mai 21 (X)	E3	11	5

[a] X indicates that the genotype was purified by controlled selfing

effect callus growth, somatic embryogenesis and plant regeneration via an effect on hormonal metabolism. In addition, Ben Amer et al. (1992) showed that the semi-dwarfism allele *Rht8* has a minor effect on wheat callus growth and regeneration ability and the allele *ppd1* (day length sensitive) has a major effect on the same tissue culture traits. It can be suggested that one or more of these genes are being expressed in the highly regenerable genotypes on at least one of the media. Of the genotypes that performed well in terms of regeneration ability and efficiency, 'Nanjing 7840', 'Luan' and 'Catbird' are of Chinese origin and 'Luan' and 'Tia.2' contain an alien grass in their pedigrees. These factors may be contributing to high regeneration potentials. It is also possible that the 1BL/1RS translocation could be partially responsible for such desirable tissue culture performance. Further work will be conducted at CIMMYT in an attempt to find the chromosomal regions responsible for high regeneration ability as this could aid in the future selection of genotypes for transformation. It is also possible that certain media components may regulate the expression of some of these genes as exemplified by 'Cettia' where regeneration potentials were significantly higher on one medium than on the other two. In contrast, this is unlikely with other genotypes, for example 'Munia' (X1), where regeneration potentials were not statistically different on the three different media used. In such cases the genotype, rather than the media, is the major factor influencing the regeneration potential of embryogenic calli. Wheat cultivars are known to differ in levels of cytokinins and auxins at seed maturation (Gale 1979; Mounla 1979). The better response of some lines, such as 'Luan', may therefore be due to the endogenous phytohormonal content and its regulatory influence.

The aim of this work was to identify bread wheat genotypes with the greatest potential for regeneration and select the best lines for transformation experiments. It is likely that after transformation, using the biolistic approach, levels of regeneration will be reduced due to disruption of cellular processes. After bombardment, 9 weeks may be required for three cycles of selection to identify transformed from non-transformed material. Hence, the selected genotypes should not only have high regeneration potential, but the latter should be maintainable for at least 9 weeks. Our study has shown that only 7 lines retained their ability to regenerate plantlets after 16 weeks of subculturing.

Two genotypes, 'Attila' and 'Luan', have been selected for initial transformation experiments because they possess high regeneration potentials and also because of their importance to CIMMYT wheat breeding programmes. 'Attila' was one of the best performing lines in tissue culture, and initial transformation experiments have shown that its plantlets can be regenerated after the three cycles of selection following particle bombardment using E3 as the callus initiation medium.

Acknowledgements The authors would like to thank B. Luna, L. Huerta and M. Ramos for their excellent technical assistance and Wolfgang Pfeiffer and Mujeeb Kazi for their useful comments on this manuscirpt. They acknowledge the Overseas Development Administration (ODA), UK, for funding S. Fennell on the APOS scheme.

References

Ahloowalia BS (1982) Plant regeneration from callus culture in wheat. Crop Sci 22:405–410

Ben Amer IM, Worland AJ, Schlegel R (1992) In vitro culture variation of wheat and rye caused by genes affecting plant growth habit in vivo. Euphytica 61:233–240

Bohorova NE, van Ginkel M, Rajaram S, Hoisington DA (1995) Tissue culture response of CIMMYT elite bread wheat varieties and evaluation of regenerated plants. Cereal Res Comm (in press)

Carman JG, Jefferson NE, Campbell WF (1988) Induction of embryogenic *Triticum aestivum* L. calli. I. Quantification of genotype and culture medium effect. Plant Cell Tissue Org Cult 12:83–95

Chu CC, Wang CC, Sun CS, Hsu C, Yin KC, Bi CV (1975) Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen source. Sci Sinica 18:659–668

Elena EB, Ginzo HD (1988) Effect of auxin levels on shoot formation with different embryo tissues from a cultivar and a commercial hybrid of wheat (*Triticum aestivum* L.). J Plant Physiol 132:600–603

Felsenburg T, Feldman M, Galun E (1987) Aneuploid and alloplasmic lines as tools for the study of nuclear and cytoplasmic control of culture ability and regeneration of scutellar calli from common wheat. Theor Appl Genet 74:802–810

Gale MD (1979) Genetic variation for hormonal activity and yield. In: Spiertz JHJ, Kramer Th (eds) Crop physiology and cereal breeding. Centre for Agric Pub and Documentation, Wageningen, The Netherlands, pp. 29–34

Galiba G, Kovacs G, Sutka J (1986) Substitution analysis of plant regeneration from callus culture in wheat. Plant Breed 97:261–263

Gosh-Wackerle G, Avivi L, Galun E (1979) Induction, culture and differentiation of callus from immature rachises, seeds and embryos of *Triticium*. Z Pflanzenphysiol 91:267–278

He DG, Yang YM, Scott KJ (1989) The effect of macroelements in the induction of embryogenic callus from immature embryos of wheat (*Triticum aestivum* L.) Plant Sci 64:251–258

Higgins P, Mathias RJ (1987) The effect of the 4B chromosomes of hexaploid wheat on the growth and regeneration of callus cultures. Theor Appl Genet 74:439–444

Hunsinger H, Schauz K (1987) The influence of dicamba on somatic embryogenesis and frequency of plant regeneration from cultured immature embryos of wheat (*Triticum aestivum* L.). Plant Breed 98:119–123

Kaleilau EK, Sears RG, Gill BS (1989a) Monosomic analysis of tissue culture response in wheat (*Triticum aestivum* L.). Theor Appl Genet 78:783–787

Langridge P, Lazzeri P, Lörz H (1991) A segment of rye chromosome 1 enhances growth and embryogenesis of calli derived from immature embryos of wheat. Plant Cell Rep 10:148–151

Lazar MD, Collins GB, Vian WE (1983) Genetic and environmental effects on the growth and differentiation of wheat somatic cultures. J Hered 74:353–357

Maddock SE, Lancaster VA, Risiott R, Franklin J (1983) Plant regeneration from cultured immature embryos and inflorescences of 25 cultivars of wheat (*Triticum aestivum*). J Exp Bot 34:915–926

Mathias RJ, Atkinson E (1988) In vitro expression of genes affecting whole plant phenotype – the effect of *Rht/Gai* alleles on the callus culture response of wheat (*Triticum aestivum* L.em. Thell). Theor Appl Genet 75:474–479

Mathias RJ, Fukui K (1986) The effect of specific chromosome and cytoplasm substitutions on the tissue culture response of wheat (*Triticum aestivum*) callus. Theor Appl Genet 71:797–800

Mathias RJ, Simpson ES (1986) The interaction of genotype and culture medium on the tissue culture responses of wheat (*Triticum aestivum* L.) callus. Plant Cell Tissue Org Cult 7:31–37

Mounla MAKh (1979) Phytohormones and grain growth in cereals. In: Spiertz JHJ, Kramer Th (eds) Crop physiology and cereal breeding. Centre for Agric Pub and Documentation, Wageningen, The Netherlands, pp 20–28

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Plant Physiol 15:473–497

O'Hara JF, Street HE (1978) Wheat callus culture: the initiation, growth and organogenesis of callus derived from various explant sources. Ann Bot 42:1029–1038

Ozias-Akins P, Vasil IK (1982) Plant regeneration from cultured immature embryos and inflorescences of *Triticum aestivum* L. (wheat): evidence for somatic embryogenesis. Protoplasma 110:95–105

Ozias-Akins P, Vasil IK (1983) Improved efficiency and normalization of somatic embryogenesis in *Triticum aestivum* (wheat). Protoplasma 116:40–44

Sears RG, Deckard EL (1982) Tissue culture variability in wheat: callus induction and plant regeneration. Crop Sci 22: 546–550

Shimada T (1978) Plant regeneration from the callus induced from wheat embryos. Jpn J Genet 53:371–374

Shimada T, Yamada Y (1979) Wheat plants regenerated from embryo cell cultures. Jpn J Genet 54:379–385

Waller RA, Duncan DB (1969) A Bayes rule for the symmetric multiple comparison problem. J Am Stat Assoc 64:1484–1503