Genomic Selection:
Key Models, Methods, and Perspectives

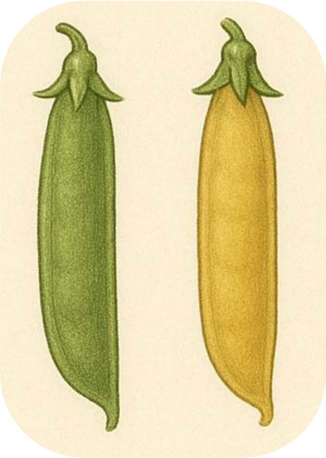
Alexandre Hild Aono
@AonoAlexandre
alexandre.aono@gmail.com

Swedish University of Agricultural Sciences (SLU)
Department of Plant Breeding

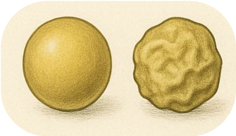
Phenotypes

Pea Traits

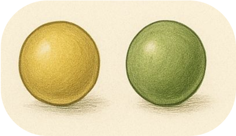
Pod Color



Seed Shape



Seed Color



Flower Color

Mendelian Traits Determined by a single or very few genes

Number of Pods

Yield

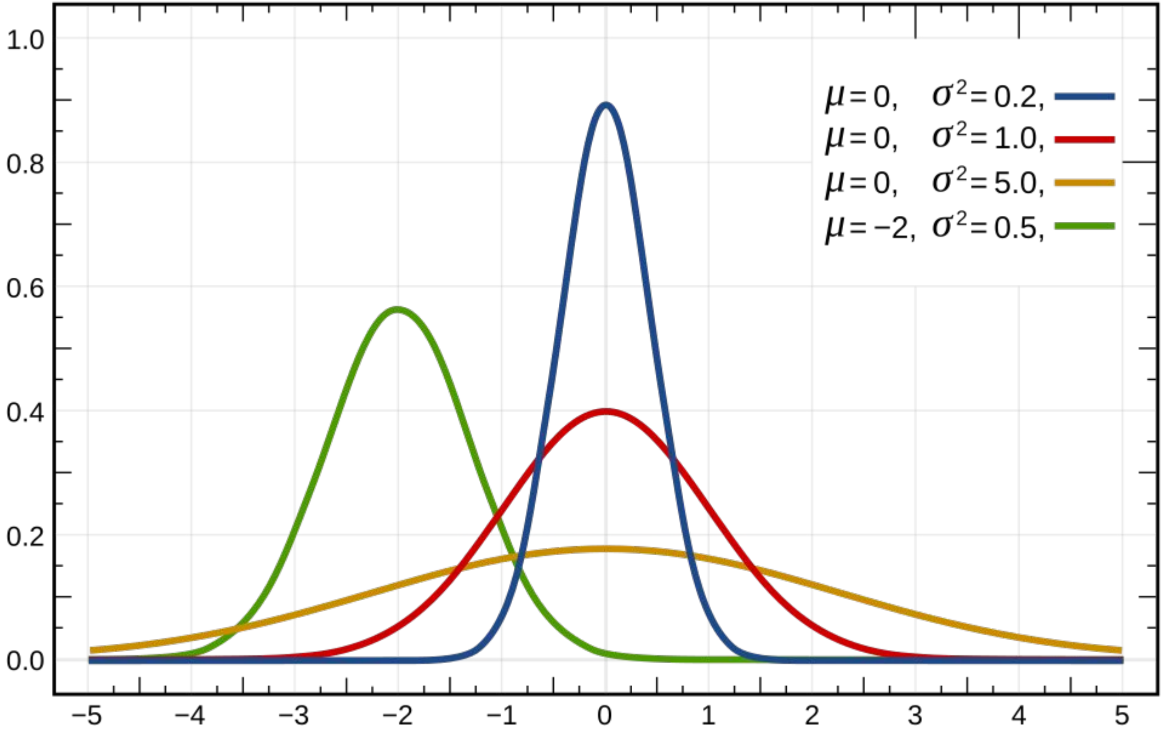
Plant Height

Quantitative Traits Determined by a large number of genes and non-genetic factors

Normal Distribution

$$Y \sim N(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$



$$\mathrm{E}(Y) = \mu$$

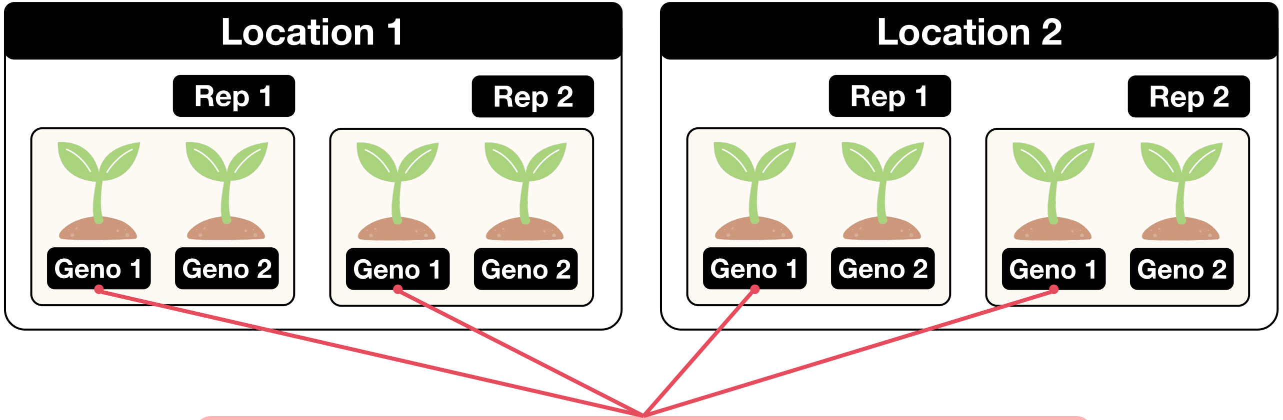
$$\mathrm{Var}(Y) = \sigma^2$$

Quantitative Traits

Directly transmitted through generations

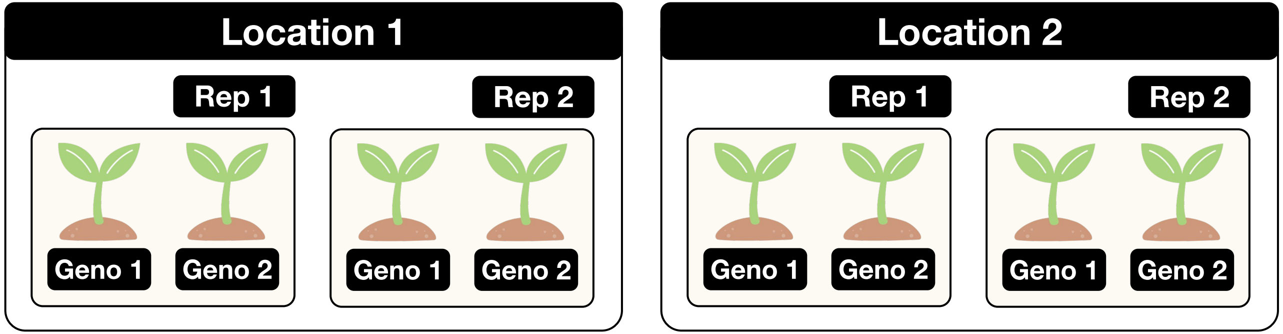
Additive Effects

`Phenotype = Fixed Effects + Genetics + Environment + Residuals`



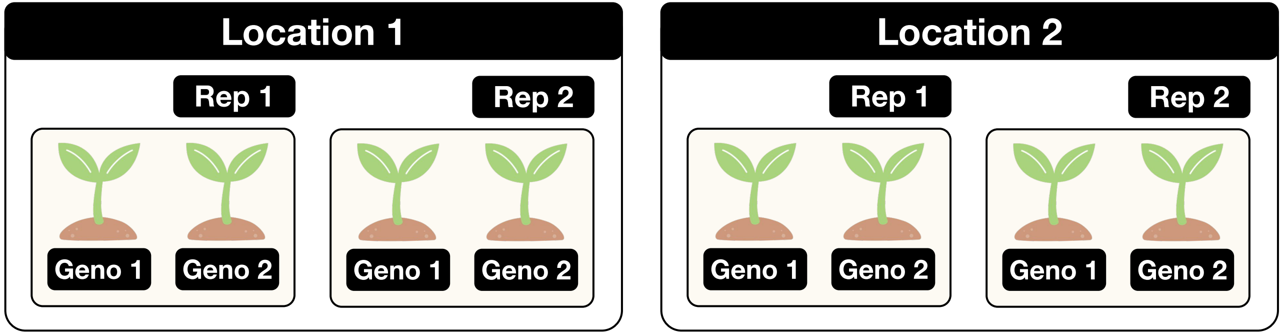
Will these plants have the same phenotype?

Quantitative Traits



$$
\begin{bmatrix} 65 \\ 64 \\ 46 \\ 45 \\ 94 \\ 94 \\ 76 \\ 74 \end{bmatrix}
=
\begin{bmatrix} 60 \\ 60 \\ 60 \\ 60 \\ 90 \\ 90 \\ 90 \\ 90 \end{bmatrix}
\begin{matrix} \text{Geno 1} \\ \\ \\ \\ \\ \\ \\ \text{Geno 2} \end{matrix}
+
\begin{bmatrix} 0 \\ 0 \\ -20 \\ -20 \\ 0 \\ 0 \\ -20 \\ -20 \end{bmatrix}
\begin{matrix} \text{Loc 1} \\ \\ \text{Loc 2} \\ \\ \text{Loc 1} \\ \\ \text{Loc 2} \\ \\ \end{matrix}
+
\begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \\ 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}
\begin{matrix} \text{Rep 1 | Loc 1} \\ \text{Rep 2 | Loc 1} \\ \text{Rep 1 | Loc 2} \\ \text{Rep 2 | Loc 2} \\ \text{Rep 1 | Loc 1} \\ \text{Rep 2 | Loc 1} \\ \text{Rep 1 | Loc 2} \\ \text{Rep 2 | Loc 2} \end{matrix}
+
\begin{bmatrix} 3 \\ 1 \\ 2 \\ 4 \\ 2 \\ 1 \\ 2 \\ 3 \end{bmatrix}
$$

Quantitative Traits



$Y_{i,j,k}$

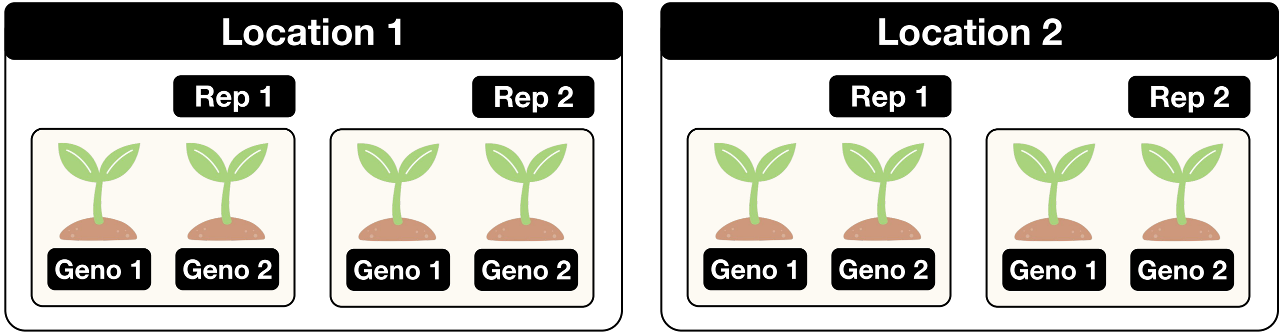
Genotype *i*

Location *j*

Repetition *k*

$$
\begin{bmatrix} Y_{1,1,1} \\ Y_{1,1,2} \\ Y_{1,2,1} \\ Y_{1,2,2} \\ Y_{2,1,1} \\ Y_{2,1,2} \\ Y_{2,2,1} \\ Y_{2,2,2} \end{bmatrix}
=
\begin{bmatrix} g_1 \\ g_1 \\ g_1 \\ g_1 \\ g_2 \\ g_2 \\ g_2 \\ g_2 \end{bmatrix}
+
\begin{bmatrix} e_1 \\ e_1 \\ e_2 \\ e_2 \\ e_1 \\ e_1 \\ e_2 \\ e_2 \end{bmatrix}
+
\begin{bmatrix} e_1:r_1 \\ e_1:r_2 \\ e_2:r_1 \\ e_2:r_2 \\ e_1:r_1 \\ e_1:r_2 \\ e_2:r_1 \\ e_2:r_2 \end{bmatrix}
+
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix}
$$

Quantitative Traits



$$Y = X_1 g + X_2 e + X_3 e{:}r + \square$$

Quantitative Traits

$$
\begin{bmatrix} Y_{1,1,1} \\ Y_{1,1,2} \\ Y_{1,2,1} \\ Y_{1,2,2} \\ Y_{2,1,1} \\ Y_{2,1,2} \\ Y_{2,2,1} \\ Y_{2,2,2} \end{bmatrix}
=
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 : r_1 \\ e_1 : r_2 \\ e_2 : r_1 \\ e_2 : r_2 \end{bmatrix}
+
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix}
$$

Linear Model

Coefficients

$$
\begin{bmatrix} 65 \\ 64 \\ 46 \\ 45 \\ 94 \\ 94 \\ 76 \\ 74 \end{bmatrix}
=
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 : r_1 \\ e_1 : r_2 \\ e_2 : r_1 \\ e_2 : r_2 \end{bmatrix}
+
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix}
$$

Linear Model

$$
\begin{bmatrix} 65 \\ 64 \\ 46 \\ 45 \\ 94 \\ 94 \\ 76 \\ 74 \end{bmatrix}
-
\left[
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} \hat{e}_1 \\ \hat{e}_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} \hat{e}_1 : r_1 \\ \hat{e}_1 : r_2 \\ \hat{e}_2 : r_1 \\ \hat{e}_2 : r_2 \end{bmatrix}
\right]
=
\begin{bmatrix} \hat{\square}_{1,1,1} \\ \hat{\square}_{1,1,2} \\ \hat{\square}_{1,2,1} \\ \hat{\square}_{1,2,2} \\ \hat{\square}_{2,1,1} \\ \hat{\square}_{2,1,2} \\ \hat{\square}_{2,2,1} \\ \hat{\square}_{2,2,2} \end{bmatrix}
$$

Linear Model - Assumption

$$
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix} \sim N \left[\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2 \end{bmatrix} \right]
$$

Independent and Identically Distributed

Correlation, Variance, and Covariance

```
X = [1    2    1    3]
Y = [4    3    2    5]
```

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Correlation, Variance, and Covariance

```
X = [1      2      1      3]
Y = [4      3      2      5]
```

$$\mathrm{cov}(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_x)(y_i - \mu_y)$$

$$\mathrm{cor}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}}$$

Covariance Matrix

$$\begin{bmatrix} \sigma^2_{X} & \sigma_{X,Y} \\ \sigma_{X,Y} & \sigma^2_{Y} \end{bmatrix}$$

Linear Mixed Effects Models

Linear Mixed Effects Model

$$
\begin{bmatrix} Y_{1,1,1} \\ Y_{1,1,2} \\ Y_{1,2,1} \\ Y_{1,2,2} \\ Y_{2,1,1} \\ Y_{2,1,2} \\ Y_{2,2,1} \\ Y_{2,2,2} \end{bmatrix}
=
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1{:}r_1 \\ e_1{:}r_2 \\ e_2{:}r_1 \\ e_2{:}r_2 \end{bmatrix}
+
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix}
$$

Linear Mixed Effects Model

$$
\begin{bmatrix} Y_{1,1,1} \\ Y_{1,1,2} \\ Y_{1,2,1} \\ Y_{1,2,2} \\ Y_{2,1,1} \\ Y_{2,1,2} \\ Y_{2,2,1} \\ Y_{2,2,2} \end{bmatrix}
=
\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} e_1 \\ e_2 \\ e_1:r_1 \\ e_1:r_2 \\ e_2:r_1 \\ e_2:r_2 \end{bmatrix}
+
\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} g_1 \\ g_2 \end{bmatrix}
+
\begin{bmatrix} \square_{1,1,1} \\ \square_{1,1,2} \\ \square_{1,2,1} \\ \square_{1,2,2} \\ \square_{2,1,1} \\ \square_{2,1,2} \\ \square_{2,2,1} \\ \square_{2,2,2} \end{bmatrix}
$$

Linear Mixed Effects Model

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} \square_1 \\ \square_2 \\ \square_3 \\ \square_4 \\ \square_5 \\ \square_6 \\ \square_7 \\ \square_8 \end{bmatrix}$$

Phenotype = Fixed Effects + Random Effects + Residuals

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

Linear Mixed Effects Model

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} \square_1 \\ \square_2 \\ \square_3 \\ \square_4 \\ \square_5 \\ \square_6 \\ \square_7 \\ \square_8 \end{bmatrix} \sim N \left[\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sigma_{\square}^2 \right]$$

R

$$\square \sim N(0, R\sigma_{\square}^2)$$

Linear Mixed Effects Model

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_u^{\ 2}\right]$$

$\boldsymbol{\Sigma}$

$$u \sim N(0, \Sigma\sigma_u^{\ 2})$$

Henderson Equations

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + \Sigma^{-1} \dfrac{\sigma_{\square}^{2}}{\sigma_u^{2}} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\square \sim N(0, R\sigma_{\square}^{2})$$

$$u \sim N(0, \Sigma\sigma_u^{2})$$

Henderson Equations

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + \Sigma^{-1} \frac{\sigma_{\square}{}^2}{\sigma_u{}^2} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\square \sim N(0, R\sigma_{\square}{}^2)$$

$$u \sim N(0, \Sigma\sigma_u{}^2)$$

Henderson Equations

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + \Sigma^{-1} \frac{\sigma_{\square}^2}{\sigma_u^2} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\square \sim N(0, R\sigma_{\square}^2)$$

$$u \sim N(0, \Sigma\sigma_u^2)$$

Example

Variance Components

Genomic Selection:

Key Models, Methods, and Perspectives

Henderson Equations

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \boldsymbol{\epsilon}$$

What about the distribution of y?

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + \Sigma^{-1}\frac{\sigma_\epsilon^2}{\sigma_u^2} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\epsilon \sim N(0, R\sigma_\epsilon^2)$$

$$u \sim N(0, \Sigma\sigma_u^2)$$

$$y \sim N(Xb,\; Z\,\Sigma\,\sigma_u^2\,Z^T + R\,\sigma_\epsilon^2)$$

Phenotypic Variance

$$y \sim N(Xb,\ Z\ \Sigma\ \sigma_g^2\ Z^T + R\ \sigma_{\square}^2)$$

Genetic Variance

Residual Variance

Broad Sense Heritability

$$\frac{\sigma_g^2}{\sigma_g^2 + \sigma_{\square}^2}$$

How well the trait will respond to selection

The Infinitesimal Model

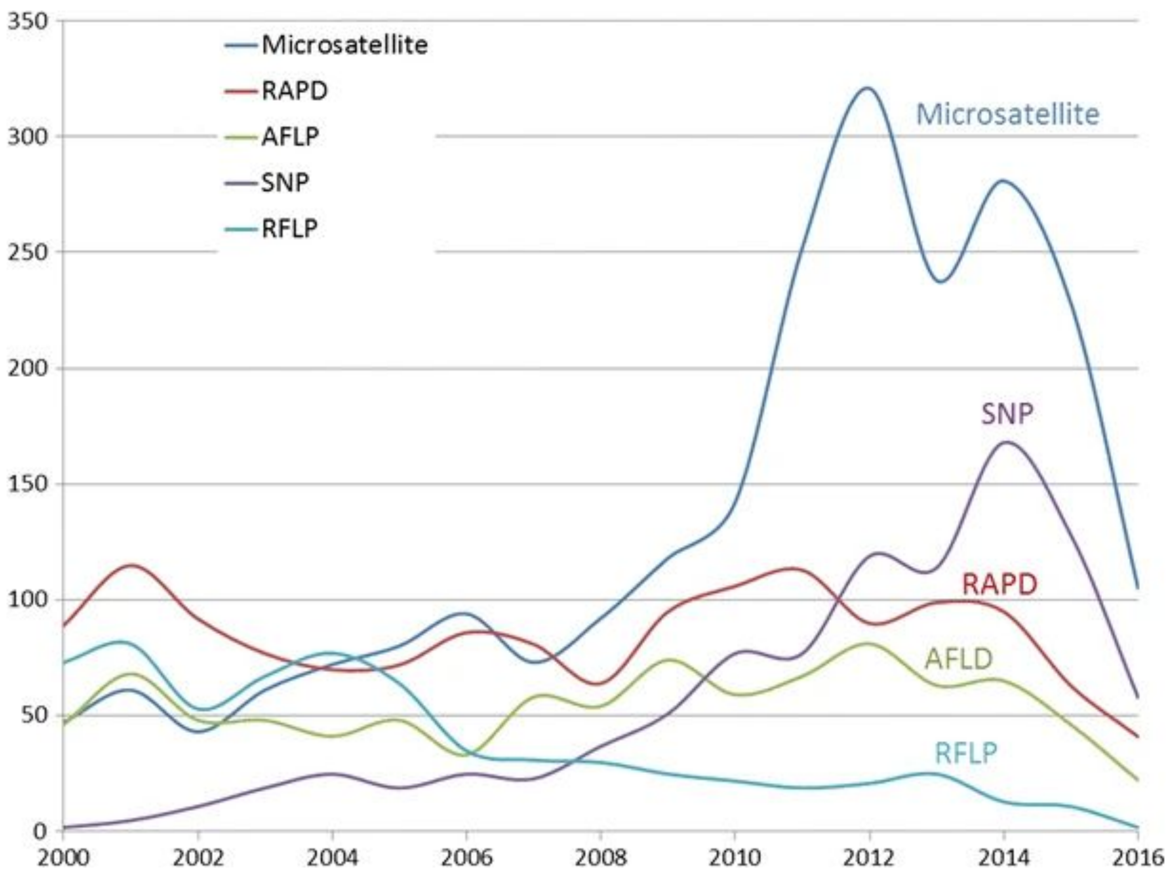
Infinitesimal Model (Fisher, 1918)

- Quantitative traits are polygenic: **very small genetic effects** over a **very large number of genes**

- Random sampling of alleles at each locus results in **continuous and normally distributed phenotypes** in the population

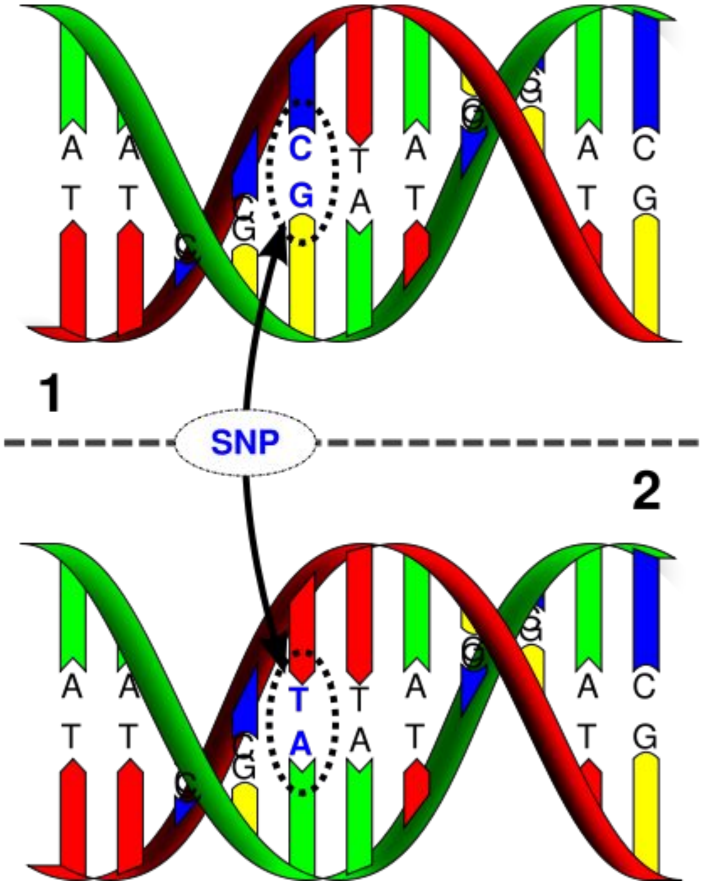
- Assumes that **phenotypic variation** due to genetics is **constant**

Molecular Markers



Garrido-Cardenas, Jose Antonio, Concepción Mesa-Valle, and Francisco Manzano-Agugliaro. "Trends in plant research using molecular markers." *Planta* 247 (2018): 543-557.

Single Nucleotide Polymorphisms



Geno	Allele 1	Allele 2	Code
G1	C	C	0
G2	C	G	1
G3	G	G	2

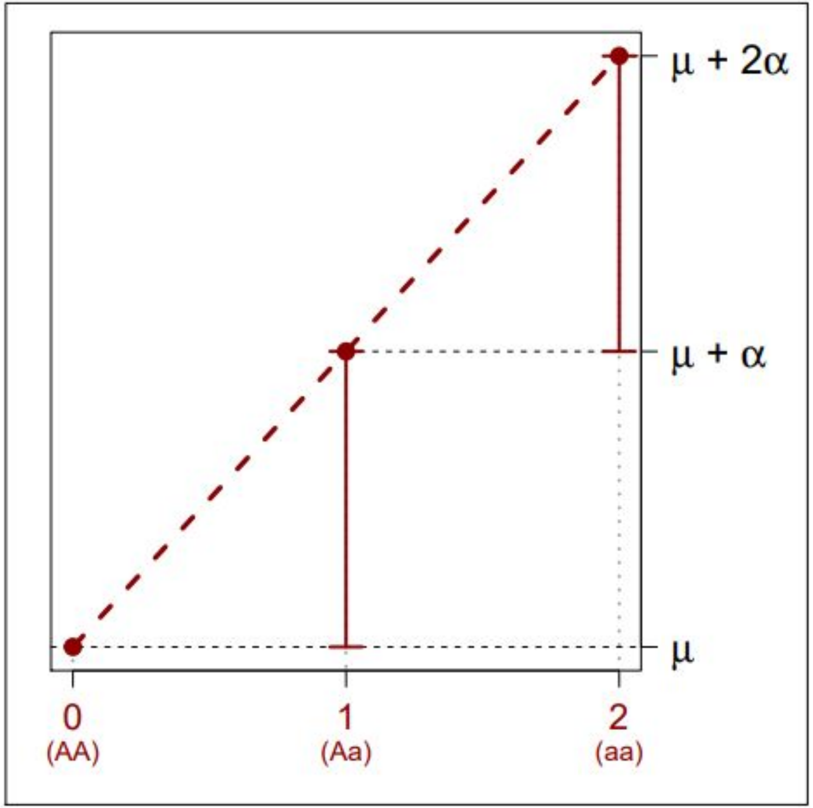
One Locus Model

Effect

$$y = \mu + M\alpha + \square$$

0 1 2

Marker Genotype



Infinitesimal Model

$$y_i = \mu + \sum_{j=1}^{m} M_{ij}\alpha_j + \epsilon_i$$

Infinitesimal Model

M_1 M_2 M_3 M_4

Ind 1

ACTGCTATGCGCGATATCGCGATCG Allele 1
ACTGCTATGCACGATATCGCGATCG Allele 2

$$Y_1 = \mu + M_{1,1}\boldsymbol{\alpha}_1 + M_{1,2}\boldsymbol{\alpha}_2 + M_{1,3}\boldsymbol{\alpha}_3 + M_{1,4}\boldsymbol{\alpha}_4 + E_1$$

Ind 2

ACTGTTATGCACGATATCGCGATCG Allele 1
CCTGTTATGCACGATAACGCGATCG Allele 2

$$Y_2 = \mu + M_{2,1}\boldsymbol{\alpha}_1 + M_{2,2}\boldsymbol{\alpha}_2 + M_{2,3}\boldsymbol{\alpha}_3 + M_{2,4}\boldsymbol{\alpha}_4 + E_2$$

Ind 3

CCTGCTATGCGCGATAACGCGATCG Allele 1
CCTGTTATGCGCGATAACGCGATCG Allele 2

$$Y_3 = \mu + M_{3,1}\boldsymbol{\alpha}_1 + M_{3,2}\boldsymbol{\alpha}_2 + M_{3,3}\boldsymbol{\alpha}_3 + M_{3,4}\boldsymbol{\alpha}_4 + E_3$$

$$Y_1 = G_1 + E_1$$

$$Y_2 = G_2 + E_2$$

$$Y_3 = G_3 + E_3$$

Infinitesimal Model

Only Additive Effects

$$y_i = \mu + \sum_{j=1}^{m} M_{ij}\alpha_j + \epsilon_i$$

The individual additive effect of each locus is very small
Epistasis and dominance effects are disregarded

Infinitesimal Model

$$m \to \infty$$

$$y_i = \mu + \sum_{j=1}^{m} M_{ij}\alpha_j + \epsilon_i$$

The number of loci affecting the trait is so large that approximates infinity, assuming that many additive effects can absorb epistasis and dominance

Infinitesimal Model

$$y_i = \mu + \sum_{j=1}^{m} M_{ij}\alpha_j + \epsilon_i$$

Genetic (Breeding) Values

Infinitesimal Model

$$y_i = \mu + \sum_{j=1}^{m} M_{ij}\alpha_j + \epsilon_i$$

Genetic (Breeding) Values

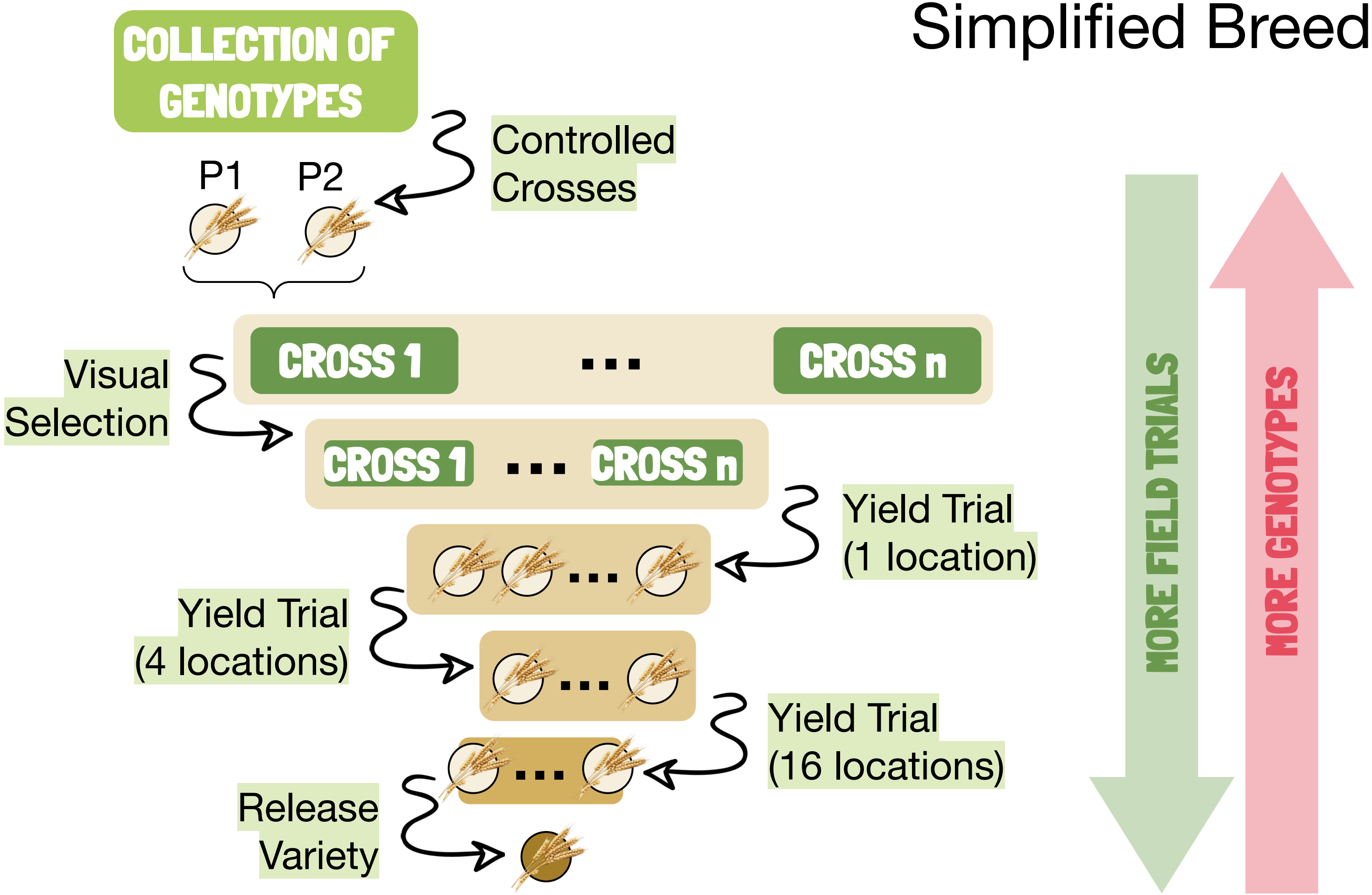
$$y_i = \mu + g_i + \epsilon_i$$

Marker-Assisted and Genomic Selection

Genomic Selection:

Key Models, Methods, and Perspectives

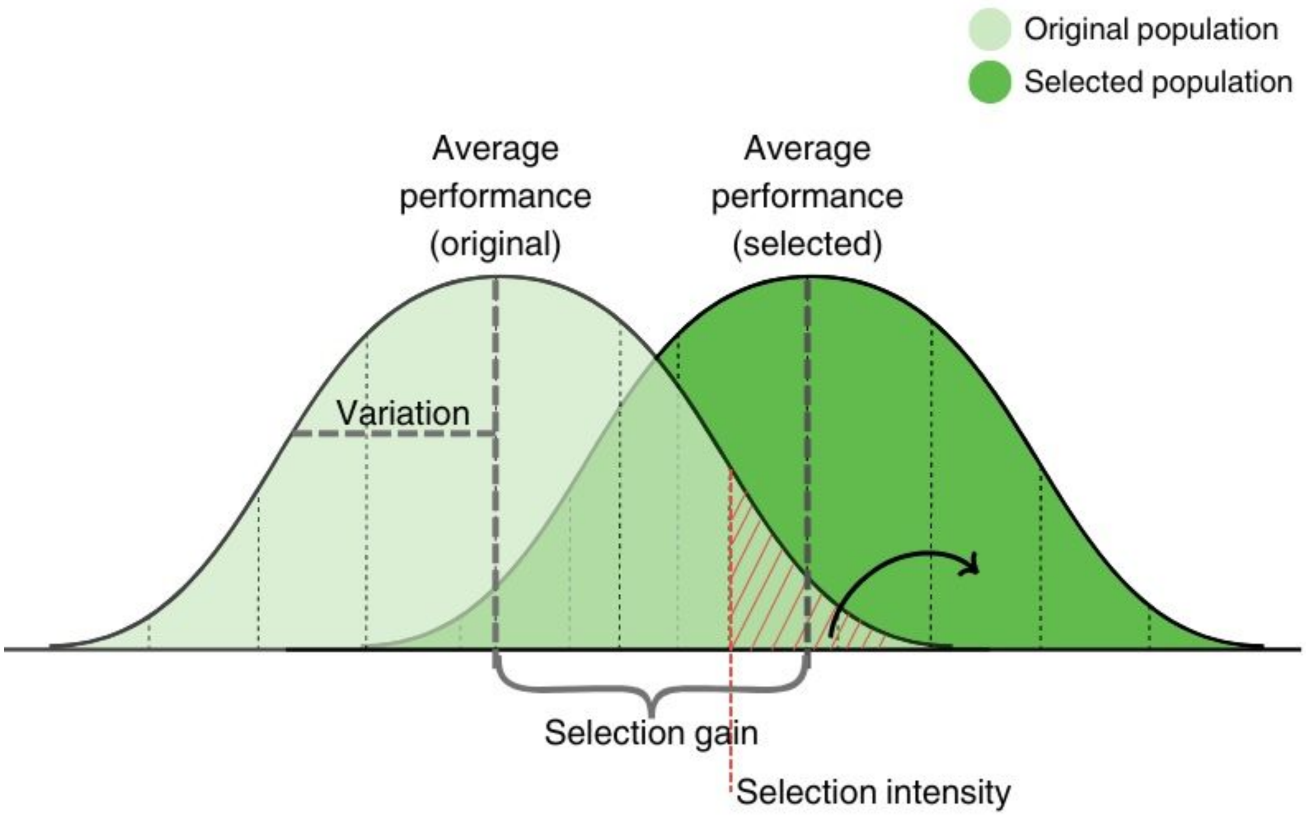
Simplified Breeding Program



How to Maximize the Genetic Gains?

- Increasing the **intensity of selection**
- Minimizing the **environmental effect** (replicate and randomize)
- Increasing the **genetic variability**
- Using **more accurate or less time-consuming** breeding schemes

Genetic Gain Obtained



Estimated Genetic Gain

$$\Delta G = \frac{i \times r \times \sigma_A}{L}$$

Genetic Gain

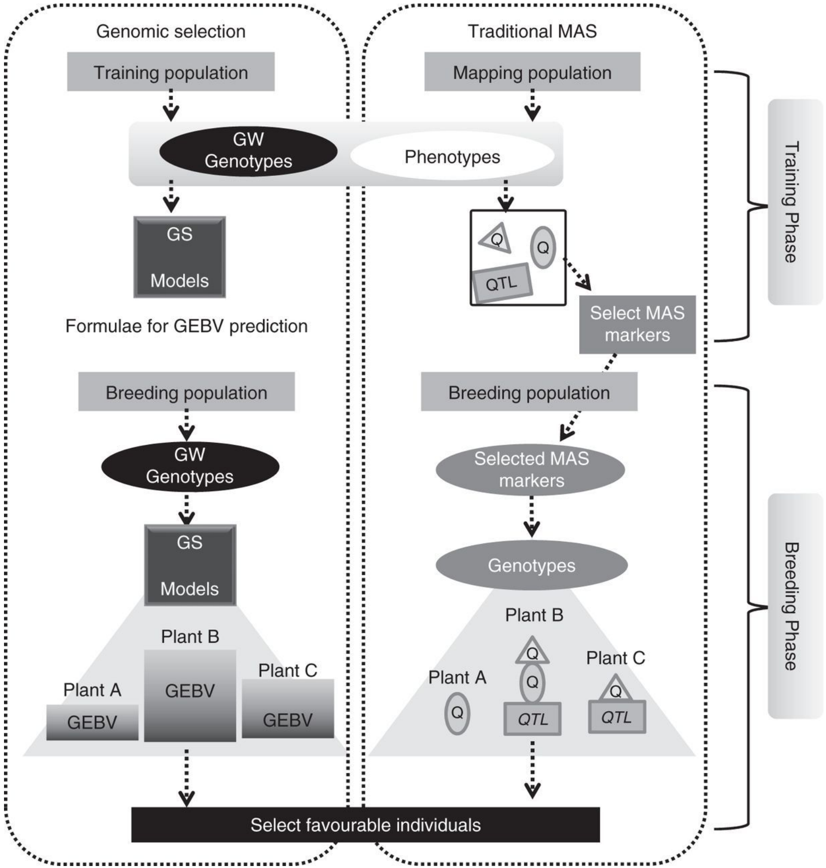
Selection Intensity

Accuracy of Selection

Genetic Variance

Generation Interval

Molecular Markers



Nakaya, Akihiro, and Sachiko N. Isobe. "Will genomic selection be a practical method for plant breeding?." *Annals of Botany* 110.6 (2012): 1303-1316.

Marker Assisted Selection

QTL Mapping

Low allelic diversity
Biparental population needs to be created
Lower resolution (few recombination events)
Low resolution, **large QTLs**

GWAS

Sometimes results are **not replicated** across populations
Results need to be **validated**
Size of population and markers

Meuwissen et al. (2001)

Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,* B. J. Hayes[†] and M. E. Goddard[†,‡]

**Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia*

Manuscript received August 17, 2000
Accepted for publication January 17, 2001

Past Perspectives

Annals of Botany **110**: 1303–1316, 2012
doi:10.1093/aob/mcs109, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

REVIEW: PART OF A HIGHLIGHT ON BREEDING STRATEGIES FOR FORAGE AND GRASS IMPROVEMENT

Will genomic selection be a practical method for plant breeding?

Akihiro Nakaya[1] and Sachiko N. Isobe[2,*]

[1]*Center for Transdisciplinary Research, Niigata University, 1-757 Asahimachi-dori, Chuo-ku, Niigata 951-8585, Japan and* [2]*Kazusa DNA Research Institute, 2-6-7 Kazusa Kamatari, Kisarazu, Chiba 292-0818, Japan*
**For correspondence. E-mail: sisobe@kazusa.or.jp*

Past Perspectives

Trends in Biotechnology

OPINION · Volume 31, Issue 9, P497-504, September 2013

Opinion

Cell PRESS

Does genomic selection have a future in plant breeding?

Elisabeth Jonas and Dirk-Jan de Koning

Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, Box 7023, 75007 Uppsala, Sweden

Past Perspectives

Trends in Plant Science

FEATURE REVIEW · Volume 22, Issue 11, P961-975, November 2017

Feature Review

Genomic Selection in Plant Breeding: Methods, Models, and Perspectives

José Crossa,[1,*] Paulino Pérez-Rodríguez,[2] Jaime Cuevas,[3]
Osval Montesinos-López,[4] Diego Jarquín,[5]
Gustavo de los Campos,[6] Juan Burgueño,[1]
Juan M. Camacho-González,[2] Sergio Pérez-Elizalde,[2]
Yoseph Beyene,[1] Susanne Dreisigacker,[1] Ravi Singh,[1]
Xuecai Zhang,[1] Manje Gowda,[1] Manish Roorkiwal,[7]
Jessica Rutkoski,[8] and Rajeev K. Varshney[7,*]

Strategy Evaluations

Plant Science 242 (2016) 23–36

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

ELSEVIER

Breeding schemes for the implementation of genomic selection in wheat (*Triticum* spp.)

CrossMark

Filippo M. Bassi [a], Alison R. Bentley [b], Gilles Charmet [c], Rodomiro Ortiz [d,*], Jose Crossa [e]

[a] *International Center for Agricultural Research in the Dry Areas (ICARDA), Av. Mohamed Alaoui, Rabat 10000, Morocco*
[b] *The John Bingham Laboratory, NIAB, Huntingdon Road, Cambridge CB3 0LE, United Kingdom*
[c] *Institut national de la Recherche Agronomique, INRA UMR1095 GDEC Clermont-Ferrand, 5 chemin de Beaulieu, 63039 Clermont-Ferrand, France*
[d] *Swedish University of Agricultural Sciences (SLU), Sundsvagen 14 Box 101, SE23453, Alnarp, Sweden*
[e] *Biometrics and Statistics Unit, International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico DF, Mexico*

Strategy Evaluations

Theoretical and Applied Genetics (2019) 132:669–686
https://doi.org/10.1007/s00122-018-3270-8

REVIEW ARTICLE

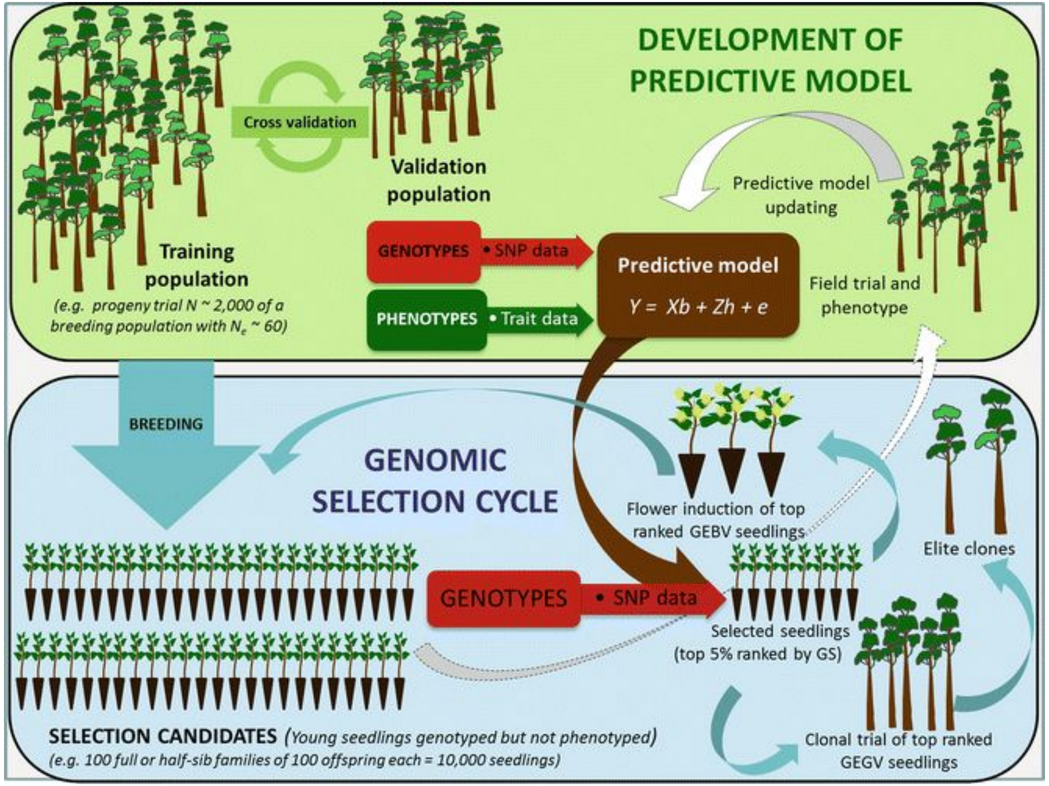
CrossMark

Accelerating crop genetic gains with genomic selection

Kai Peter Voss-Fels[1] · Mark Cooper[1] · Ben John Hayes[1]

Received: 9 September 2018 / Accepted: 12 December 2018 / Published online: 19 December 2018

Genomic Selection



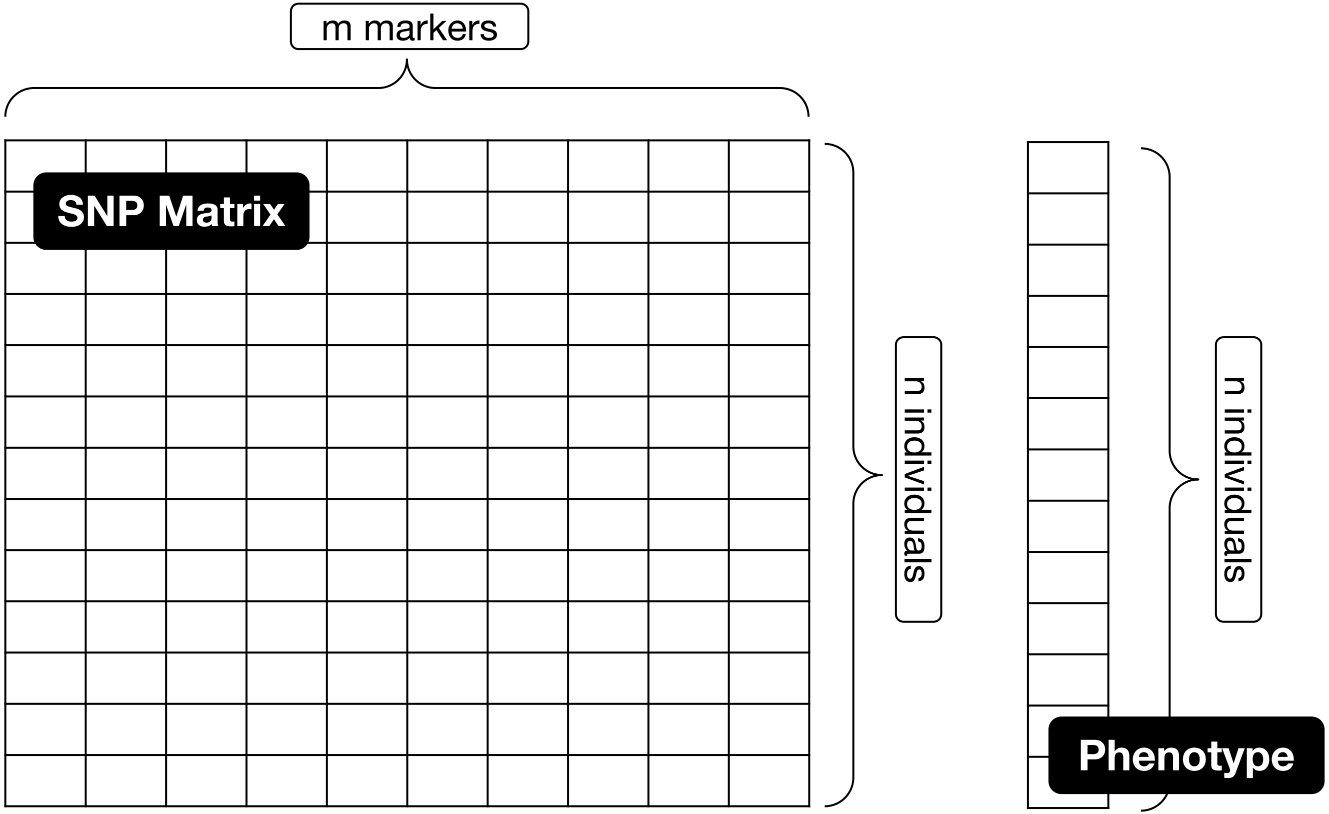
Grattapaglia, Dario. "Status and perspectives of genomic selection in forest tree breeding." *Genomic selection for crop improvement: New molecular breeding strategies for crop improvement* (2017): 199-249.

Genomic Selection Applied to Breeding Programs

- Reducing the time to develop cultivars
- Increasing the effective size and selection intensity
- Increasing the genetic gain per unit time

Genomic Prediction Models

Genomic Prediction



Daetwyler, Hans D., et al. "Genomic prediction in animals and plants: simulation of data, validation, reporting, and benchmarking." *Genetics* 193.2 (2013): 347-365.

Henderson Equations

$$\mathbf{y = Xb + Zu + \square}$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + \Sigma^{-1} \dfrac{\sigma_{\square}^{2}}{\sigma_{u}^{2}} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\square \sim N(0, R\sigma_{\square}^{2})$$

$$u \sim N(0, \Sigma\sigma_{u}^{2})$$

SNP-BLUP

SNP Matrix (label pointing to $\mathbf{Z}$)

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + I \dfrac{\sigma_{\square}^{2}}{\sigma_u^{2}} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

All SNPs explain the same proportion of variance on each trait

$$\square \sim N(0, R\sigma_{\square}^{2})$$

$$u \sim N(0, I\sigma_u^{2})$$

G-BLUP

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \square$$

$$\begin{bmatrix} X^T y \\ Z^T y \end{bmatrix} = \begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + G^{-1}\dfrac{\sigma_{\square}^{2}}{\sigma_{u}^{2}} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix}$$

$$\square \sim N(0, R\sigma_{\square}^{2})$$

$$u \sim N(0, G\sigma_{u}^{2})$$

Genomic Relationship Matrix

G-BLUP

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_u^2 \right]$$

$$G_{jk} = \sum_{i=1}^{m} \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$$

G-BLUP

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_u^2\right]$$

Going SNP by SNP

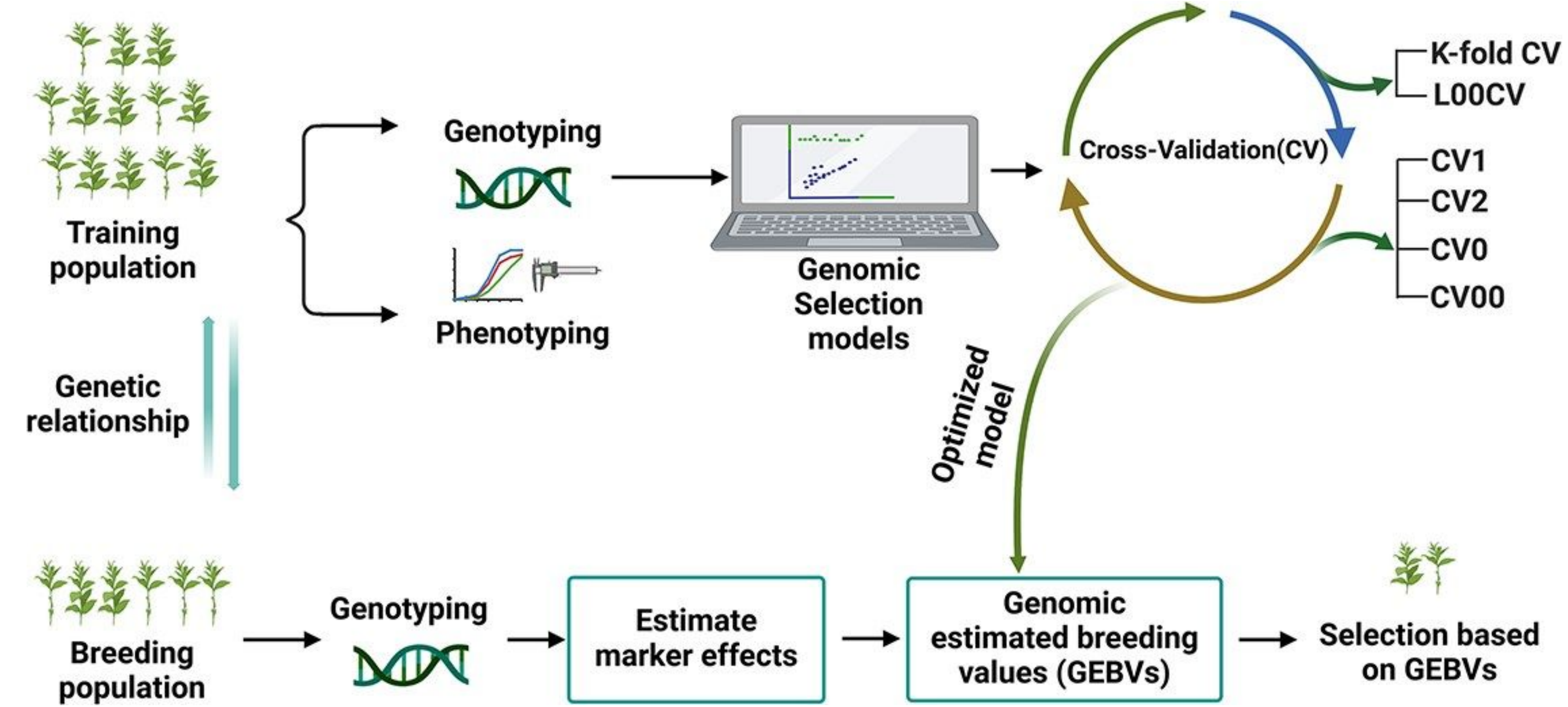
$$G_{jk} = \sum_{i=1}^{m} \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$$

Genomic relationship between individual j and individual k

Minor allele frequency of SNP i

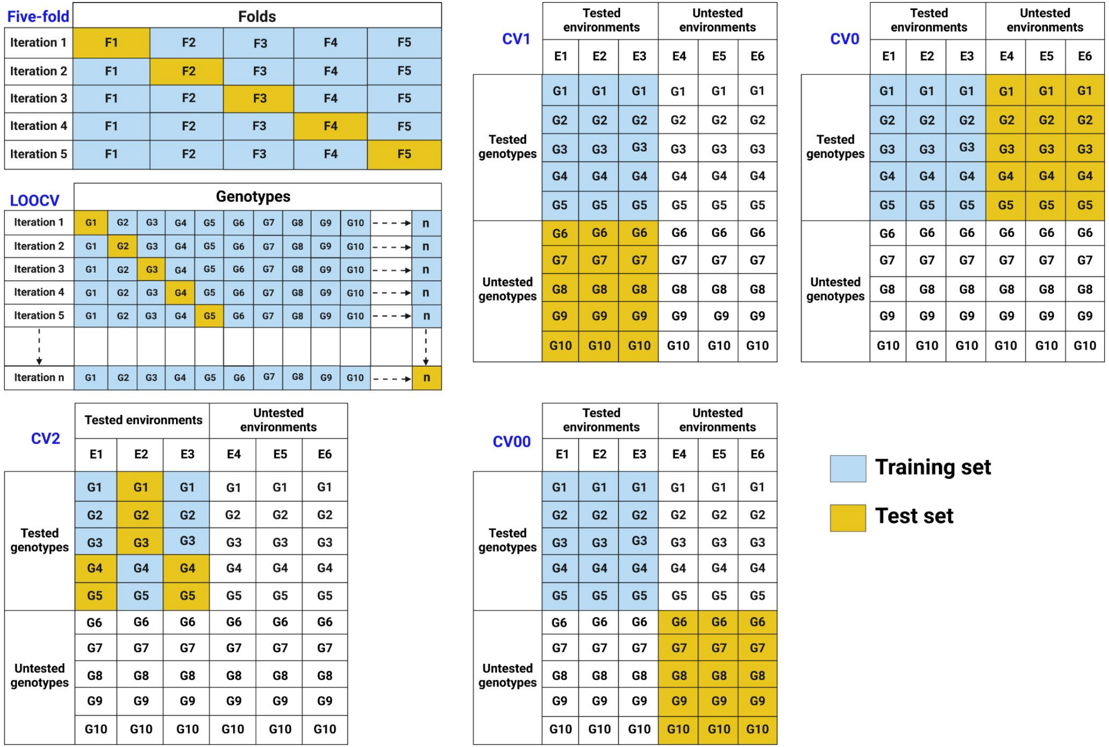
Cross Validation Strategies

Cross Validation Strategies



Mbebi, Alain J., et al. "Advances in multi-trait genomic prediction approaches: classification, comparative analysis, and perspectives." *Briefings in Bioinformatics* 26.3 (2025): bbaf211.

Cross Validation Strategies



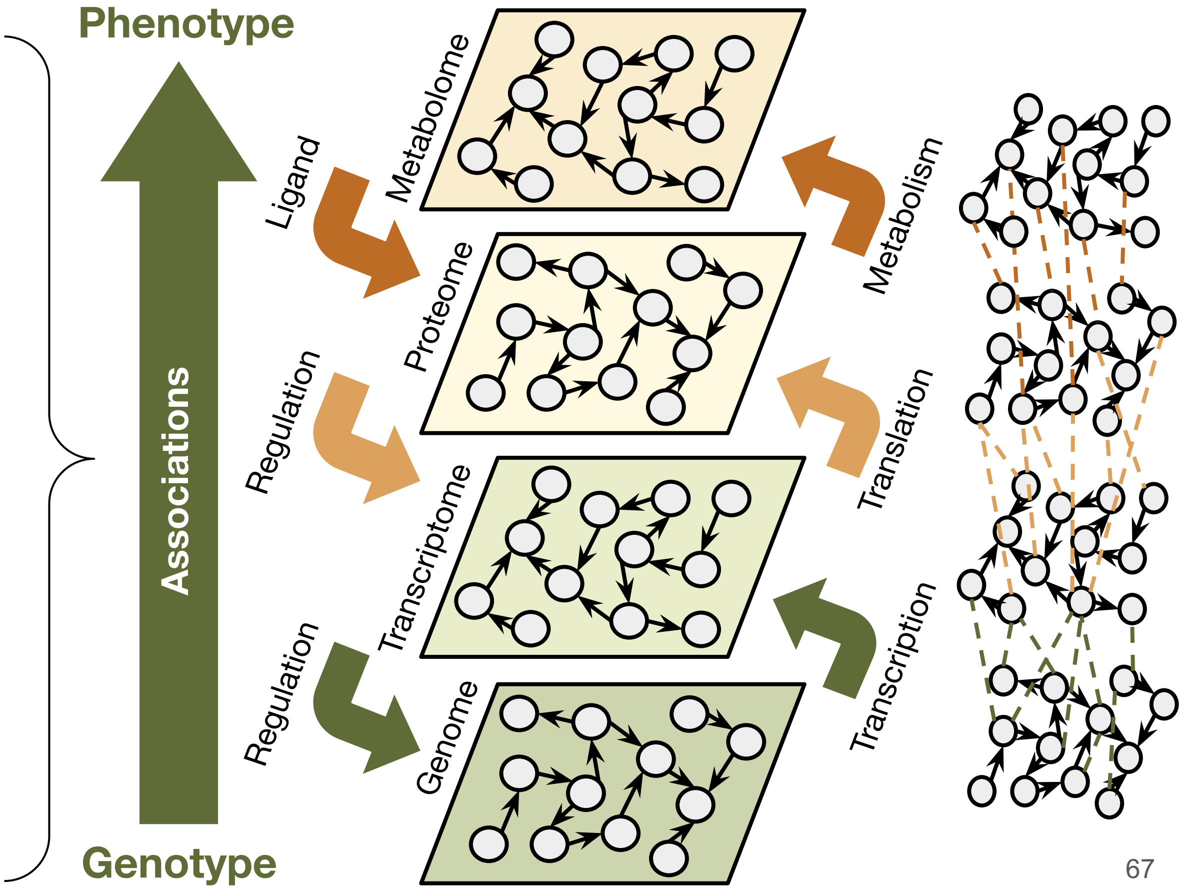
Alemu, Admas, et al. "Genomic selection in plant breeding: Key factors shaping two decades of progress." *Molecular Plant* 17.4 (2024): 552-578.

Factor Affecting Prediction Accuracy

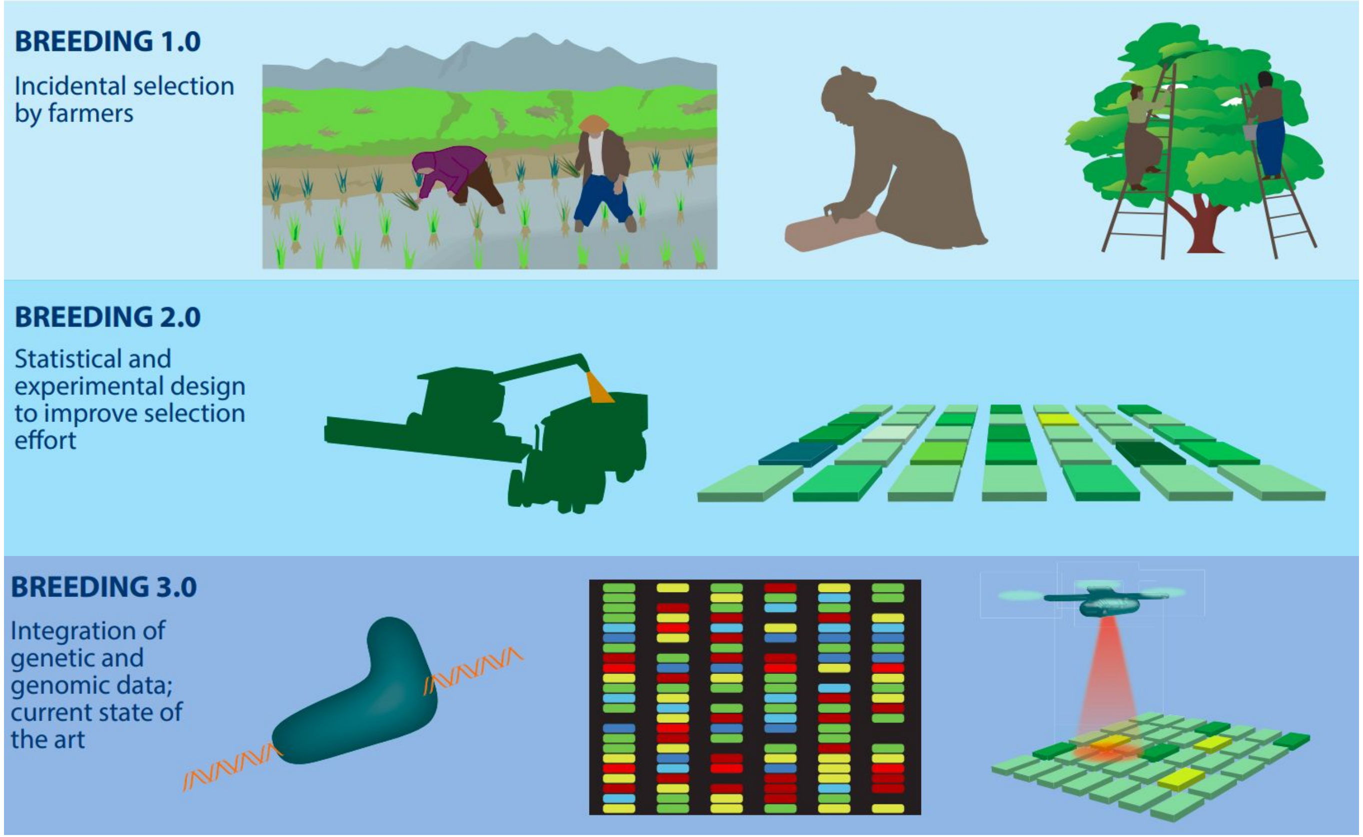
- Marker density
- Population diversity and effective population size
- Training set
- Genetic relationship between training population and selection candidates
- Rare alleles
- Missing data and imputation
- Ploidy
- GxE

Multi-Omics Approaches

Unraveling Complex Phenotypes

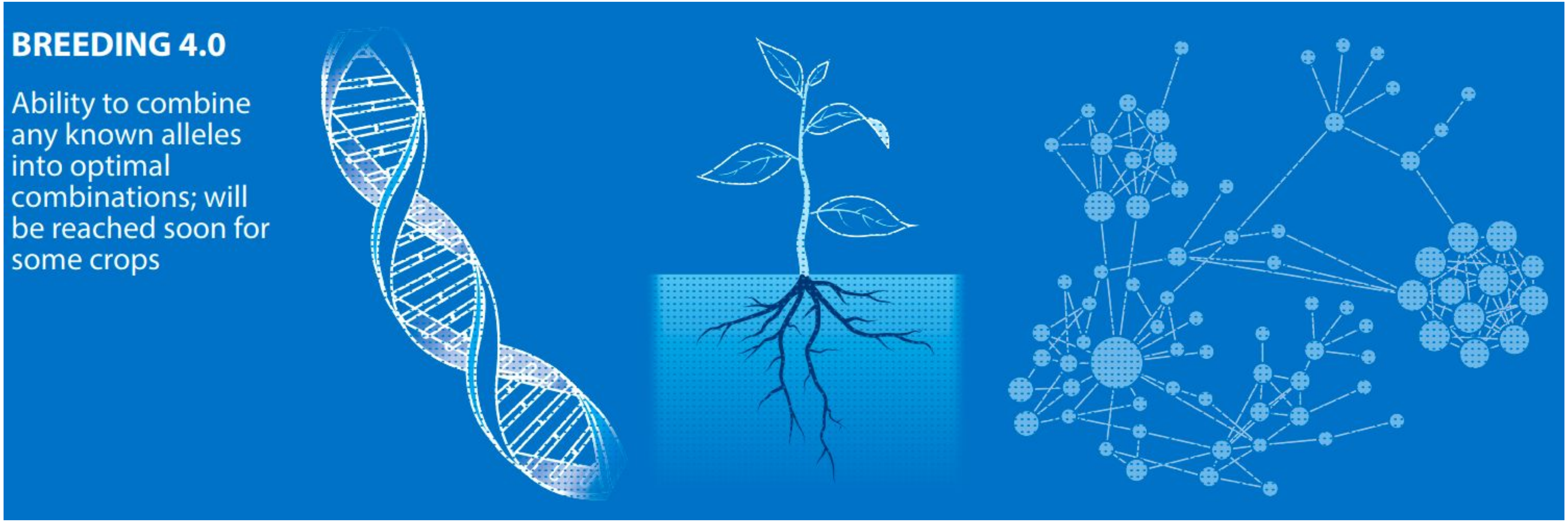


Breeding Stages



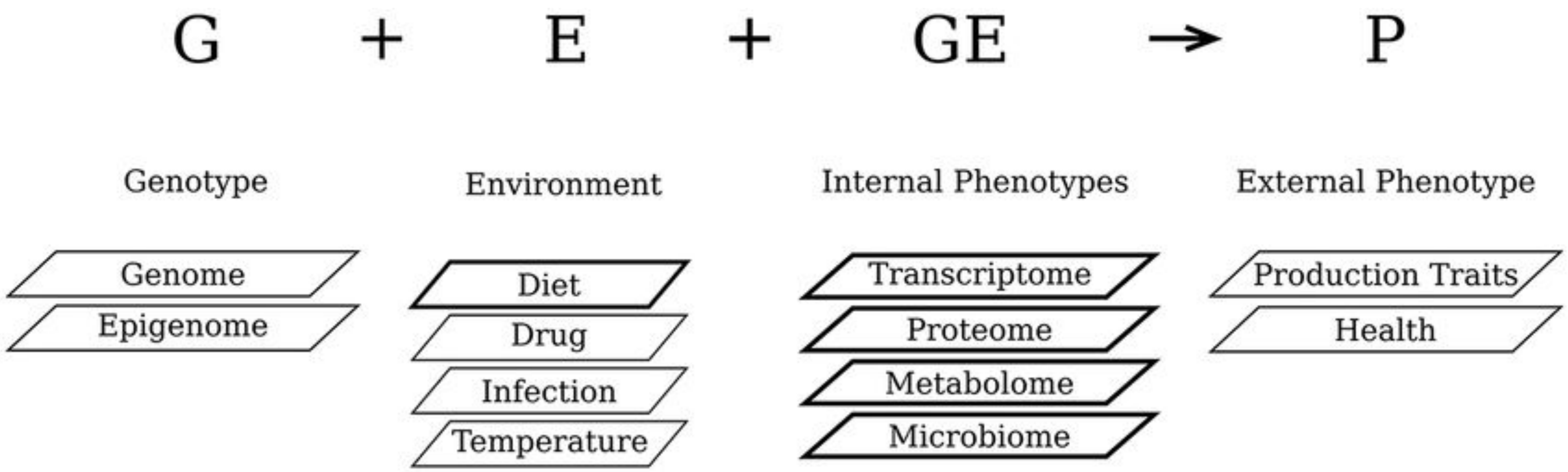
Wallace, Jason G., Eli Rodgers-Melnick, and Edward S. Buckler. "On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics." *Annual review of genetics* 52 (2018): 421-444.

Breeding Stages



Wallace, Jason G., Eli Rodgers-Melnick, and Edward S. Buckler. "On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics." *Annual review of genetics* 52 (2018): 421-444.

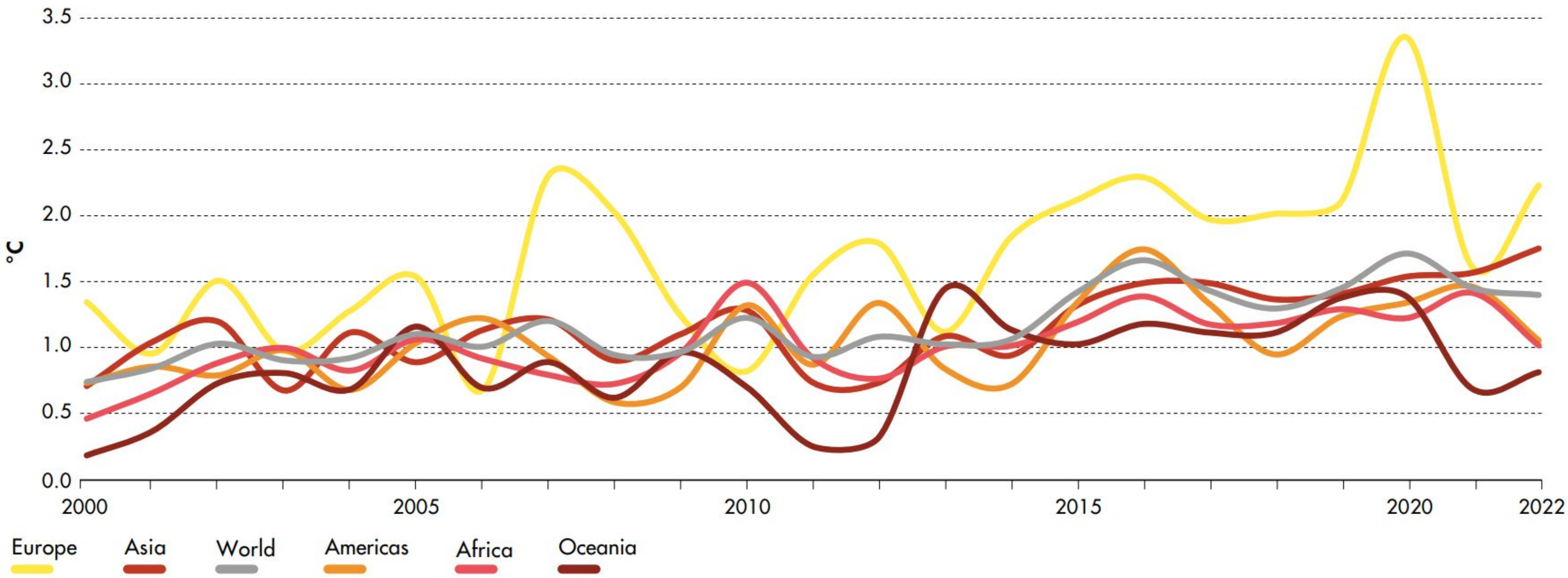
Unraveling Complex Phenotypes



Benis, Nirupama, et al. "Multi-level integration of environmentally perturbed internal phenotypes reveals key points of connectivity between them." *Frontiers in Physiology* 8 (2017): 388.

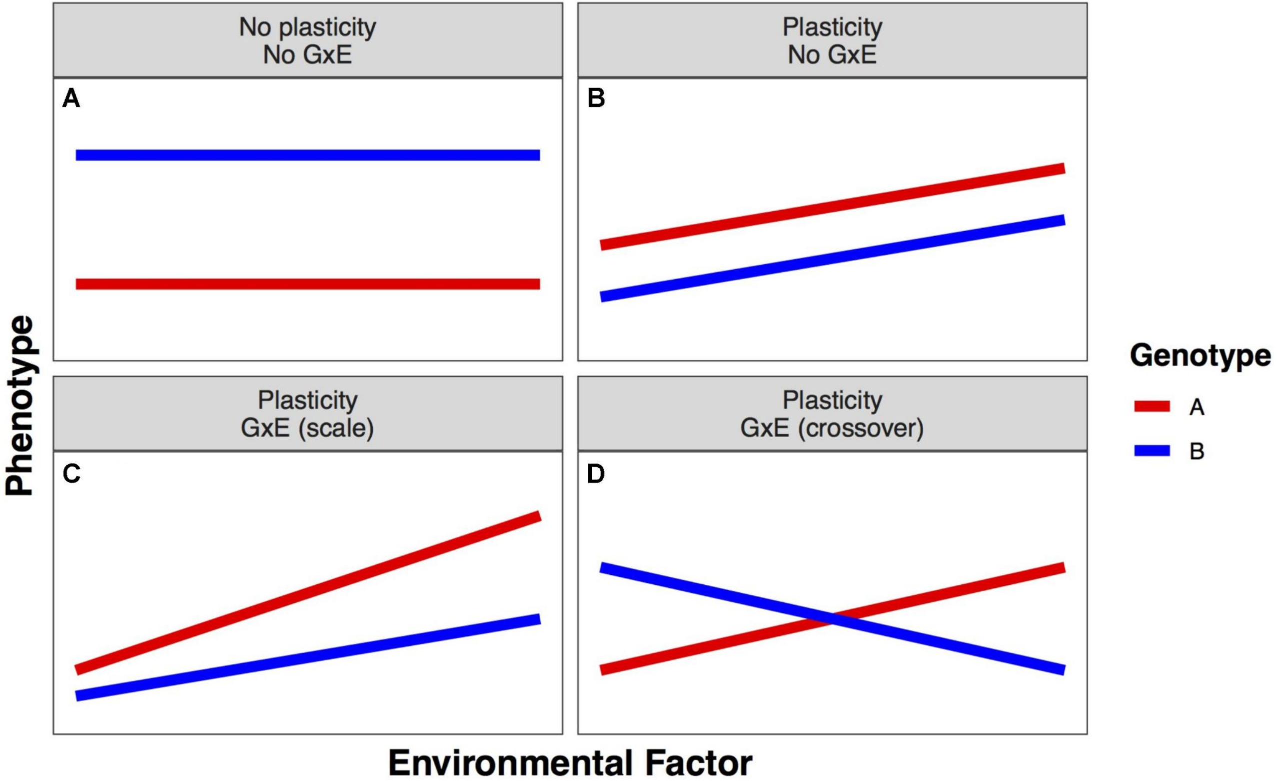
Enviromics

Environmental Changes



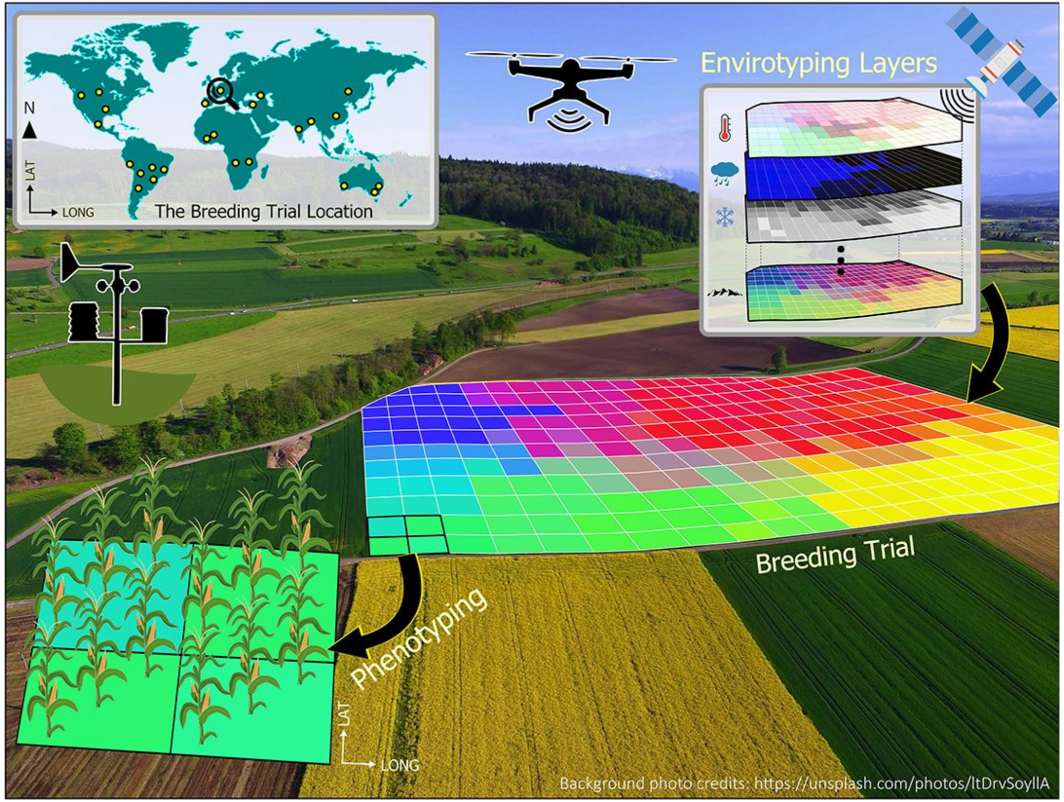
Food and Agriculture Organization of the United Nations

GxE



Kusmec, Aaron, Natalia de Leon, and Patrick S. Schnable. "Harnessing phenotypic plasticity to improve maize yields." *Frontiers in Plant Science* 9 (2018): 1377.

Environmental Changes



Resende, Rafael Tassinari, et al. "Enviromics in plant breeding." *Frontiers in Plant Science* 13 (2022): 935380.