MORE ON THE INTRODUCTION OF TEMPERATE MAIZE INTO EUROPE:
LARGE-SCALE BULK SSR GENOTYPING AND NEW HISTORICAL ELEMENTS

P. Dubreuil1,2,3, M. Warburton2, M. Chastanet4, D. Hoisington2, A. Charcosset1,*

1 INRA-CNRS-UPS-INAPG, Station de Génétique Végétale, Ferme du Moulon, 91190 Gif sur Yvette, France
2 CIMMYT, Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico
3 PROMAIS, Association pour l'étude et l'amélioration du maïs, 39 Chemin Virebent, 31200 Toulouse, France
4 CNRS, UMR 8171, Centre de Recherches Historiques et Juridiques - Université Paris I, 9 rue Malher, 75181 Paris cedex 04, France

Received May 30, 2005

ABSTRACT - The commonly accepted scenario for the spread of maize into Europe is a single introduction followed by a slow adaptation to temperate climates. With the purpose of clarifying the origins and modalities of maize introduction in Europe, we conducted an extensive survey of 275 maize populations from both American and European origins by using microsatellite (SSR) analysis on pools of individuals. Our data strongly support two major sources, one from the Caribbean and one from northeastern America, giving rise to most of the open pollinated varieties cultivated in Europe. A detailed historical analysis confirms a first introduction of maize in southern Europe by Columbus, and suggests that introduction(s) of temperate northeastern American maize should have occurred at the beginning of the 16th Century at the time of Spanish or French expeditions. In addition, our results reveal that maize varieties cultivated at middle latitudes in Europe likely resulted from hybridization between the southern and northeastern European varieties.

KEY WORDS: Molecular markers; Genetic diversity; Adaptation; Landraces; Evolutionary relationships.

INTRODUCTION

Maize was first brought to Europe from the West Indies by Columbus in 1493 (ANGIERA, 1907, 1st complete ed. 1530). Material from this first introduction was unadapted to the European temperate climate, and its cultivation probably remained confined to warm European areas such as southern Spain (i.e. Andalusia) and the Mediterranean periphery. Surprisingly, no more than a few decades later, maize was being extensively cultivated in northern European regions, such as Germany in 1539 (FINAN, 1948; GAY, 1984, 1999). Until recently, European maize has been considered by historians (HERMARDINQUER, 1963) and linguists (FRANCOINE, 2000) to be the product of a single introduction into southern Spain. This scenario, which would require rapid adaptation, seems unlikely given the time frame (a few decades) from the initial introduction to the cultivation of maize in northern European regions. Recently, the analysis of a set of more than 200 American and European landraces at several RFLP (Restriction Fragment Length Polymorphism) loci was performed by analyzing bulked samples of 15 individuals from the same landrace. This bulk strategy made it possible to estimate accurately within-population allelic frequencies based on the RFLP probe hybridization signal intensities of the population samples (DUBREUIL et al., 1999). Results indicated that European maize diversity traced back to at least two distinct introductions, one of tropical maize from the West Indies and other(s) from temperate maize from northeastern America (REBOURG et al., 2003). A joint historical analysis showed that the second introduction(s) should have occurred much earlier than previously hypothesized by BRANDOLINI (1970), at a time consistent with temperate maize cultivation in Germany. This bulk approach was shown to be a powerful means to analyze numerous populations from European gene banks, and brought new, relevant information for the maintenance and utilization of maize diversity. In order to facilitate further large scale analyses of diverse maize populations, we adapted the bulk sample approach to the specificities of microsatellites, Simple Sequence Repeats (SSRs) and evaluated its accuracy using control samples. Besides practical considerations, use of SSRs is significant because of their greater accuracy in distinguishing among accessions from a similar origin, to resolve evolutionary relationships between landraces and trace major paths of dispersal of maize from its center of domestica-

* For correspondence (charcos@moulon.inra.fr).
tion (Matsuoka et al., 2002b). We used this approach to further investigate the origins of maize in Europe by analyzing 275 European and American maize accessions with 24 SSR loci. A thorough historical investigation was conducted in parallel to further establish the modalities and timing of maize introductions into Europe.

**MATERIALS AND METHODS**

**Plant materials**

In order to confidently identify American sources of European maize, we chose accessions representing a broad range of genetic diversity within both continents. In Europe, we first grouped landraces into 8 geographical areas, including all primary centers of introduction (Spain, Portugal, and Italy) and traditional areas of cultivation. Within each group, accessions were then sampled to represent major morphologically distinct landraces (Gauthier et al., 2002). A detailed description of the 131 European populations has been published previously (Gauthier et al., 2002; Rebourg et al., 2003). Because all landraces in the Americas, regardless of their origin or adaptability to European conditions, could potentially have contributed to the development of maize in Europe, we sampled 144 accessions among all identified racial groups from the Americas. This sample includes maize types adapted to temperate conditions from southwestern and northeastern USA, highland Mexico, core Andes, southern Chile, and Argentina, as well as maize types adapted to tropical conditions from lowland Mexico and Guatemala, the Caribbean, and Central America. This set of American populations includes 88 populations considered in a preliminary study (Rebourg et al., 2003) and 56 additional central American populations that represent complementary origins. Fig. 1 displays the locations of all accessions sampled in this study. Passport data for the plant materials, including accession numbers, landrace designations, geographical coordinates, and elevation (when recorded) are available on request from the corresponding author.

**SSR analysis**

DNA was extracted at INRA, Le Moulon, Gif-sur-Yvette, France. Each population was represented by two pools of 15 plants each. As described in Rebourg et al. (2001), each pool was prepared by bulking an equal amount leaf material from each 15 individuals. DNA extraction was performed after the pooling step. Molecular analyses were carried out at CIMMYT’s Applied Biotechnology Center, El Batán, Mexico. Each accession was represented by 2 independent DNA samples, except for accessions

---

**FIGURE 1** - Geographical distribution of accessions considered in the study. When exact geographical coordinates were unknown, we used coordinates of cities representative of the countries or regions where the populations were collected.
5, 645, 687, and 719, which were represented by only one sample. All SSR loci were amplified separately by PCR by using a standard protocol available on the CIMMYT web site, www.cimmyt.cgiar.org. Fluorescently labeled PCR products were electrophoresed in an ABI Prism® 377 automatic DNA sequencer (Applied Biosystems, Foster City, CA), and analyzed with Genescan v3.0 and Genotyper v2.1 (Applied Biosystems). The Genescan software was used for checking gel quality, sizing bands and generating input files for the Genotyper software. This latter application was used to define bins (allele categories) and extract raw data needed for subsequent analyses. The Genotyper category filtering tools were not used, and all the bands falling into a defined bin were kept for further steps of allele frequency estimation, regardless of whether the bands were known alleles or extra bands, and had strong or faint intensities. SSR loci were mostly chosen with trinucleotide repeats or greater, since these have been shown to be less sensitive to stuttering phenomena than are di-nucleotide SSRs (DANIELS et al., 1998; PERLIN et al., 1995). In order to identify SSR loci that were the most appropriate for estimating allele frequencies within pooled-DNA samples, we performed a pilot study similar to that described by DUBREUIL et al. (1999). Briefly, we compared expected and estimated allele frequencies (see the next section) at each locus over several controlled DNA pools. Four pairs of unrelated distant inbred lines were considered (CML78-CML249; CML202-CML366; CML206-CML299; CML176-216). For each pair, bulks with proportions of the first inbred of 0.025, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50 were prepared from appropriate numbers of leaf disks with equal size. More complex pools were obtained by varying the contribution of inbred CML78 relatively to the seven other lines in the same proportions. We retained only the loci for which (i) the number of extra bands (stutter bands, +A bands, and background noise) recorded as true alleles was minimal, (ii) the correlation coefficient between expected and estimated allele frequencies was higher than 0.8, and (iii) the competition effects between alleles were not large. These criteria allowed us to retain 24 out of 46 loci assayed on controlled DNA pools (Table 1). Additional information on SSR loci used in this study can be found on the MaizeGDB web site, www.maizegdb.org.

### Allele frequency estimation

Raw data from Genotyper were first submitted to a preliminary filtering operation to remove the background noise. Bands that comprise the background noise have small peak heights that barely emerge from the baseline. They are thus associated with low modulation score values (MSV, see Applied Biosystems;
Genotype User’s Guide) and are therefore easily identified. Since some bands with large peak heights can sometimes be associated with low MSV, only bands that had both a low MSV (<10%) and a small peak height (<500) were removed. Allele frequencies within DNA pools were estimated from the remaining bands, either directly from raw data, or after correcting them for overlaps with stutter bands if present. For this purpose, we adopted a simple deconvolution method originally described by Ledec et al. (1995). This method requires the estimation of a parameter that measures the intensity of the stuttering phenomenon, and which is defined as the ratio of peak heights between the α-stutter band (the stutter band one repeat unit smaller than the allele) and the allele itself. Data from the pilot study were used to estimate $r$ values at loci sensitive to stuttering. Each locus-specific $r$ value was estimated as the average of individual-allele estimates. We assumed that $r$ was constant for a given locus among alleles and among experiments. Among the selected loci, 58% were insensitive to stuttering ($r = 0$), 21% revealed low sensitivity to stuttering ($r = 0.03$–0.12), and another 21% revealed moderate to high sensitivity to stuttering ($r = 0.17$–0.55) (Table 1). In case of stuttering, for each lane, observed peak heights were corrected sequentially starting from the band with the highest molecular weight (MW), by using the following formulas:

$$Q'_{n} = \hat{Q}_{n},$$
$$Q'_{n-1} = \hat{Q}_{n-1} - r\hat{Q}_{n},$$
$$Q'_{n-2} = \hat{Q}_{n-2} - r\hat{Q}_{n-1},$$
$$...$$
$$Q'_{n-m} = \hat{Q}_{n-m} - r\hat{Q}_{n-m+1},$$

where $n$ is the length in repeat units of the band with the highest MW, $m$ is the number of repeat units between the two extreme bands, $\hat{Q}$ and $Q'$ are the observed and corrected peak heights, respectively, and $r$ is the stutter ratio estimated at this locus. Corrected peak heights were set to 0 if negative. Allele frequencies were then estimated from corrected peak heights within each pool as done for RFLPs (Rebourg et al., 2001), and those found lower than the minimal expected frequency (~0.033) within a pool of 15 plants were set to 0. Finally, frequencies within accessions were estimated as the average frequencies of the two pools of an accession.

In this study, many loci followed a mixed model of evolution combining stepwise mutation (SM) and infinite allele (IA) models as previously reported (Matsuda et al., 2002a). Whereas a pure SM model can only generate one set of bands, multiples of the number of base pairs contained in the repeat unit, the mixed model can generate as many different sets as the number of base pairs of the repeat unit. In this situation, each set of bands was treated separately. Methods used for filtering raw data and estimating allele frequencies in the presence of stutter bands have been implemented in a SAS program available on request from the authors.

**PCA analysis**

Genetic relationships among maize accessions were investigated by Principal Components Analysis (PCA). The analysis was performed on the among accession variance-covariance matrix of allele frequencies with the PRINCOMP procedure of the SAS software (SAS, 2000). Missing allele frequencies within accessions (1.43%) were replaced by corresponding average frequencies within the whole set of accessions before running this analysis.

**Genetic distance estimation and classification**

Since most SSR loci exhibited a complex model of evolution, the distance between two accessions was estimated by the Modified Rogers’ Distance (Rogers, 1972), which is free of any assumptions about mutation model at marker loci. We used the following unbiased estimator adapted from Ghiéarde et al. (1998) for multiallelic loci:

$$D_{ij} = \frac{1}{2L} \sum_{l=1}^{L} \sum_{a=1}^{A_l} (\hat{f}_{il} - \hat{f}_{jl})^2 - \frac{1}{2L} \sum_{l=1}^{L} \sum_{a=1}^{A_l} \left( \frac{\hat{f}_{il} \hat{f}_{jl}}{N_l^{-1}} + \frac{\hat{f}_{jl} \hat{f}_{il}}{N_j^{-1}} \right),$$

where $L$ is the number of loci, $\hat{f}_{il}$ ($\hat{f}_{jl}$) the allele frequency of the allele $a$ at the locus $i$ for the accession $j$, and $N_i$ ($N_j$) the number of individuals sampled in the variety $j$.

Pairwise interaccession distances were computed using a SAS program (LCMDV, www.cimmyt.cgiar.org/english/webp/research/abc/index.htm). The resulting distance matrix was subjected to the Ward’s hierarchical clustering algorithm included in the LCDMV computer program and a dendrogram was obtained with the TreeView software available at http://taxonomy.zoology.gla.ac.uk/rod/rod.html.

**Genetic diversity estimation**

Nei’s unbiased genetic diversity (Nei, 1978) was computed for each locus ($H_e$) and for all the loci ($H_o$) as

$$H_e = \frac{1}{L} \sum_{l=1}^{L} H_{dl}$$

$$H_o = \frac{1}{L} \sum_{l=1}^{L} H_{dl} = \frac{\sum_{i=1}^{N_f} A_i f_{il}}{2n_i - 1} \left( 1 - \sum_{a=1}^{A} \left( f_{il}^2 \right) \right),$$

where $f_{il}$ is the frequency of allele $a$ at locus $l$ in the whole population, $A_i$ is the number of alleles detected at this locus, $L$ is the total number of loci analysed and $n_i$ is the number of individuals characterised for locus $i$. Genetic diversity within a given population $i$ was estimated similarly at each locus ($H_e^i$) and for all loci ($H_o^i$). In this case, $f_{il}$ is the frequency of allele $a$ at locus $l$ within the population $i$ considered and $A_i$ is the number of alleles detected at this locus within this population.

**Model-based analysis of population structure**

Groupings of populations identified from PCA and clustering analysis were further investigated by using the computer program STRUCTURE (Pritchard et al., 2000) (http://purl.bsd.uchicago.edu/). For each individual plant sampled within a population, this program provides an estimate of the fraction of its genome derived from different ancestral genetic pools. It is therefore useful in deciding whether distance-based clusters are valid or not, and which genetic pools are likely to derive from a hybridization between other pools. Since the method requires multilocus genotypes of individual plants to be known, and we only had information on allele frequency at each locus, we created dummy genotypes of each population as follows: 40 multilocus haplotypes were sampled within the estimated allele frequency distribution for each population, and randomly associated by pair to reconstruct 20 multilocus genotypes from each population. Creating those dummy genotypes relied on two implicit assumptions: Hardy-Weinberg equilibrium within populations and absence of linkage disequilibrium between loci within populations. The program STRUCTURE was run with a burn-in period of 30,000 and 10^6 iterations, as recommended by the authors (Pritchard et al., 2000).
RESULTS AND DISCUSSION

Accuracy of the pooled sample method for estimating SSR allele frequencies

Although DNA pooling methods have proven to be accurate for estimating maize diversity at RFLP loci, one could expect several drawbacks when applying them to SSR loci. First, as a consequence of slippage of the Taq polymerase on the repeated sequences, SSR alleles are often multiple banded. As a result, when several alleles are amplified simultaneously, overlapping patterns can make it difficult to distinguish true alleles from stutters. Second, the occurrence of null alleles, sometimes at a high frequency, may decrease the accuracy of frequency estimates. Third, differential amplification among alleles can greatly skew actual patterns of frequencies within pooled samples. In order to test for the accuracy of this method and to evaluate the potential drawbacks of applying it to SSRs, we performed a pilot study on controlled pools of inbred lines. Out of 46 loci assayed, 8 were not investigated further due to a high level of missing data including true null alleles and failed PCR amplifications. The remaining 38 loci were individually checked and 14 additional loci were discarded for the following reasons: lack of polymorphism, unexplained banding pattern complexity, unstable stutter ratios, or strong competition effects between alleles, leaving 24 remaining loci. The correlation between the expected and estimated allele frequencies in the controlled pools of inbred lines for the 24 selected loci varied from 0.79 for locus phi093, which showed rather strong competition, to 0.99 for the loci phi014 and phi115 (Table 1). At the level of populations, the average distance between the pools within a given population was very low (0.034). In comparison, only 0.08% of all pairwise inter-population distances were lower. Thus, allele frequencies estimated within the pools representing a given population were very similar, indicating that the pooled-sample method is highly repeatable and that characterization of a single pool could be performed in subsequent experiments. Finally, we used a set of common accessions to compare SSR distances estimated by using these 24 SSR loci with those estimated previously with 29 RFLP loci (Rebourg et al., 2003). The simple correlation coefficient (0.73; Mantel test at p<0.001) can be considered high given the number of loci considered for each technique and confirms that possible PCR-related problems are unlikely to affect final results.

Genetic diversity at SSR loci

The number of alleles per locus varied between 2 (locus phi046) and 13 (locus phi085) and averaged 7.8. Considering the large number of distinct genetic origins included in this set of populations, mean number of alleles per locus was quite low, and contrasted with estimates reported in previous studies that included a broad American diversity. For instance, the average number of alleles per RFLP locus found by (Rebourg et al., 2003) was 13.1 within a subset of the populations considered in this study. The low number of alleles can be attributed to both the type of SSR loci that were used for genotyping the populations, and the way the alleles were identified. In fact, all SSR loci except one (phi112) were chosen among tri-nucleotide or higher repeat motifs in order to limit the stuttering often associated with di-nucleotide SSR loci. However, tri-nucleotide or higher repeat SSR loci are less polymorphic than dinucleotide repeat SSRs (Vigouroux et al., 2002). In addition, alleles have been identified from the quantitative analysis of banding patterns obtained on bulks of 15 individuals. Such a poolplexing approach might discard some bands among the faintest in intensity.

In contrast, Nei’s diversity (Nei, 1978), which is barely influenced by rare alleles, ranged between 0.40 and 0.85, with a mean of 0.62, and was fairly similar to previously reported values in maize for RFLP (Rebourg et al., 2003) or SSR loci (Matsuoka et al., 2002a; Labate et al., 2003). Diversity within populations (Nei, 1978) encompassed a wide range of values from 0.09 to 0.56. Interestingly the least diverse populations were mainly from northeastern Europe (0.32 on average) and US (Northern Flint material, 0.35 on average), whereas the most diverse populations were mainly from Mexico and Guatemala (0.47 on average). On average, populations from the Americas were significantly (t-test at p<0.001) more diverse (0.44) than those from Europe (0.38), consistent with the evolutionary history of maize from its domestication in Mexico, its diffusion in the Americas, and its subsequent introduction in Europe.

Structure among American accessions

We first performed clustering analyses and PCA on a set of 144 North American, South American and Caribbean populations (Fig. 2). Major trends in the structure of this material were highly consistent with those recently reported by Matsuoka et al. (2002b) using a somewhat different approach (a
single individual per population and 99 loci). If the Chilean populations, not included in Matsuoka et al. (2002b), were excluded from our analysis, both studies showed a central position of Mexican and Caribbean materials and a clear differentiation of north and south American genetic pools. This is consistent with the domestication of maize in Mexico followed by southwards and northwards dispersion. Cluster analysis (results not shown) further confirmed the distinction between Mexican and Caribbean materials and revealed the split of Mexican material in two groups previously identified by a morphological survey: a Mexican Pyramidal group and a northern group (Goodman and Brown, 1988; Sanchez G. et al., 2000). In order to further test for the presence of these two groups, we included them in a model-based structure analysis (Pritchard et al., 2000) and compared this a posteriori classification with racial classification. This analysis confirmed the presence of two ancestral groups that contributed, respectively, to 40% and 60% of the northern complex genome and 58% and 42% of the Pyramidal complex genome. The distinction between the two racial complexes was therefore very loose at the molecular level. Finally, the position of Chilean populations between Peruvian and Ecuadorian material and populations from the northeastern American complex suggest that Chilean material analyzed in this study could result from hybridization between these two materials. This finding was further supported by a population structure analysis of the three origins. This analysis revealed three ancestral groups and showed that Chilean populations inherited 11% from Andean and 32% northeastern American ancestral groups, whereas Andean and northeastern American populations inherited 15% of their genome from external origins. The modalities of this hybridization need further investigation.

American sources of European maize

Principal component analysis (not shown) and clustering analysis (Fig. 3) showed several strong similarities between groups of American and European populations. First, seven southern Spanish populations were included in a group mostly composed of Caribbean material. This was consistent with historical data on the introduction of maize in the south of Spain by Columbus after his second trip in the Caribbean region. Our results confirm that this material did not spread widely, consistent with former hypotheses based on isozyme (Revilla et al., 1998) and RFLP (Gauthier et al., 2002) analyses.

A second striking result was the similarity between northeastern American populations and northeastern European populations. Maize was cultivated in northeastern America prior to the Discovery of the area by Europeans. Traditional varieties from this region form a very specific genetic group both at the morphological level (relatively short size due to earliness, high tillering, presence of long husk leaves, cylindrical ears of 8 to 10 rows) and at the molecular level (see above and Doebley et al., 1986; Rebourg et al., 2003). It also displays sweet and flour kernel mutant types; however, flint kernel type is predominant among this material. This morphological characteristic lead to the “Northern Flint” designation (Brown and Anderson, 1947), most used within the maize geneticists community. On the other hand, archeologists and ethnologists (see below) describe this genetic group as “8-10 rowed
maize” or “eastern complex” to better account for variation within this group. In contrast to the Caribbean material firstly introduced in Europe, Northern Flints were better adapted to cultivation in temperate Europe climates, largely because of their low sensitivity to photoperiod and early flowering in temperate regions. The flint kernel texture may also have contributed a better ability to germinate and grow in cold spring conditions (DERIEUX et al., 1989).

Finally, a wide range of European material cultivated in particular in the Pyrenees or in Spanish Galicia display no close similarity with any American material (Fig. 3). This material appeared to be intermediate between the Caribbean and Northern Flint genetic origins described above because it included specific alleles from both origins. Estimates from the STRUCTURE software indeed revealed that Pyrenean and Galician populations were composed of high proportions of European Northern Flint (17%) and southern Spanish (27%) genetic pools. A PCA performed on all accessions (not shown) revealed that three SSR alleles had a major contribution in the definition of the first axis, namely phi041-197, phi072-151 and phi085-238. It can be noted that allele phi041-197 had formally been found to be specific to Northern Flours and Flints materials and that alleles phi072-151 and phi085-238 have been found to be specific to early pop-corns, known to be allied to Northern Flours and Flints (GONZALEZ ULGADE, 1997). We investigated the geographical patterns of frequencies for these three alleles. Fig. 4 illustrates the results obtained with phi085-238. This pattern was similar among all three alleles, and is consistent with present knowledge (MATSUOKA et al., 2002b) on the origins of Northern Flint materials in southwestern American material, itself originally derived from Mexican material (results not shown). But more importantly, this analysis helped in tracing back the diffusion of Northern Flint material in northeastern Europe. It also revealed the important contribution of this material in
the establishment of maize cultivation in other European regions such as the Pyrenees, Galicia (Spain) and Northern Italy via hybridization with materials from other origins, as hypothesized by Revilla et al. (2003) based on isozyme markers. In other words, Northern Flints appear to have been a key source of alleles for adaptation to temperate and cool regions and the contribution of this material has clearly been a determinant for the cultivation of maize in Europe.

Historical analysis of origins and dispersion of Northern Flints

Historically, first references to the specific morphology of Northern Flint material (see above) were made by two German herbalists, Bock in 1539 (New Kreuterbuch von Underscheid, quoted by Fignon, 1948), and Fuchs in 1542 (Fuchs, 1549), who mentioned regular ears with 8 to 10 rows. This last author also observed that maize was relatively common in Germany at that time ‘in all gardens, almost everywhere’. We found that the same typical morphology (ears with 8 to 10 straight rows of kernels), was mentioned by Matthiolus in 1570 (Matthiolus, 1579, 1st Latin ed. 1570) for maize cultivated in Italy. Further detailed descriptions by this author (Matthiolus, 1571) about maize use for cooking polenta in the Alpine regions suggest that maize had been established there already for many years. Note that this author was also the first one to acknowledge the American origin of maize and to rule out the Eastern origin considered at that time due to the common “Turkish Korn” designation. Cultivation of Northern Flint material was therefore well estab-
lished in several regions of the periphery of the Alps by 1540-1570.

The introduction of northeastern American populations in Europe shortly after the Discovery of America by Europeans was hypothesized by Finan (1948), but with no consideration of the expeditions that could have led to this introduction. It was reconsidered recently by Rebourg et al. (2003) who pointed out the possible contributions of Giovanni Verrazano and Jacques Cartier who discovered the North Eastern American coast in 1524 and 1534, respectively. In order to investigate possible alternative hypotheses, we reconsidered the extension of Northern Flint material at the time of Discovery. Recent archaeological data have revealed that maize varieties that display the characteristic features of Northern Flints were not only cultivated in northeastern America prior to the Discovery of the New World but everywhere along the eastern coast, from the St Laurent Bay down to the north of Florida (Ruhl, 1993). Illustration (Fig. 5) of maize cultivation by Native Americans in “Virginia” (presently North Carolina) in 1585 shows a wide range of planting times at the same location (from ripe to newly sown corn in nearby fields). This was likely made possible by the early flowering of these varieties. From there, Spanish expeditions to Florida and Georgia in the very early 16th Century (Bernand and Gruziniski, 1991) may have given opportunities to introduce Northern Flint varieties in Spain. The strong political and commercial relationships between Spain and Germany at this period (Pérez, 1994) might have then promoted the diffusion of this maize into the regions mentioned by Bock and Fuchs. Similarly, commercial relationships between Spain and Italy, Venice in particular, likely facilitated the diffusion of this same maize to the Italian regions mentioned by Matthiolus.

Concerning the expeditions to Northeastern America of Giovanni Verrazano (1524) and Jacques Cartier (1534 and then 1535-1536), Rebourg et al. (2003) have shown that both explorers could have introduced maize in France. They showed that the possible contribution of Verrazano in the European history of maize had remained unknown because of a terminological problem. In his travel report, Verrazano indeed described maize as ‘legumes’, which was the ancient term for pulses. Similar descriptions of maize occurred at this period, as illustrated by different records about the French expedition to Florida later in the mid 16th Century. Maize, also named “mil” (a French term for pearl millet), was described as having “a grain as big as a pea” and was presented as “one of the main legums (sic) of their food” (Lussagnet, 1958). Until recently, the only author to hypothesize that Verrazano had described maize was Parker (1968, 1st ed. 1910), “in the light of subsequent descriptions by other explorers”. Verrazano very likely brought maize as food for his return travel, although the evidence is incomplete in his report (as edited by Mollat du Jourdin and Habert, 1982). Our investigations showed that he reported on his discovery in Lyon in 1524 to the French king and to the Italian merchants of the city who funded his trip. Only a single
letter from a member of this community refers to what he brought back: “a young boy” and “some drugs” (MOLLAT DU JOURDIN and HABERT, 1982), which may have included maize. If this is the case, the interest of Italian merchants in agriculture and the central place of Lyon as an international trading center should have facilitated the diffusion of the plant in Europe (GASCON, 1971). It can be noted that archives of the early 17th Century (first data in 1608) show that maize was then well integrated in the agricultural system of the south of Bresse, near Lyon (PONSOT, 2003). Experimentation of rice cultivation in this region (BRUWERIN-CHAMPIER, 1998, 1st Latin ed. 1560) illustrates the strong interest of local farmers for new plants in the 16th Century, which likely facilitated the settlement of maize.

Original American northeastern varieties were replaced rapidly by more productive types in the south of the present US, after the Spanish colonization which started in 1565 (BERNAND and GRUZINSKI, 1991). Therefore, no living sample representative from the original precolombian maize varieties cultivated in this region is presently available. This makes it difficult to use genetic analyses to distinguish between the ‘Spanish’ and ‘French’ scenarios in the introduction of Northern Flints in Europe. It can, however, be reconsidered in future work by genotyping archaeological samples (JAENICKE-DESPRES et al., 2003).

CONCLUSION AND PERSPECTIVE

The pooling protocol we developed here for SSRs seems to be remarkably reliable for characterizing maize populations. The general gain of efficiency of SSR markers compared to RFLP can facilitate large scale studies. The results that were obtained in this study confirm the suitability of SSRs as a tool for understanding the mechanisms of maize adaptation and dispersion from its center of origin. The patterns of maize dispersion and evolution in America we found are consistent with those of MATSUOKA et al. (2002b), and support the presence of Northern Flint-related material in the south of South America. The combined survey of European and American material at SSR loci and a thorough historical analysis, confirms two primary sources of European maize: a Caribbean source and a northeastern American coastal source. Both Spanish and French expeditions could have introduced this material into Europe at a time consistent with reliable records on its cultivation in the Alps and their periphery, beginning in the first half of the 16th Century. Adaptation of the crop to the new environmental conditions found in midwestern Europe (cool and humid) appears to have been the result of a hybridization between these two groups, rather than a slow dispersion and adaptation process starting in southern Spain, as previously suggested. The dynamics of adaptation following this hybridization calls for further research using markers for genes involved in climatic adaptation.

ACKNOWLEDGEMENTS - We are grateful to B. Gouesnard, C. Rebourg and S. Tabia for their contribution to the choice of the genetic materials considered in this study, and to C. Dillmann for help in generating dummy genotypes from the observed allele frequencies within the populations. We are also grateful to J. Cayouette, G. Crawford, G.R. Hammel, J.P. Hart, M. Mauzé, M. Scarry and W. Sturtevant for their helpful advice on North American botany, archaeology and ethology. We are very grateful to Maud Tenaille and Ed Coe for their extremely useful comments on the manuscript. We are grateful to PROMAIS members AGROW France (Monstanto SAS), CARGILL SEMENCES (Monsanto SAS), CAUSSADE SEMENCES, CEPECO SEMENCES, EURALIS GENETIQUE, MASADOIR SEMENCES, NOVARTIS SEEDS (Sygenta), LIMAGRAIN GENETICS, PIONEER GENETIQUE, S.D.M.E. K.W.S. France, VERNEUIL RECHERCHE (LIMAGRAIN), and to the French Ministry of Foreign Affairs for supporting this program, and to J. Beigbeder and P. Ruaud for helpful coordination and discussion. This paper is dedicated to J.-P. Monod and A. Gallais for their unfailing support throughout this research program.

REFERENCES


