

## Plant and Animal Breeding



Miguel Pérez-Enciso

## Outline

- Breeding principles.
- Within breed improvement.
- Crossbreeding and heterosis.
- Genetic modification.

## Facts



There exists

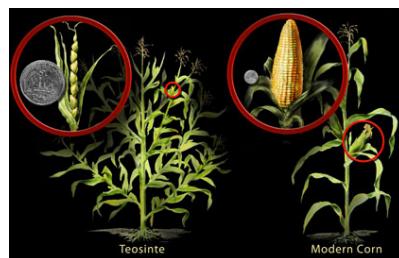
enorMous  
variability in nature



Part of this variability is inherited,  
i.e., offspring resemble parents  
more than unrelated individuals.

## Main events: domestication

Teosinte → Corn

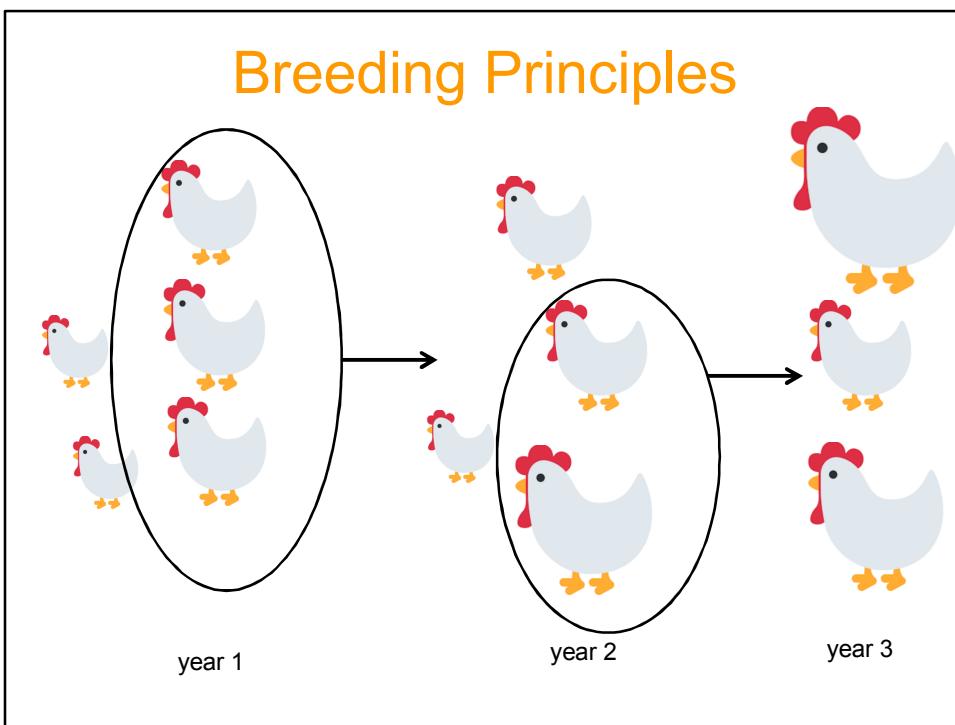
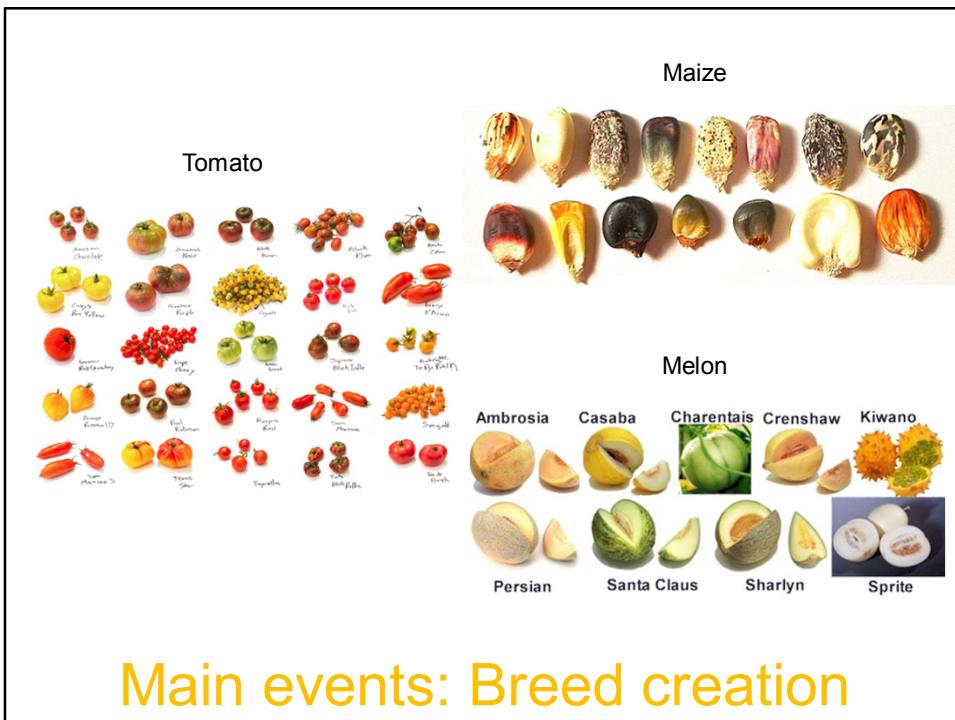


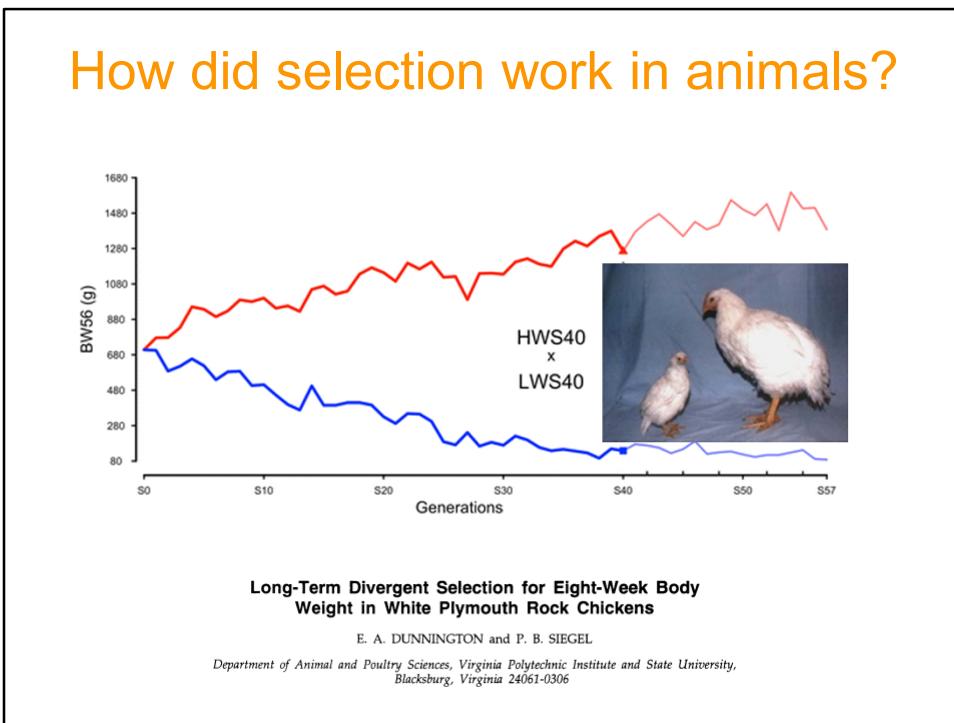
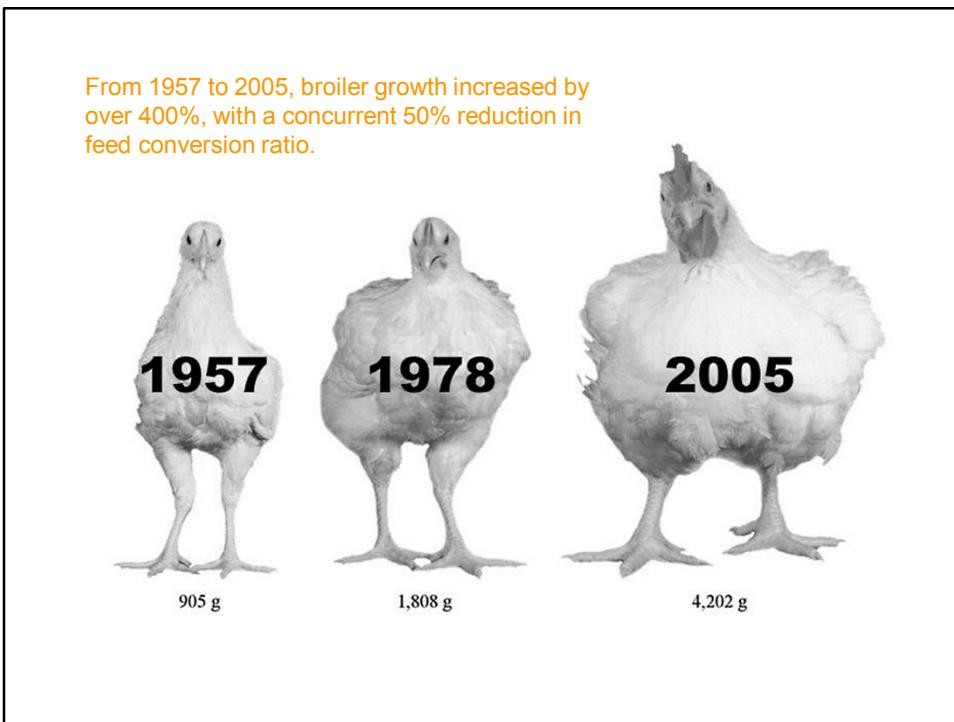
Mesoamerica, ~ 9,000 years ago



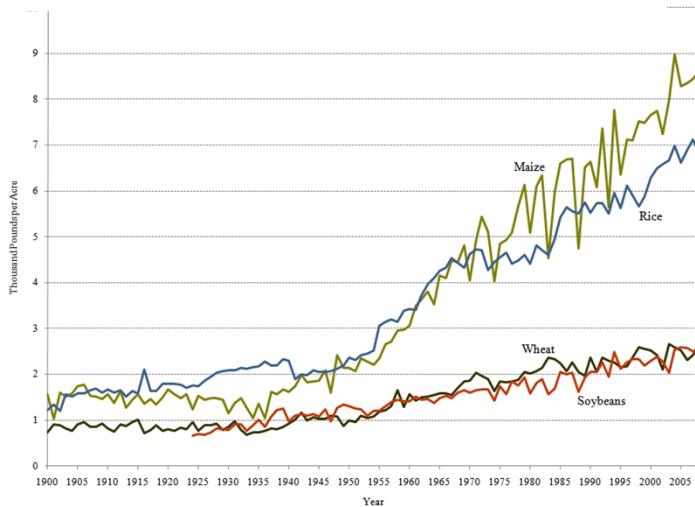
Wild boar

Asia and Anatolia,  
~ 9,000 years ago





## Increase in average performance: cereals

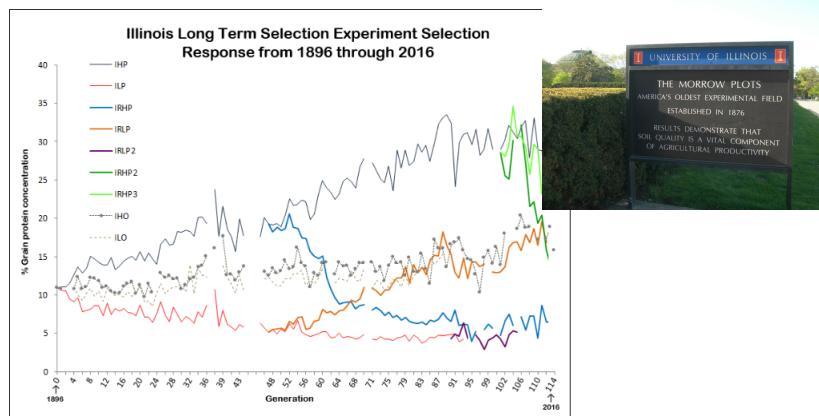


### Illinois Long-term Selection Experiment

#### Project Overview:

In 1896, University of Illinois scientists initiated an experiment that is now the longest running continuous genetics experiment in higher plants. Beginning with an open pollinated variety of maize, Bur's White, the founders selected for ears with the highest or lowest concentrations of grain protein or oil. Over a century has elapsed, and one-hundred fifteen cycles of recurrent selection has created twelve populations that vary significantly in their grain protein and oil composition.

Our lab maintains this experiment, but our research focus is on the protein selections. Illinois High Protein (IHP) and Illinois Low Protein (ILP) represent the known extremes for protein concentrations in maize (see figure below). After forty-eight cycles of forward selection, breeders reversed the direction of selection by selecting low protein ears from IHP and high protein ears from ILP, thereby creating two new lines, Illinois Reverse High Protein (IRHP) and Illinois Reverse Low Protein (IRLP). Collectively, these four strains are referred to as the Illinois Protein Strains (IPS). The IPS provide a unique source of genetic material with which to make comparisons both within and among the strains and generations, at the DNA, mRNA, and protein levels, with a goal of elucidating the dynamics of evolutionary changes during these intervals. They are also a useful resource for discovering genes that contribute to seed protein accumulation, not only in maize, but other crop species.



Selection responses in the Illinois Protein Strains. Illinois High Oil (IHO) and Illinois Low Oil (ILO) are part of the same source population, but were not selected for protein. Thus, they illustrate the effects of random genetic drift. Selection for oil was discontinued. IRHP: Illinois Reverse high protein; IRLP: Illinois Reverse low protein.

# Main breeding concepts

- **Selection objective** ( $y$ ): The characteristic of interest.
  - **Heritability** ( $h^2$ ): The fraction of phenotypic variance explained by genes =  $\text{Var}(g) / \text{Var}(y)$ .
  - **Selection intensity** ( $i$ ): mean standardized value of selected individuals.
  - **Generation interval** ( $L$ ): the average age of parents when their selected offspring are born.

## Main breeding assumption

$$y = \mu + g + e$$

## Phenotype

# Genotype

'fixed' effects  
(water, soil, ...)

## Environmental error ('noise')

## Heritability

It measures how more similar are related individuals than unrelated ones

Is human height inherited?

Is human life span inherited?

Is the number of legs inherited?

## Heritability

It measures how more similar are related individuals than unrelated ones

Is human height inherited?

Is human life span inherited?

Is the number of legs inherited? Does show genetic variance?

## Heritability

limits

$$0 \leq h^2 \leq 1$$

$$h^2 = \frac{Var(g)}{Var(y)}$$

$h^2 = 0$

No hope

$0 < h^2 < 1$

- Most traits fall here:
- Low  $h^2$ : stress resistance, reproduction
- High  $h^2$ : conformation, growth

$h^2 = 1$

What you  
see is what  
you get

How to  
estimate  
heritability?  
Simple

Genetic relationship coefficient: it measures probability of genes being identical in relatives

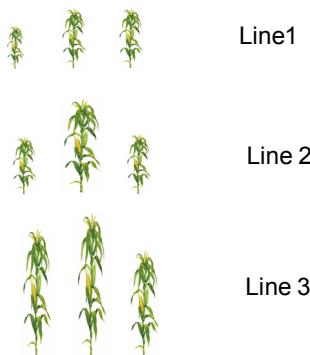
Obtained from pedigree or marker information

$$\text{Cov}(y, y') = \text{Cov}(g+e, g'+e') = \text{Cov}(g, g') = r \text{Var}(g)$$

$$h^2 = \text{Cov}(y, y') / r \text{Var}(y)$$

# How to estimate heritability in inbred lines ?

Easy

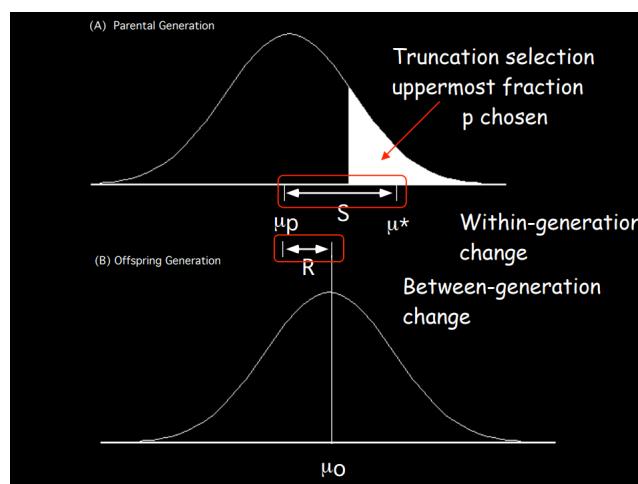


$$Var(e) = \sum_i Var(y_i) / n$$

$$h^2 = \frac{Var(g)}{Var(g) + Var(e)}$$

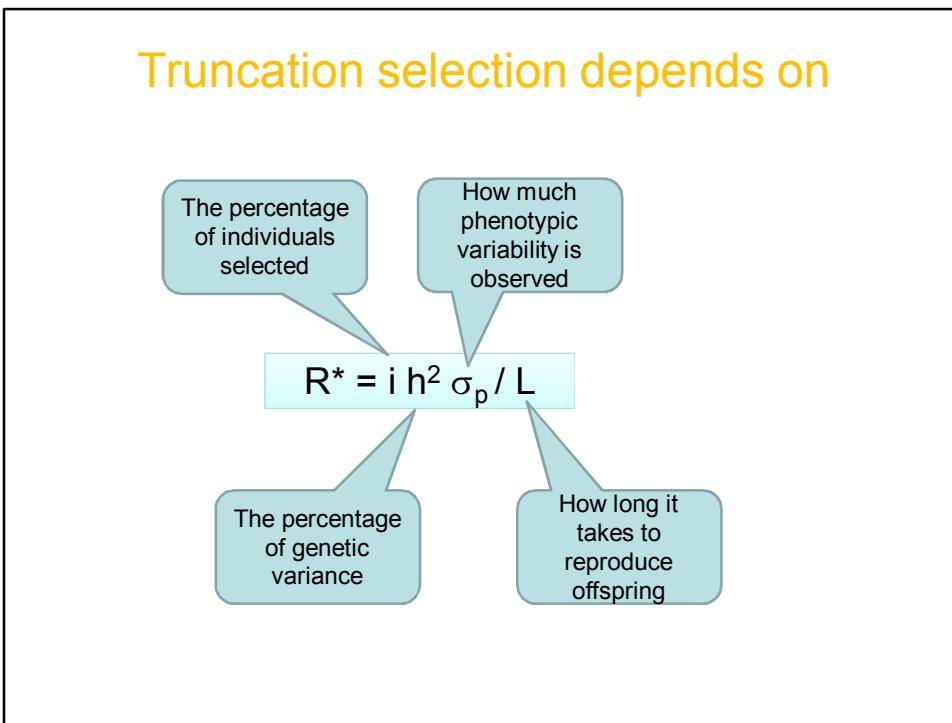
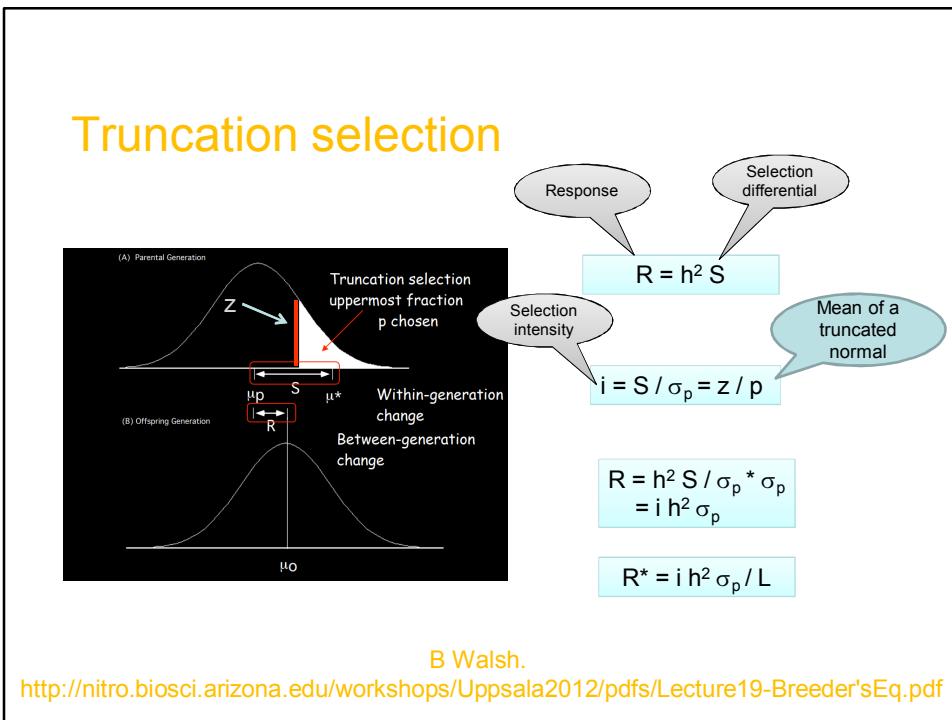
$$Var(g) = \sum_i Var(\mu_i) / n$$

## Truncation selection

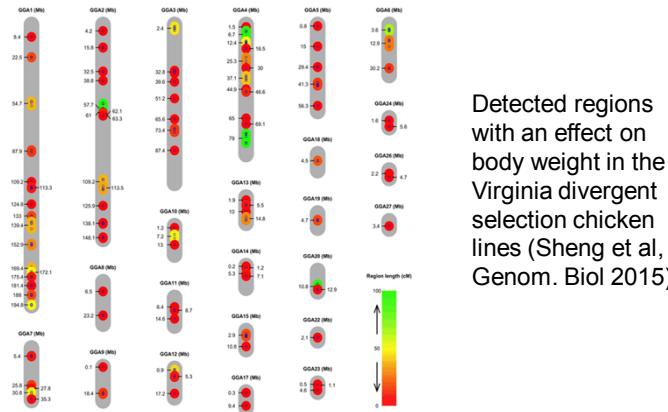


B Walsh.

<http://nitro.biosci.arizona.edu/workshops/Uppsala2012/pdfs/Lecture19-Breeder'sEq.pdf>



**IMPORTANT!!! Many genes affect the traits of interest**



**'The infinitesimal model' is a common assumption**



It presupposes that quantitative traits are explained by a large number of genes, each acting individually and of small effect. In addition, quantitative traits are also modified by the environment.

XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. Communicated by Professor J. ARTHUR THOMSON. (With Four Figures in Text.)

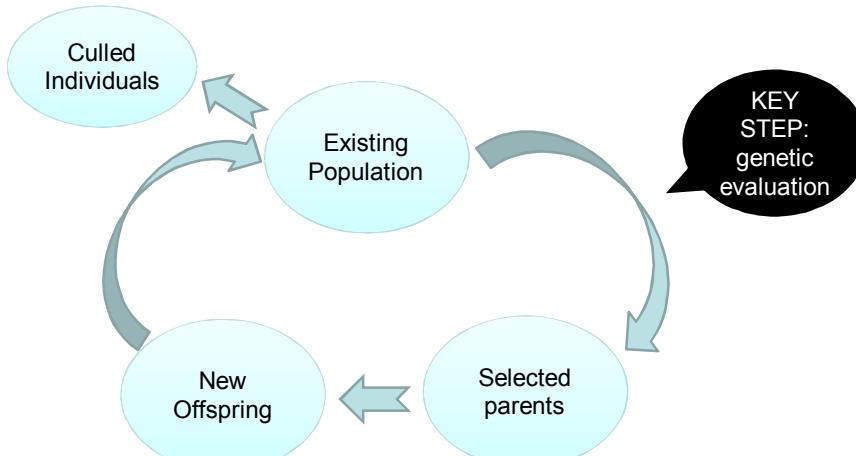
(MS. received June 16, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

## Generalization

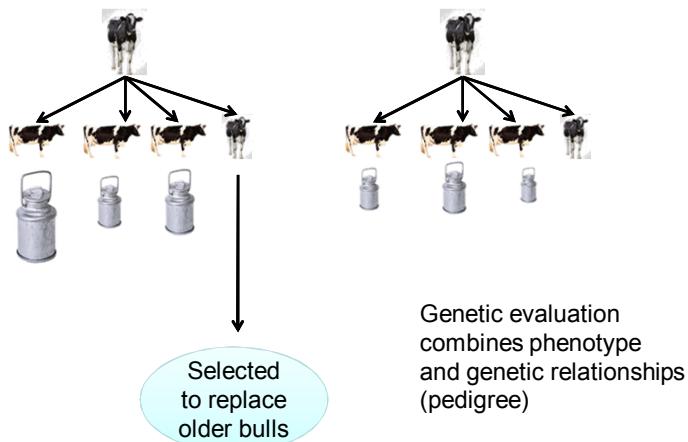
$$\begin{aligned}
 R^* &= i h^2 \sigma_p / L \\
 &= i \sigma_a^2 / \sigma_p^2 \sigma_p / L \\
 &= i \sigma_a / \sigma_p \sigma_a / L \\
 &= i \rho \sigma_a / L
 \end{aligned}$$

Accuracy of prediction:  
correlation between  
selection criterion and  
breeding value

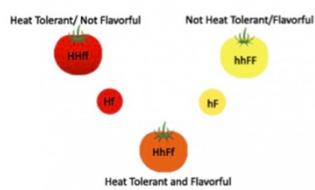
## How things work in practice



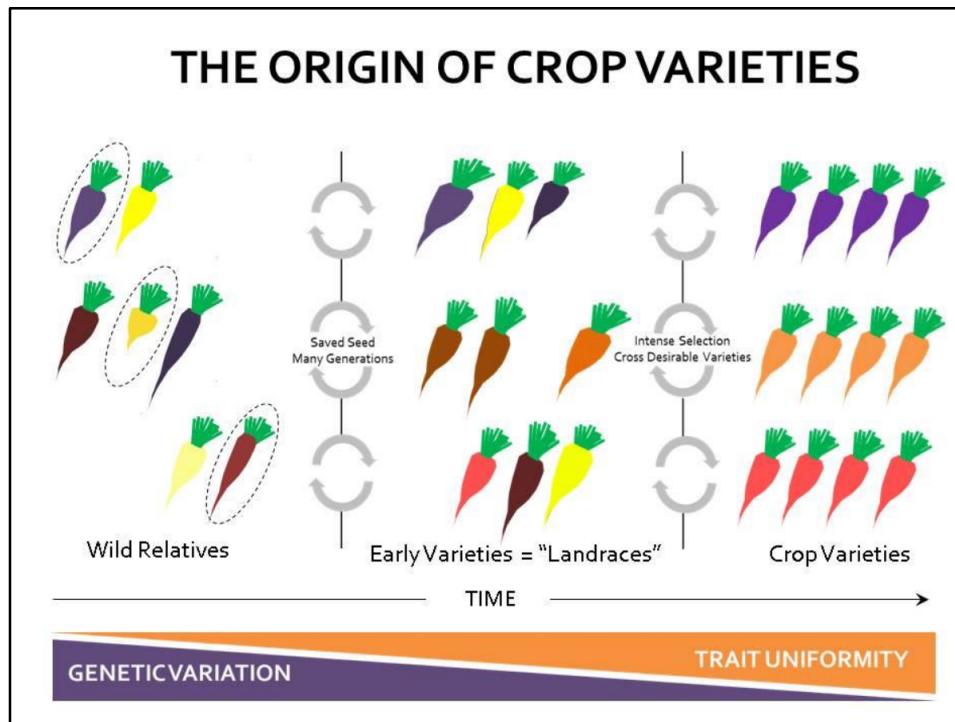
## How Animal improvement works



## How Plant improvement works



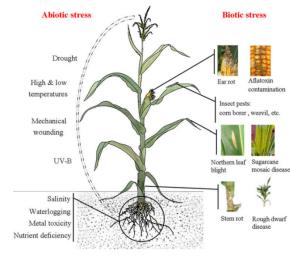
Sloooooow process



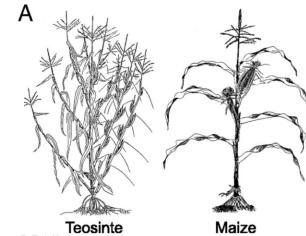
## First thing you need: A breeding goal

- We need to define a **breeding goal** or set of phenotypes we wish to improve.
- Usually a set of phenotypes weighed by some constants that define the economic value per unit increase. This makes a '**selection index**'.
- They are moving targets, not fixed, but need to be relatively stable over the years because genetic improvement is slow.

## First thing you need: A breeding goal



Consumer preferences (taste, size...)



Plant architecture

## Usual plant breeding goals



Milk yield

## Usual breeding goals in animals



Egg production



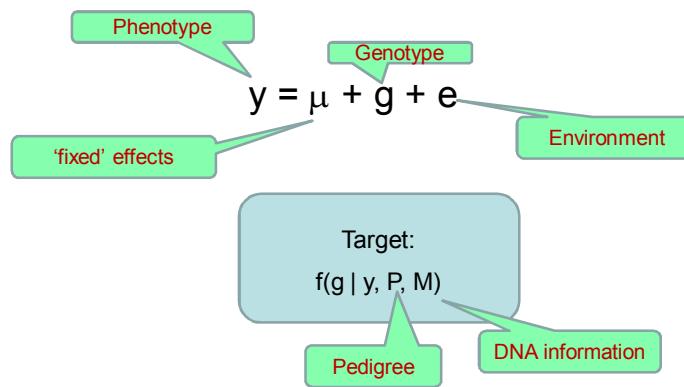
Wool diameter

## Different objectives same species

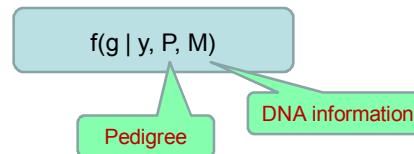


## Second thing you need: genetic evaluation

Genetic evaluation aims at predicting the offspring potentially produced by a given individual, which is called the '**breeding value**' ( $g$ ) of that individual.



## Second thing you need: genetic evaluation



Genetic Evaluation uses:

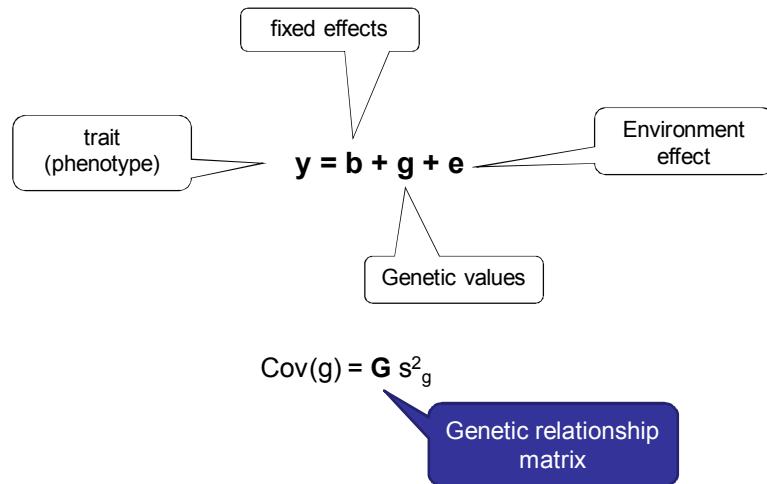
- Own performance
- Performance of relatives
- Marker information

## How things work in practice: genetic evaluation with BLUP

- BLUP stands for 'Best Linear Unbiased Prediction'.
- It was invented by Charles Henderson, the founder of modern breeding methods in the late 50s.
- The rationale is to use all information in a coherent approach, and use the same equations irrespective of the dataset.
- It is linear on phenotypes, it is best in that minimizes quadratic error and is unbiased.



## How things work in practice: genetic evaluation with BLUP



## How things work in practice: Mixed Model Equations

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

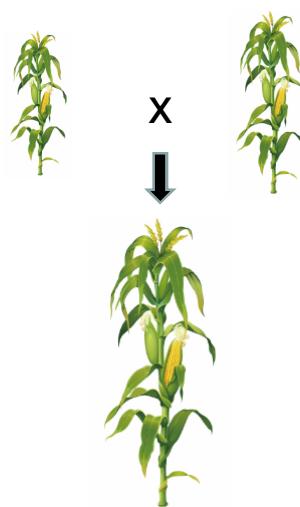
Annotations for the Mixed Model Equations:

- Incidence matrix** (top left)
- Incidence matrix** (top middle)
- phenotypes** (top right)
- Constant that depends on  $h^2$**  (bottom left)
- BLUP breeding values** (bottom right)

## How things work in practice: Cross breeding and heterosis

What is  
heterosis?

What is heterosis?



Or hybrid vigor  
means the  
offspring is better  
than average  
parents

## Why heterosis?

Assume a quantitative trait affected by a biallelic locus



Mean phenotypic values for each genotype

Biallelic locus, b & B with frequencies p & q

Value	$-a$	$d$	$a$
Frequency	$q^2$	$2pq$	$p^2$

## Why heterosis?

Genotypes



bb



bB



BB

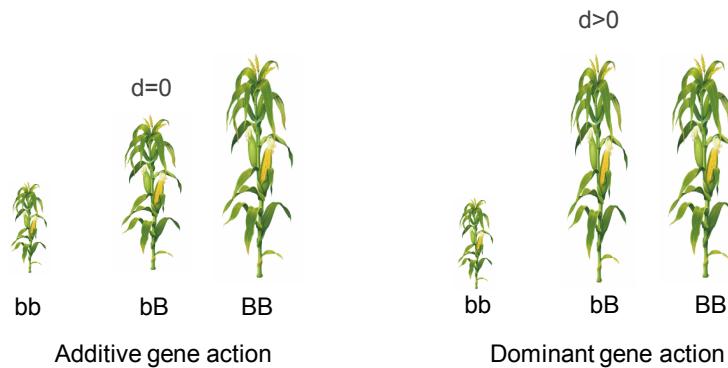
d = dominance effect

Value

Value	$-a$	$d$	$a$
Frequency	$q^2$	$2pq$	$p^2$

a = additive effect

## Why heterosis?



Heterosis occurs only if  $d \neq 0$

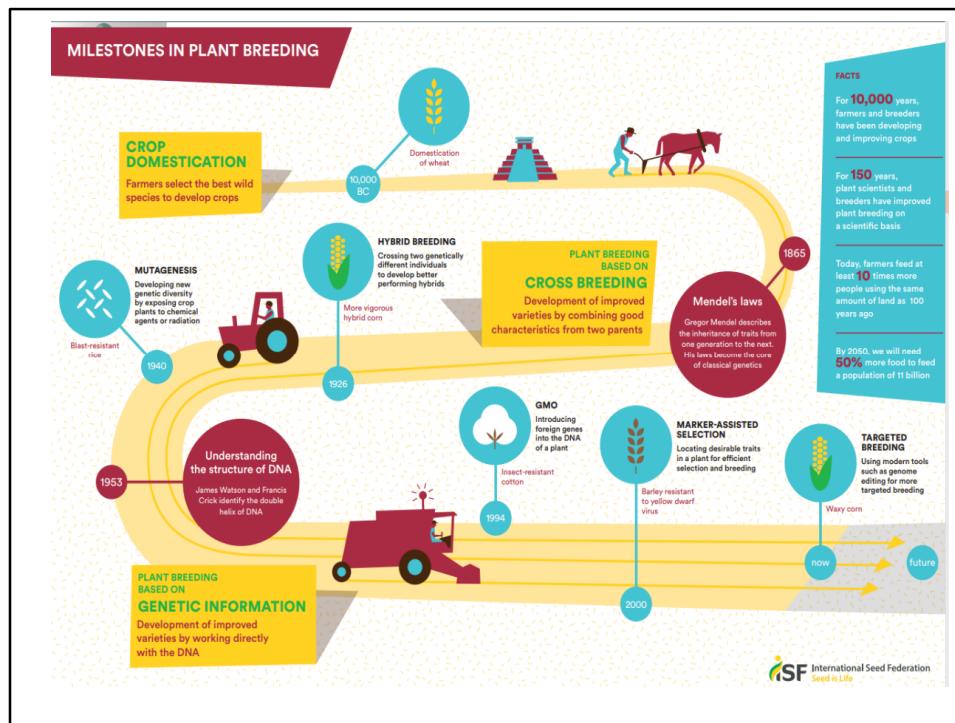
## Why is it called hybrid vigor?

- Because, when  $d > 0$ , the hybrid is better than the average of the two parents.
- It normally refers to crosses between pure lines.

Exercise: make a R function to compute expected mean of crosses between lines with different allele frequencies.

## Why hybrid lines are so widely used?

- To benefit from hybrid vigor.
- They allow combining properties of different pure lines.
- Companies retain pure lines and sell F1 seeds.
- First used in corn in the 1925's in Iowa (before Green revolution)



## Gene editing and modification

- Most traits are influenced by numerous genes, many of them uncharacterized, this makes it difficult to genetically modify causal genes.
- Some genes and their causal mutations are actually known, which has permitted transgenesis and gene-editing, mainly in plants.
- Societal concerns mainly in Europe are important, which causes a big loss in competitiveness of European companies (although most are multinational).

## Why can be useful?

- Speed up introgression of desired alleles, say disease resistance in sensitive lines.
- To be combined with pesticide treatments (Monsanto glyphosphate – resistant soya).
- Create new varieties with new characteristics.



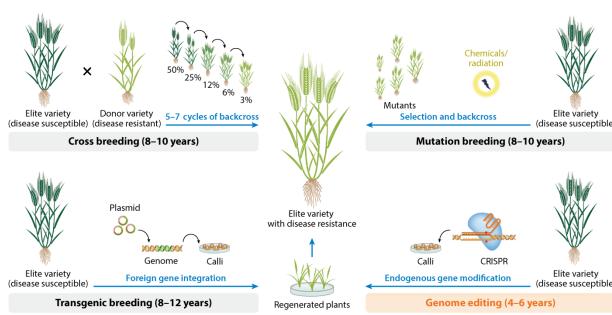
*Annual Review of Plant Biology*

## CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture

Kunling Chen,<sup>1,\*</sup> Yanpeng Wang,<sup>1,\*</sup> Rui Zhang,<sup>1</sup>  
Huawei Zhang,<sup>1</sup> and Caixia Gao<sup>1,2</sup>

<sup>1</sup>State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome  
Editing, Institute of Genetics and Developmental Biology, The Innovative Academy of Seed  
Design, Chinese Academy of Sciences, Beijing, China 100101; email: cxgao@genetics.ac.cn

<sup>2</sup>University of Chinese Academy of Sciences, Beijing, China 100864



Chen, K., et al. 2019.

*Annu. Rev. Plant Biol.* 70:667–97.

Annual Reviews

## Summary

- ✓ Artificial selection is very effective in changing the phenotypes.
- ✓ Many genes involved: no exhaustion of genetic variability (this is called the INFINITESIMAL MODEL).
- ✓ Genetic improvement is produced by both selection within lines and by crossing between different lines.
- ✓ Genetic modification is popular in plants but not so much in animals because plant breeding is largely based in inbred lines.