Principles of Field Experimentation

Berhanu Tadesse

Maize Breeder, CIMMYT-Kenya

Technicians Training from August 7-10, 2023

The Concord Hotel, Nairobi, Kenay

CGIAR

CIMMYT

Outline

- Introduction
- Steps in experimentation
- The aim of the experiment - choice of treatments
- Choice of experimental material
- Experimental vs observational unit
- Principles of experimentation
- The experimental design process
- Common Experimental Designs for maize breeding trials
- General Breeding Scheme in CIMMYT maize
- Check Strategy at CIMMYT
- Summary

Introduction

- An **experimental study** is one in which the researcher deliberately **manipulates** certain **study conditions** in order to induce and measure a response in **experimental material**.
  - The manipulated conditions are often referred to as **treatments** and
  - the **materials** to which the conditions are applied are defined as the **experimental units**.
- ***Design of an experiment*** is planning of an experiment to obtain an appropriate measurement (data) and drawing inference out of the measurement (data) with respect to any problem under investigation.

Steps in experimentation (Little and Hills 1978)

- **Define the problem**: Biotic and abiotic stresses reducing productivity of maize
- Determine the **objectives** (Purposes): High yielding genotypes under multiple stress environments
- Select the **treatments**: Newly developed maize hybrids
- Select the **experimental material**: Experimental fields across multiple locations
- Select the **experimental design**: RCBD, alpha lattice, p-rep, etc
- Select the **experimental unit** and number of replications: 1 row plot, 2 row plot
- Ensure proper **randomization** and layout:
- Ensure proper means of **data collection**: Manual data collection, electronic data capture, precision phenotyping, etc
- Outline the **statistical analysis** before doing the experiment
- **Conduct the experiment**
- **Analyze** the data and interpret the results
- Prepare complete **reports**

The aim of the experiment - choice of treatments

- ***The aim of experiment refers to the experimental purpose or objective*** such as selecting high yielding and widely adapted hybrids for variety release
- The research question clearly defines the conditions that will be manipulated in the experiment, and these manipulations are commonly called treatments.

Research aim (hypothesis)	Treatments
Which genotype is the highest yielding in a given environment?	Treatments will be different varieties grown in a field trial, and the response will be measured as the yield of plants.
Is there any reduction in disease infection achieved with a new more resistant variety compared to a control variety?	Treatments will be the two varieties; a new variety compared to a control or standard variety, and the response will be level of disease infection.
Which chemicals are effective in controlling insect damage on plants?	Treatments will be levels of different insecticide applied to insects feeding on the plant species of interest, and the response will be measured as plant damage.

Choice of experimental material

- Experimental material defines the **sample space for the experiment**
- In agricultural experiments this can be:
  - **An area of land on a farm**, which is typically divided into smaller units, or plots of land.
  - **A set of pots** to be grown in glasshouse conditions, comprising the characteristics of pot size, growing medium and pot placement on benches in the glasshouse.
  - A collection of insects that can be housed in **cages or petri dishes**.
- An informed choice of experimental material must be made to ensure that conclusions from the study are valid for the wider population.

CGIAR CIMMYT

Experimental unit

- The experimental unit is the **smallest unit to which a treatment can be applied**, such that any two units receive different treatments.
  - A number of varieties will be exposed to a range of plant spacing treatments, and the resulting yield of plants will be measured. The study will be undertaken on an area of land on a farm, which is divided into smaller units, or plots of land.
    - **Treatment**: <u>variety by plant spacing</u>;
    - **experimental unit**: <u>plot</u>
  - A <u>new variety</u> will be compared to a <u>standard check variety</u>, and the response will be level of disease infection. The varieties will be grown in pots on benches in the glasshouse.
    - **Treatment**: <u>variety</u>;
    - **experimental unit**: <u>pot</u>

Experimental Unit here is the plot

Experimental Unit here is a tray containing 9 plants

CGIAR

Observational unit

- The **observational unit** is the **smallest unit on which a measurement is taken**.
- It may be the same as the experimental unit, or on a finer scale than the experimental unit.
- Let's again consider the examples given above.
  - Variety by plant spacing is the treatment and plot is the experimental unit. **If yield is measured for the whole plot**, then <u>plot is both the experimental and the observational unit.</u>
  - Variety is the treatment and pot is the experimental unit. **If measurements are taken on each plant**, then <u>plant is the observational unit, but pot is the experimental unit.</u>

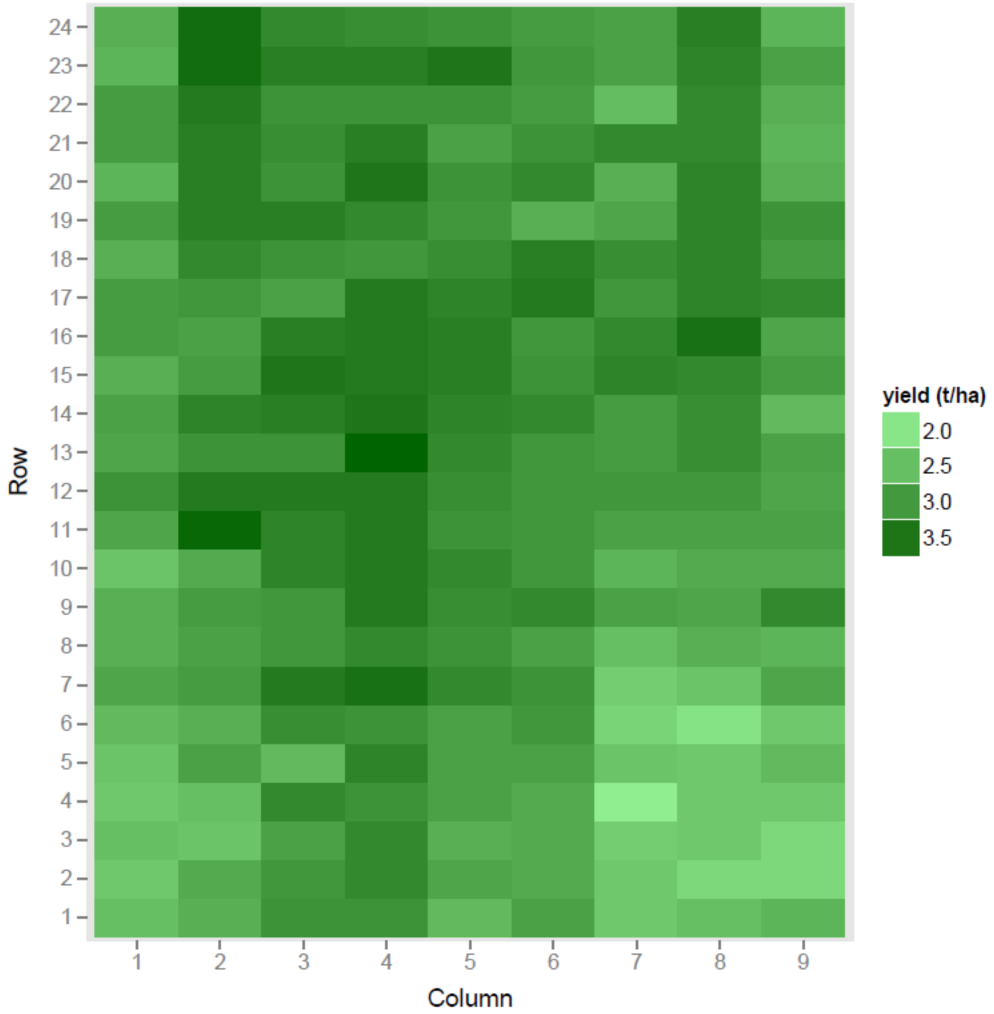
Observation unit is plot for GY

Observation Unit is pot if measurement is taken from each pot

CGIAR

The experimental design process

- Experimental design is simply the **process of allocating treatments to experimental units**.
- **The very nature of non-uniform experimental material and experimental methods dictates the need for established design principles.**
- Sound and robust experimental design is based on a knowledge of the **aims of the experiment**, the **treatments under study**, and an **understanding of the nature of the experimental material** to which the treatments will be applied.
- The experimental units may exhibit **variability due** to several sources
  - inherent variability in the material,
  - variation induced in the units through experimental processes and
  - measurement errors.

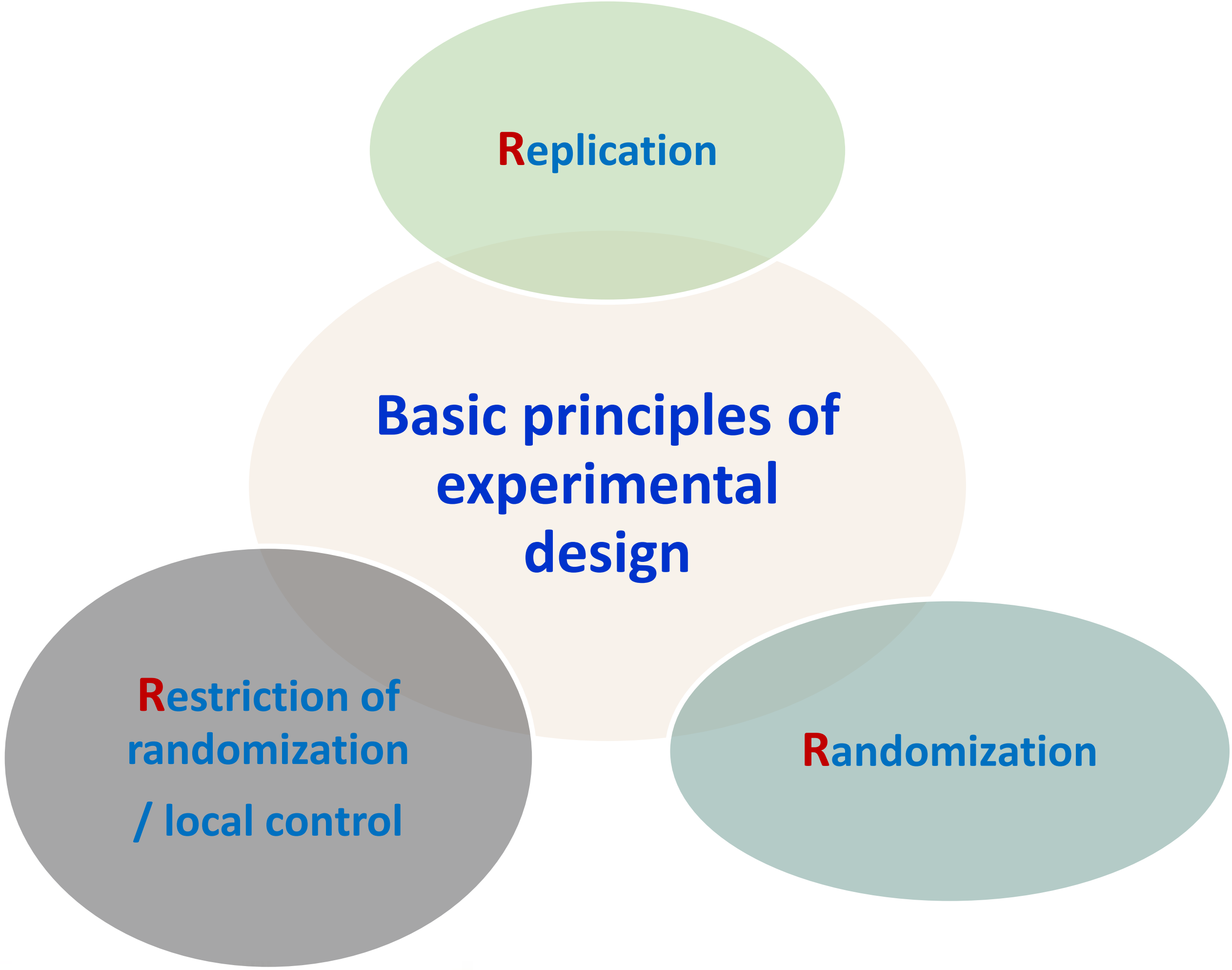


Grain yield of one variety grown in the whole field

Source: Alison Kelly Valeria Paccapelo, 2017

Principles of experimentation – 3R

- Almost all experiments involve **<u>three basic principles</u>**, *viz*., ***randomization***, ***replication*** and ***restriction of randomization/local control/blocking***.
- These three principles are **complementary to each other** in trying **to increase the accuracy** of the experiment and to **provide a valid test of significance**, retaining at the same time the distinctive features of their roles in any experiment.

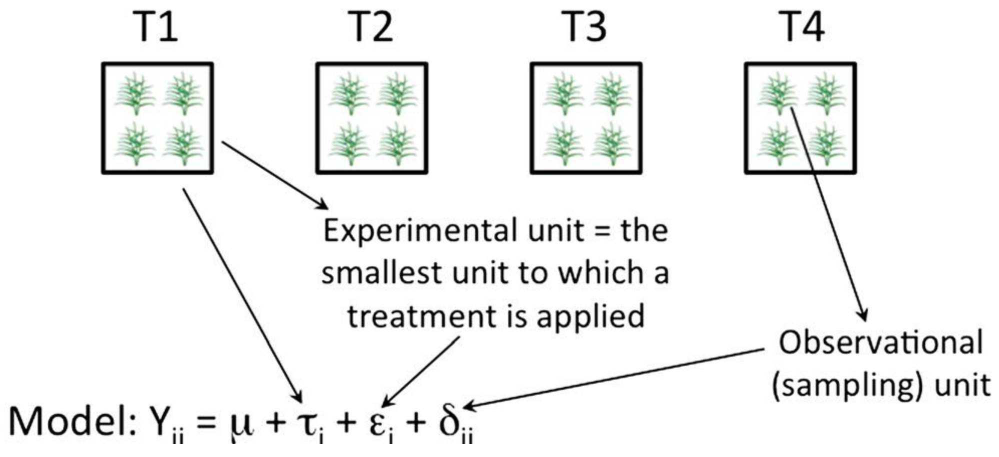


CGIAR CIMMYT

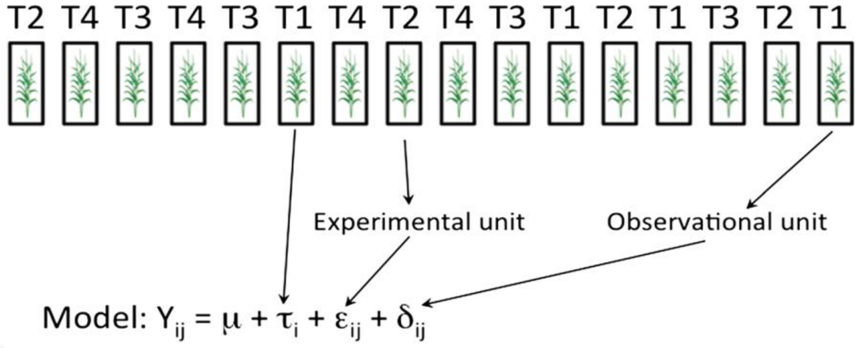
Replication

- A *replicate* is **a copy of a treatment**, such that the number of replicates of a treatment is the number of experimental units to which a treatment is applied.
- Replication of the experimental units ensures two important outcomes surrounding the experimental results.
  - **Firstly**, it **increases the accuracy** of the estimate of treatment effects, as each treatment effect is averaged across more than one experimental unit.
  - **Secondly**, the variation between experimental units containing the same treatment provides a **measure of background variability** of the experimental units

Experiment 1: Each T1, T2, T3, T4 appear only once



Experiment 2: T1, T2, T3, T4 appeared four times



CIMMYT

Randomization

- Randomization is the **process of allocating treatments to experimental units**.
- Randomization ensures that **we avoid biased or unrepresentative results** from the experiment.
- It is the insurance policy against the underlying variation in the experimental units.
- Bailey [2008] describes four types of bias, each illustrated here with a plant improvement example:
  - ***Systematic****:* allocating the varieties 1, 2, …, 20 to plots 1, 2, …, 20, i.e., variety 1 to plot 1, variety 2 to plot 2 etc. in the first replicate for all trials in a multi-environment trial series.
  - ***Selection****:* compositing the grain samples from the varieties with lower plot yields but not those with higher plot yields.
  - ***Accidental****:* measuring a grain quality trait on varieties which reach maturity before others.
  - ***Cheating****:* allocating an irrigation treatment to a lower lying part of the field than a non-irrigation treatment.

Restriction of randomization/local control

- **Experimental units** (e.g., field plots) **vary** from one another prior to the application of treatments e.g., <u>fertility and moisture gradients in the field</u>; lighting and air conditioning in glasshouse etc.
- **If this variability is ignored** in the design (or analysis) then the measurement **error** (residual variation) can be **inflated**
- ***Blocking* the experimental units into groups** that are considered to be **homogeneous** attempts to control known (or anticipated) local variation
- **Complete blocks contain an experimental unit for each treatment,** but incomplete blocks do not.

Moisture gradient in field experiment, Western Kenya, 2023

Principles of experimentation in practice

- Imagine that we have a field with a nitrogen gradient
- A breeder wants to evaluate 12 maize varieties

Less nitrogen → More nitrogen

A	D	G	J
A	D	G	J
A	D	G	J
A	D	G	J
B	E	H	K
B	E	H	K
B	E	H	K
B	E	H	K
C	F	I	L
C	F	I	L
C	F	I	L
C	F	I	L

What is the problem with this experimental design?

- Varieties A, B and C will experience worse growing conditions than J, K and L
- We won't know if their yield is lower because A, B and C are worse varieties, or simply because they had less nitrogen

Principles of experimentation in practice

- Randomly re-allocated varieties to experimental units:

Less nitrogen → More nitrogen

Completely randomized design

A	E	H	K
B	F	I	L
C	E	J	D
A	G	E	L
C	G	J	G
B	F	I	J
C	A	F	I
D	H	H	K
A	C	K	H
E	B	G	L
B	F	K	L
D	D	I	J

Slightly better, but A, B and C are still more often in low-nitrogen plots

What can we do?

Restrict the randomization > blocking

Principles of experimentation in practice

Solution:

Group experimental units into homogeneous groups **=Blocks**

Randomize varieties within groups

Blocking reduces the experimental error (σ^2)

> «Blocking can be thought as an insurance policy against disturbances that may or may not arise during the course of the experiment»

(Casley et al. 2015)

Less nitrogen → More nitrogen

Block 1	Block 2	Block 3	Block 4
A	C	E	H
F	K	B	C
C	A	F	D
I	F	C	B
B	D	J	E
E	J	A	I
J	E	H	L
G	G	K	A
K	L	G	K
D	H	D	J
L	B	I	F
H	I	L	G

Common Experimental Designs for maize breeding trials

- The Completely Randomized Design (CRD)
- Randomized Complete Block Design (RCBD)
- Alpha-Lattice Designs
- Augmented Designs
- Partially Replicated Designs

The Completely Randomized Design (CRD)

- The simplest design, suitable for a set of **homogeneous experimental units**
- Treatments **randomly** assigned to the units (***completely*** at random)
- **Problem:** in biological experiments, experimental units are almost never homogeneous

1 D	2 A	3 B	4 C	5 D
6 A	7 D	8 A	9 C	10 B
11 A	12 C	13 A	14 D	15 C
16 B	17 D	18 C	19 B	20 B

Welham et al 2014 (4.1A, p.72)

Randomized Complete Block Design (RCBD)

- All experimental units (e.g., field plots) within a block are considered homogeneous, i.e., similar in all respects that affect plant growth;
- Each block contains a *complete* set of treatments;
- Within a block the treatments are randomly allocated to the experimental units.

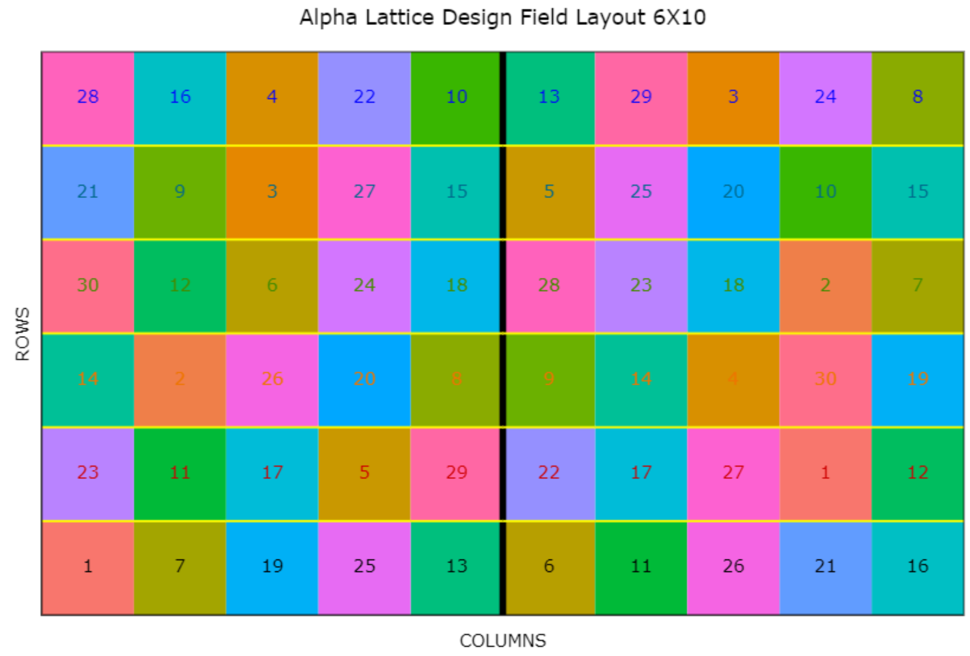
Less nitrogen Block 1	Block 2	Block 3	**Less nitrogen** Block 4
A	C	E	H
F	K	B	C
C	A	F	D
I	F	C	B
B	D	J	E
E	J	A	I
J	E	H	L
G	G	K	A
K	L	G	K
D	H	D	J
L	B	I	F
H	I	L	G

Treatments: 12 (A to L); 4 replication

CGIAR

Alpha-Lattice Designs

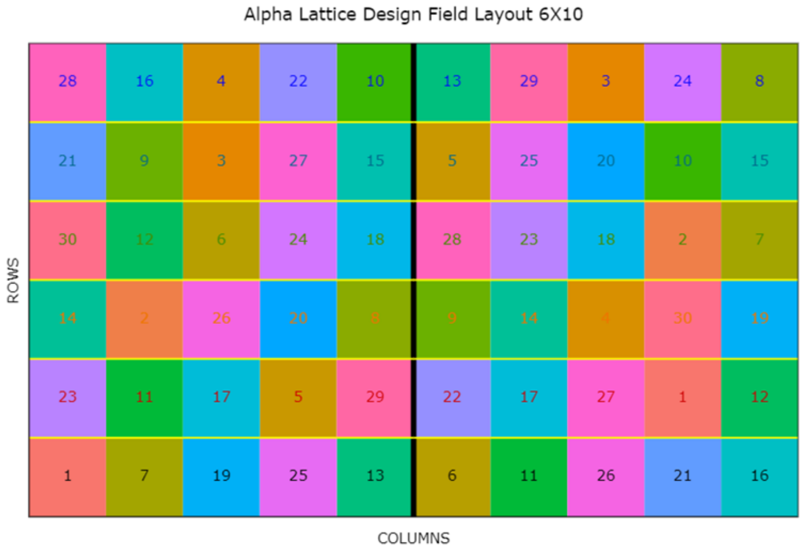
- This is one of the **most flexible experimental designs** in Plant Breeding (Patterson & Williams 1976, Patterson et al. 1978).
- It is applicable when the number of treatments (cultivars/genotypes) is **large**.
- The following are the characteristics of alpha lattice design:
  - Not all the treatments appear in one incomplete block
  - Randomly allocate treatments to plots within each incomplete block

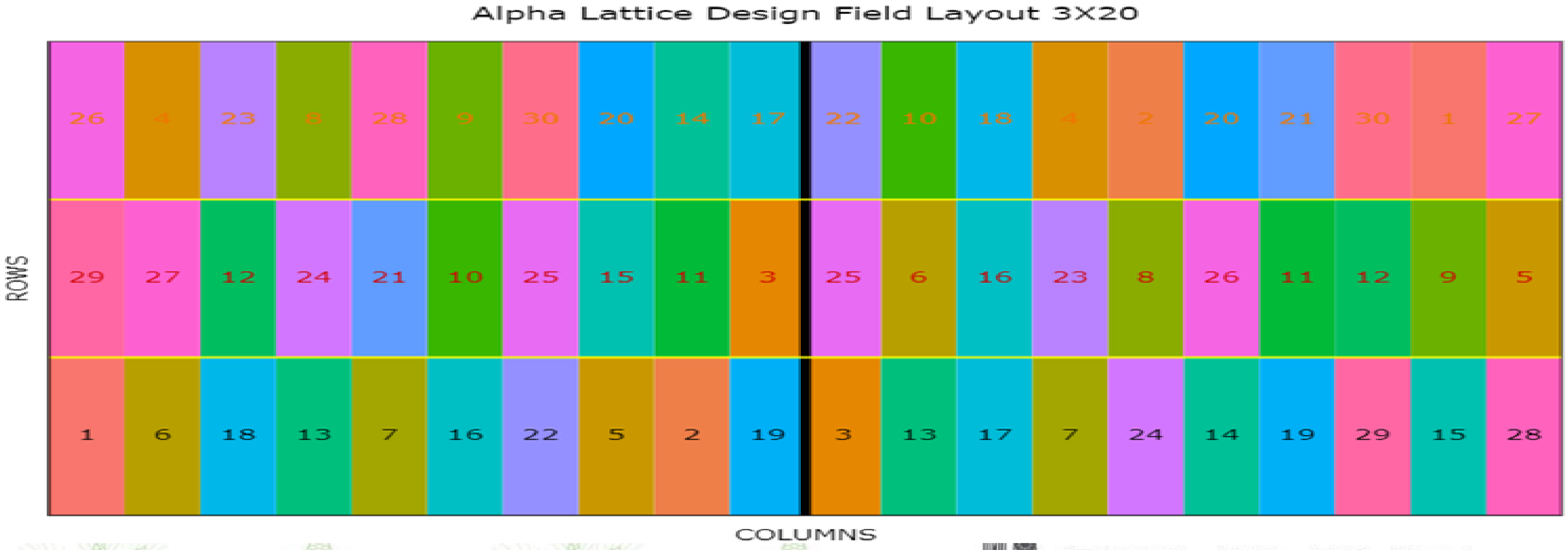


An alpha-lattice design for *v* = 30 varieties and *r* = 2 replicates, *s* = 6 blocks within each replicate and *k* = 5 plots within a block arranged as 6 rows by 10 columns,. Replicates are delineated by the vertical bold line and blocks by the yellow line lines

Alpha lattice design – example

ID	REP	IBLOCK	UNIT	ENTRY
1	1	1	1	1
2	1	1	2	7
3	1	1	3	19
4	1	1	4	25
5	1	1	5	13
6	1	2	1	23
7	1	2	2	11
8	1	2	3	17
9	1	2	4	5
10	1	2	5	29
11	1	3	1	14
12	1	3	2	2
13	1	3	3	26
14	1	3	4	20
15	1	3	5	8





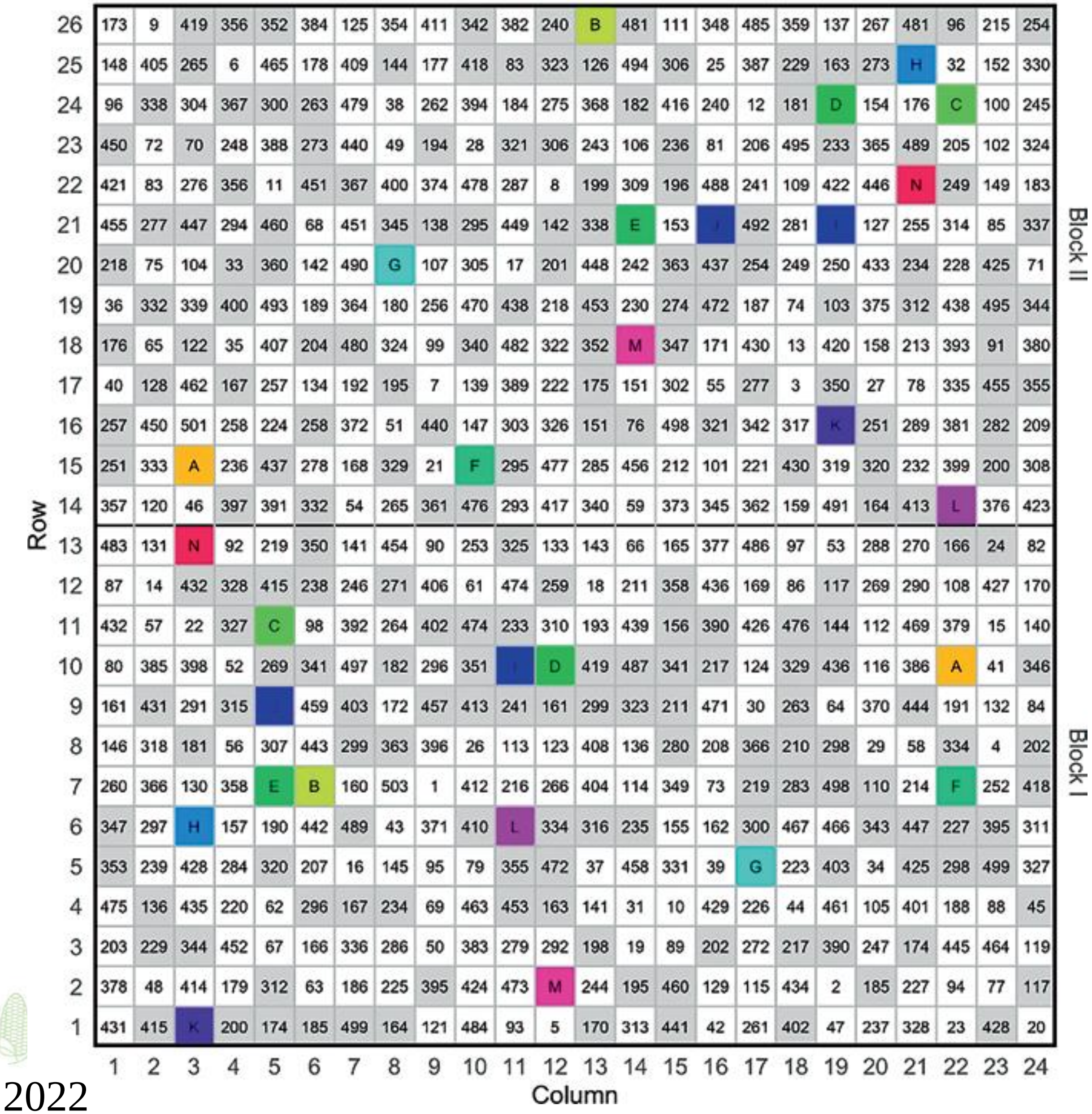
Augmented Designs

- Augmented designs are widely used in the design of **early-stage variety trials** which have **large treatment numbers**
- Augmented designs contain a combination of **replicated** and **un-replicated treatments**
- The **replicated treatments accounts** for spatial
- The systematic repetition of the **replicated treatments enables estimation of the block effects and residual (error) variance** resulting in more precise estimates of the treatment comparisons of interest.

Row \ Column	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
8	5	112	49	119	109	6	26	28	48	59	3	70	47	4	23	25
7	17	82	20	95	58	22	16	C1	92	88	117	12	107	105	15	100
6	38	C2	108	93	56	39	114	62	106	64	2	52	89	C1	110	31
5	103	19	72	86	18	33	101	75	87	32	120	60	71	40	C2	11
4	98	7	54	66	65	34	51	104	35	81	C2	96	14	37	63	50
3	113	53	43	68	42	1	29	C1	116	C1	85	46	73	30	36	84
2	77	74	57	C2	90	97	115	118	76	80	83	24	13	78	91	99
1	45	94	9	102	10	111	79	69	8	27	61	67	55	44	41	21

Partially Replicated Designs

- Partially replicated (*p*-rep) designs are introduced as an **alternative to augmented** designs for **early-stage variety trials**.
- **Key principle** is to replace the replicated checks in an augmented grid design with test lines. This increases the response to selection due to an increased replication of the lines under selection.
- See the example for partially replicated design for v = 504 varieties in 624 plots, arranged in 26 rows by 24 columns. The bold horizontal line delineates the Blocks. Colors represent different check lines.
- The gray shaded plots are those allocated with 2 replicate varieties



Source: K. L. Mathews and J. Crossa, 2022

CGIAR

Sparse phenotyping

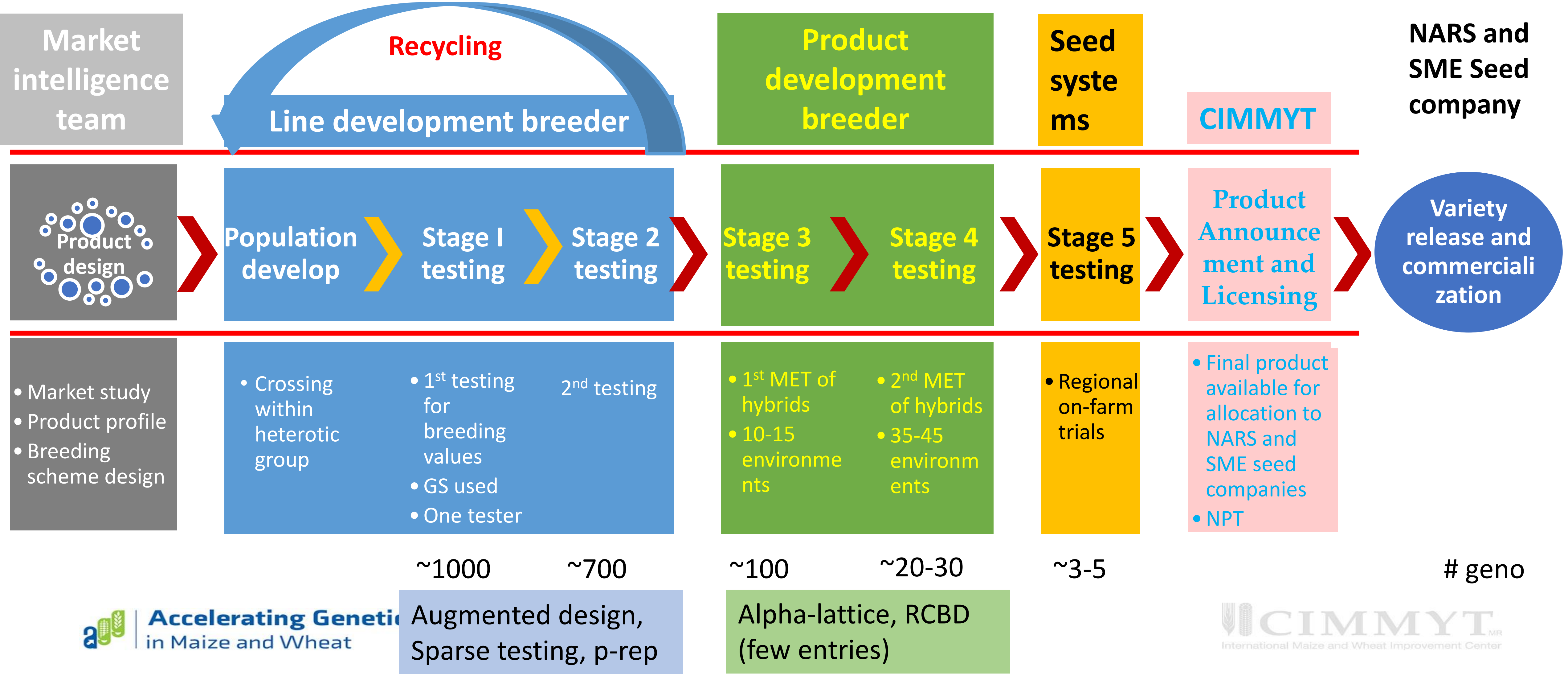
- **Sparse phenotyping** refers to early-stage multi-environment breeding trials in which not all genotypes of interest are grown in each environment
- **Using genomic-enabled prediction, the non-observed genotype-in-environment** combinations can be predicted.
- In sparse test crossing, the genomic relationship matrix (GRM) is used to for prediction

- Total number of Stage 1 entries = 928
- # Checks = 4
- # entries evaluated in all sites = 52
- # Entries to be evaluated in sparse design (928-52 = 876)
- # entries evaluated at one of the three site (876/3 = 292)

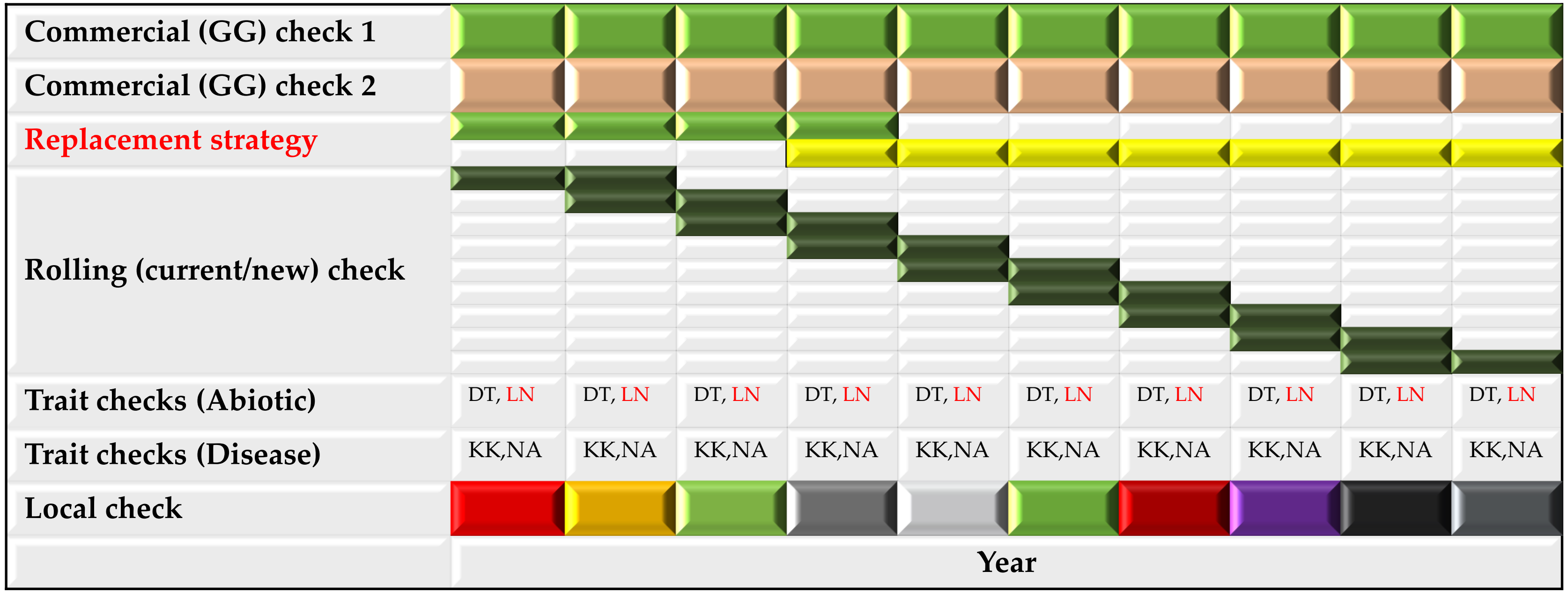
	Env1	Env2	Env3	Total
Full set (1)	52	52	52	
Set 2	292	---	---	
Set 3	---	292	---	
Set 4	---	---	292	
# checks	4	4	4	
tal # entreis /si	348	348	348	
Entries	348	348	348	932
Reps	2	2	2	2
Locations	1	1	1	3
Rows	696	696	696	2,088
Cost ($)	3480	3480	3480	**10,440**

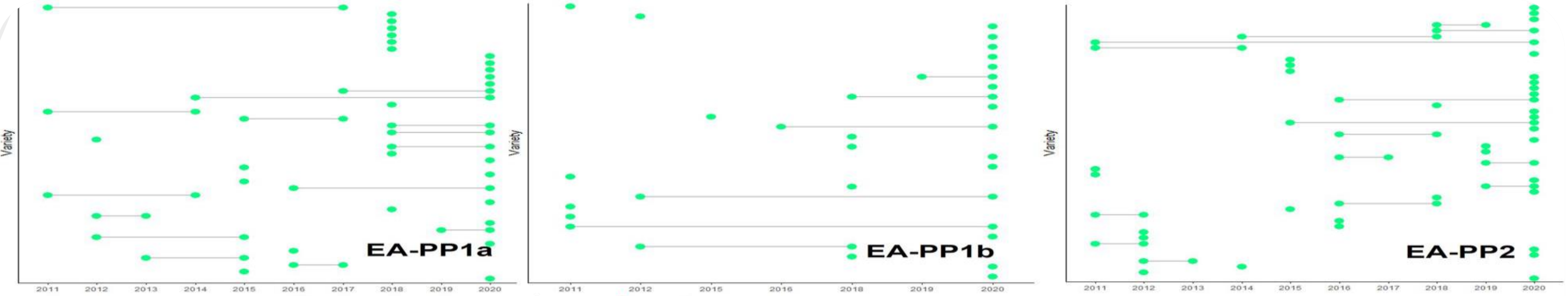
CGIAR CIMMYT

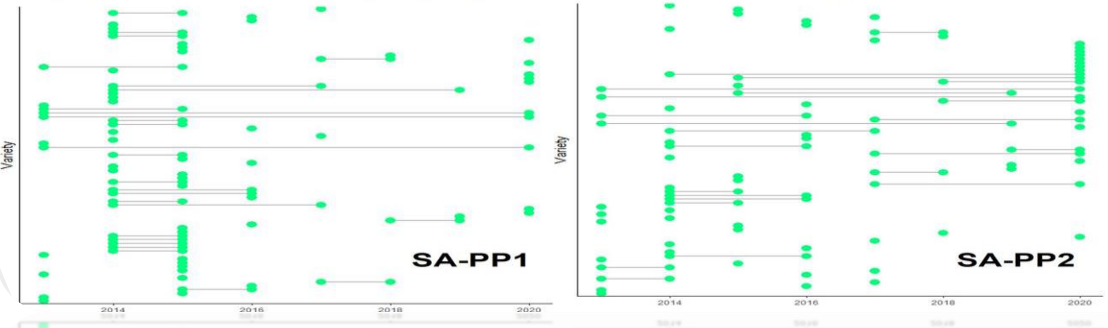
General Breeding Scheme in CIMMYT maize



Check Strategy at CIMMYT







Source: **Prasanna et al. (2022)** Genetic trends in CIMMYT's tropical maize breeding pipelines.

- **Rolling checks** (current) are recently announced hybrids after Stage 5 testing.
- **Abotic stress tolerant checks**: used only under managed stress environments (drought, low N).
- **Disease trait checks**
- **Cross-cutting check:** Maturity checks
- **Check connectivity** across years ensured to facilitate analysis of genetic trends in CIMMYT maize breeding pipelines.

Summary

- Experimental units are the physical entity which can be assigned, at random, to a treatment.
- We need replication to quantify the natural variation between **experimental units** (higher replication leads to increased precision of estimated effects).
- If experimental units are fully homogeneous (almost never the case), treatments can be randomly assigned to experimental units.
- We use blocks to restrict the randomization, forming more homogeneous groups of experimental units.

Thank you for your interest!