



(HOS 6932)– Survey of Breeding Tools and Methods

Theoretical and Practical Aspects for Genomic Selection Implementation

Felipe Ferrão

Research Assistant Scientist

ferrao@ufl.edu

felipeferrao

February, 2023

Introduction

In the last class ...

- Importance of training and testing data set
- Metrics to measure the predictive ability
- RR-BLUP vs. GBLUP

In this class ...

- Why and how to switch to genomic selection?
- Example in blueberry
- Final discussion

Introduction

Copyright © 2001 by the Genetics Society of America

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

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

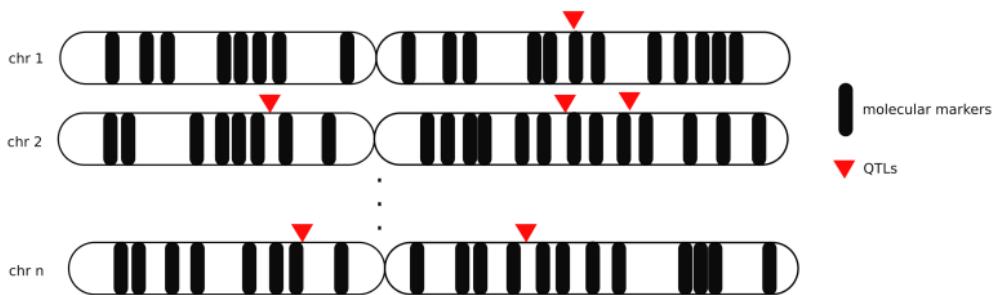
*Research Institute of Animal Science and Health, 8200 AB Lelystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia

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



Introduction

General Idea



$$\boxed{\text{•}} = \boxed{\text{DNA}} \boxed{a} + \boxed{e}$$

↓ ↓

Genomic Estimated
Breeding Values

$$\boxed{\text{•}} = \boxed{\text{plant}} \boxed{u} + \boxed{e}$$

↓ ↓

Relationship Matrix

Introduction

Background

- After this pioneer study genomic prediction has popularized in plant and animal breeding
- Organization of breeding programs can be very different between crops:
 - ▷ Clonally propagated species
 - ▷ Perennial crop
 - ▷ Genomic Selection in Hybrid Breeding
 - ▷ Self-pollinated crops
 - ▷ Family structures
- It is a dynamic process: implementation can change over time according to the evolution of technologies, knowledge, and breeding goals.

Breeder's Equation

Central Question

Why and how to switch to genomic selection?

Simple answer: GS can impact all terms in the Breeder's equation:

$$\Delta G = \frac{i \times r \times \sigma_g^2}{L}$$

- ΔG = genetic gain per cycle
- i = intensity of selection
- r = accuracy of selection
- L = breeding cycle time

Discuss element-by-element in the Breeder's equation

Breeder's Equation

Central Question

Why and how to switch to genomic selection?

Simple answer: GS can impact all terms in the Breeder's equation:

$$\Delta G = \frac{i \times r \times \sigma_g^2}{L}$$

- ΔG = genetic gain per cycle
- i = intensity of selection
- r = accuracy of selection
- L = breeding cycle time

Discuss element-by-element in the Breeder's equation

Breeder's Equation

Central Question

Why and how to switch to genomic selection?

Simple answer: GS can impact all terms in the Breeder's equation:

$$\Delta G = \frac{i \times r \times \sigma_g^2}{L}$$

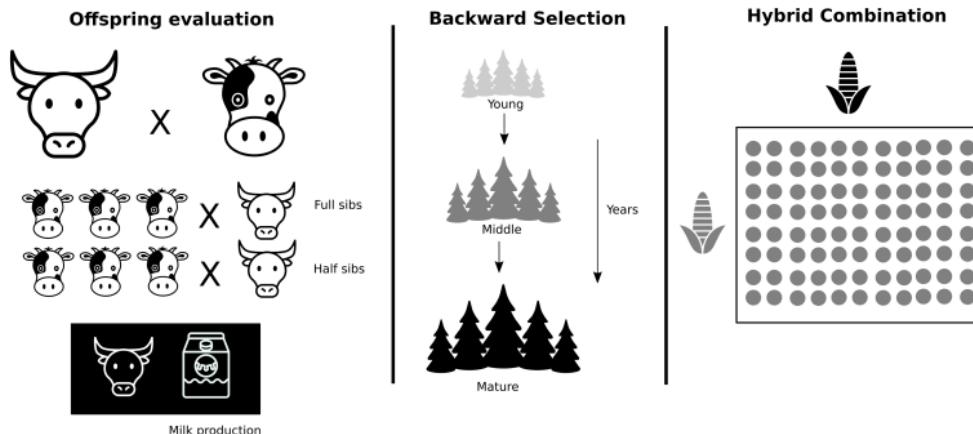
- ΔG = genetic gain per cycle
- i = intensity of selection
- r = accuracy of selection
- L = breeding cycle time

Discuss element-by-element in the Breeder's equation

Breeder's Equation

Breeding cycle time

- Breeding cycle is constrained by the way in which phenotype is collected
- ↑ Genetic Gain by ↓ breeding cycle

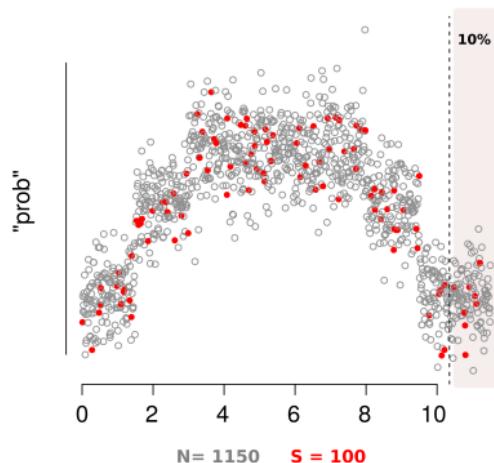


Breeder's Equation

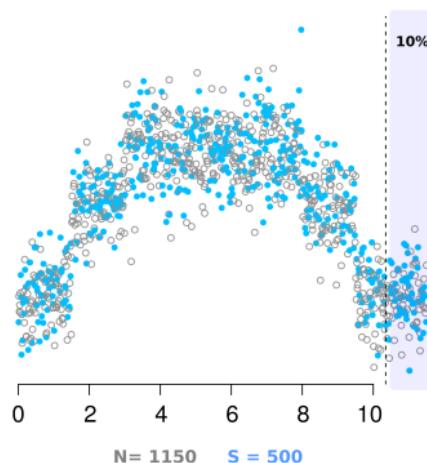
Intensity

- Lower costs with the same or higher testing capacity
- ↑ Genetic Gain by ↑ intensity

Phenotypic Selection



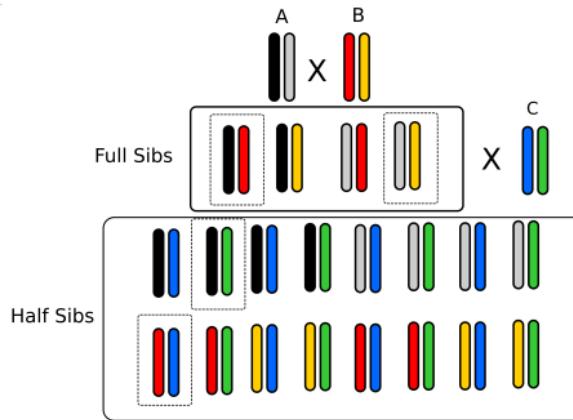
Genomic Selection



Breeder's Equation

Accuracy

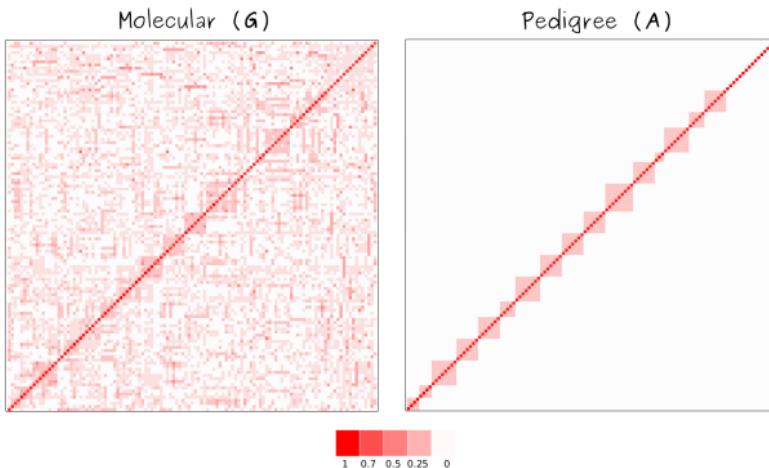
- Without genomic, full sibs have the same estimated breeding values
- Some genotypes can more related to the other, due to the inheritance of random samples of alleles from both parents.
- Mendelian Sampling variance
- ↑ Genetic Gain by ↑ accuracy



Breeder's Equation

Genetic Diversity

- Best cross: combination between highest BV and diversity
- Genomic relationship are more precise than realized kinship
- ↑ Genetic Gain by ↑ diversity



Dealing with accuracy

Background

- Predictive accuracy is a key metric for successful of genomic selection
- There are many deterministic formulas
- Check the hands-on

Expected accuracy

$$r = \sqrt{N_p h^2 [N_p h^2 + M_e]^{-1}}$$

- Daetwyler et al. (2013)
- N_p is the number of individuals in the TRN pop
- h^2 is the heritability of the trait
- M_e is the number of independent chromosome segments

Dealing with accuracy

Background

- Predictive accuracy is a key metric for successful of genomic selection
- There are many deterministic formulas
- Check the hands-on

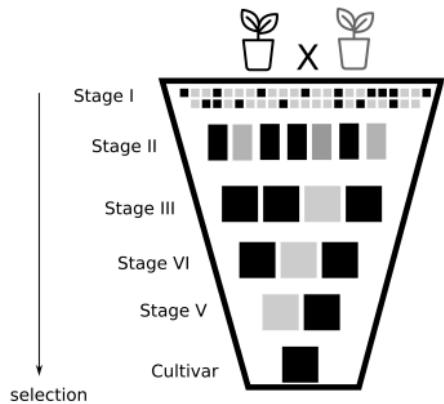
Expected accuracy

$$r = \sqrt{N_p h^2 [N_p h^2 + M_e]^{-1}}$$

- Daetwyler et al. (2013)
- N_p is the number of individuals in the TRN pop
- h^2 is the heritability of the trait
- M_e is the number of independent chromosome segments

Dealing with accuracy

Simulation Example:



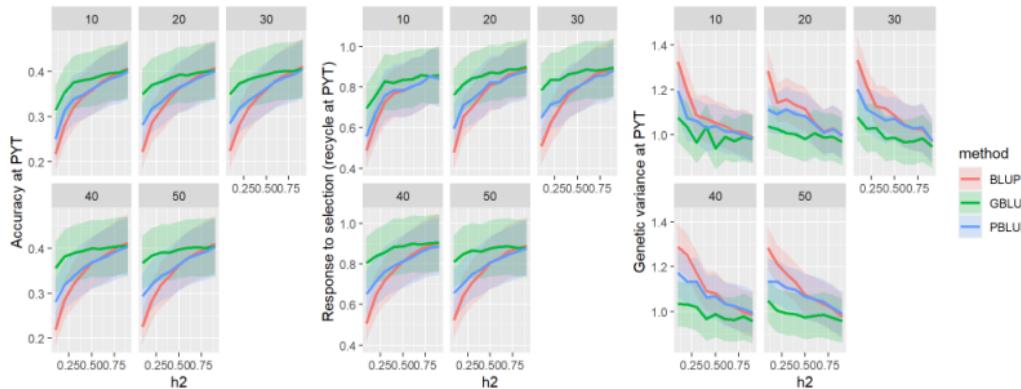
- Simulated 1000 materials in a clonal species
- Excellence in Breeding Platform
- Traits with different h^2
- Different family structures
- Selection based in GBLUP, PBLUP and BLUP

CGIAR – YouTube Channel – Basics of Genomic Selection

Dealing with accuracy

Simulation

Heritability vs. Population Structure vs. Accuracy

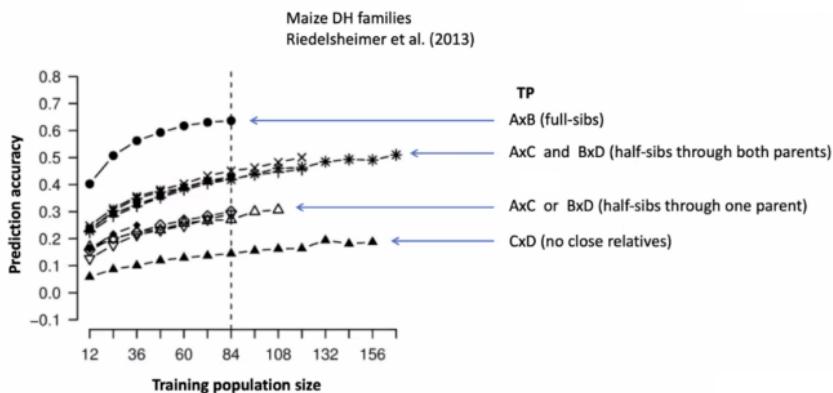


- GBLUP is the best for low-to-middle h^2
- With $\uparrow h^2$, we have \downarrow differences between ABLUP, BLUP or GBLUP.
- Lower gains in small families (large M_e)
- GBLUP tends to reduce faster the genetic variability

Dealing with accuracy

Maize

Predicting progeny of A x B with different training population set

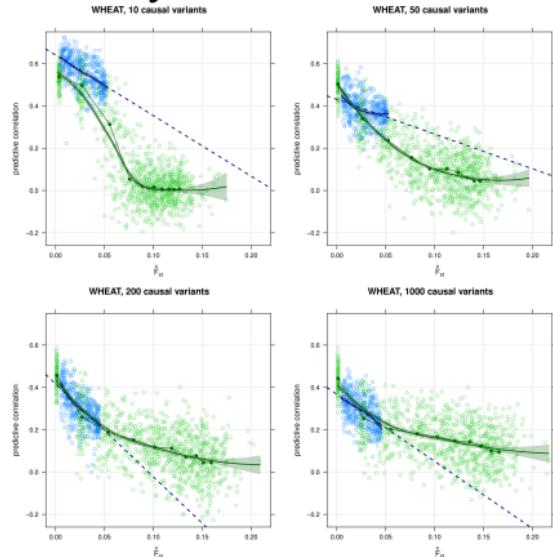


Riedelsheimer, 2013. <https://doi.org/10.1534/genetics.113.150227>

Dealing with accuracy

Wheat

Accuracy vs. Genetic Distance



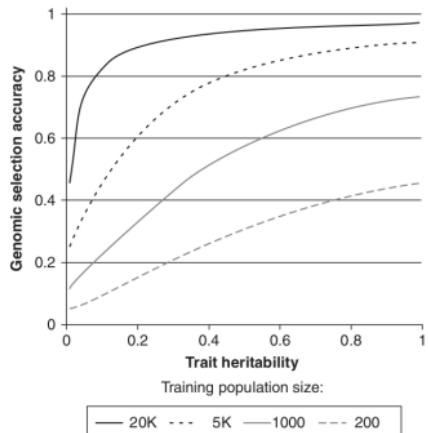
- ↑ genetic distance between TRN and target populations, ↓ accuracy
- Simple genetic architecture has larger accuracy, but faster decay.

Scutari et al., 2016. <https://doi.org/10.1371/journal.pgen.1006288>

Dealing with accuracy

Barley

Heritability vs. Population Size vs. Accuracy



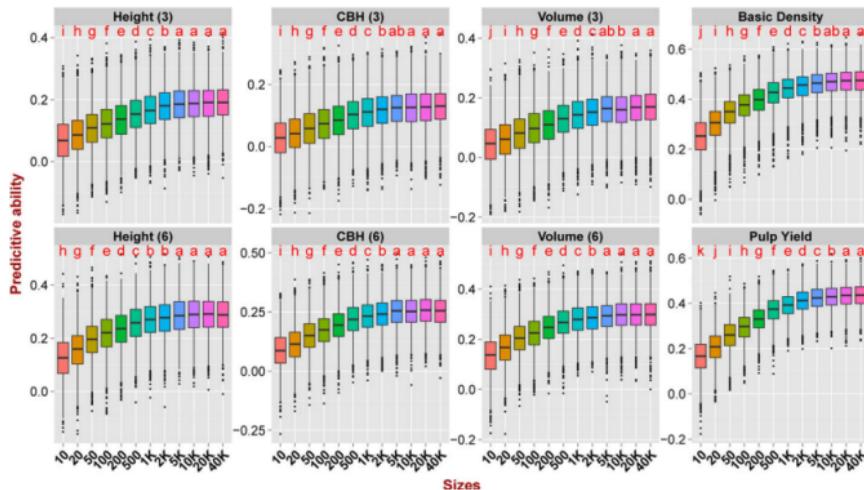
Predicted accuracy of genomic selection assuming the characteristic of elite barley:: effective population size of 50 and genetic map length of 1100 cM.

Hayes et al., 2009. DOI: 10.1017/S0016672308009981

Dealing with accuracy

Eucalyptus

SNP density vs. Accuracy

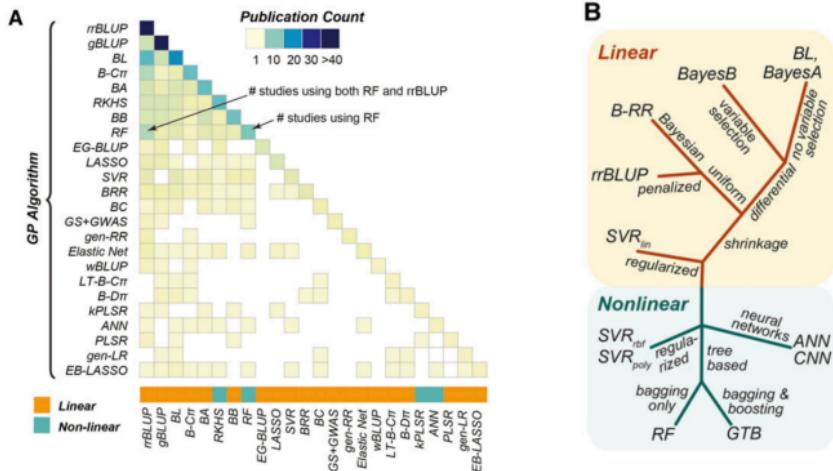


- Depends of LD and h^2 of the trait
- Most of the case, we can see a plateau

Tan et al., 2017. DOI 10.1186/s12870-017-1059-6

Predictive Models

Genomic Selection Models

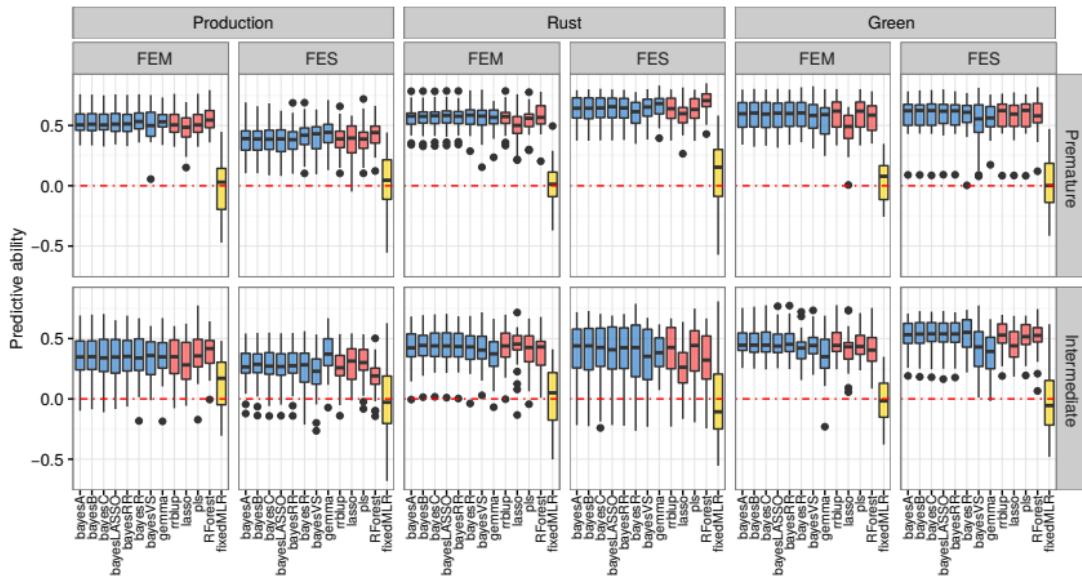


- Different genetics and biological assumptions
- rrBLUP and GBLUP are the most popular

Predictive Models

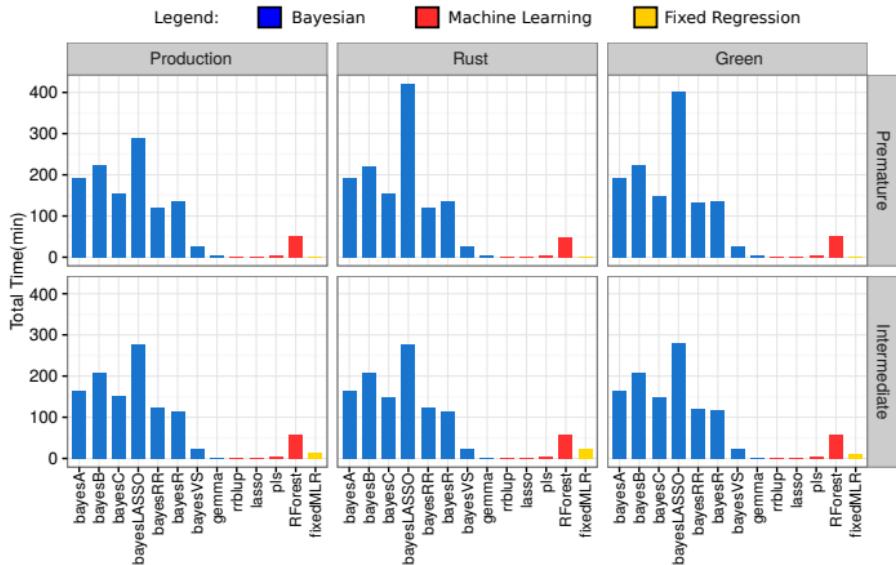
Coffee example

Legend: █ Bayesian █ Machine Learning █ Fixed Regression



Predictive Models

Coffee example



Final Consideration

My personal opinion for practical implementation:

- Genomic selection is a tool to assist breeding and not the other way around
- Have a very good understanding about your crop !!
- Implementation require a solid breeding program
- Genotyping might be not so easy and cheap
- Collect good phenotype is imperative
- Test new methods is valid. But don't forget to include an additive GBLUP as a benchmark
- It is multidisciplinary topic.

References

Other references

-  R2D2 Consortium et al., 2021. Why and How to Switch to Genomic Selection: Lessons From Plant and Animal Breeding Experience. <https://doi.org/10.3389/fgene.2021.629737>
-  Excellence in Breeding Platform. <https://excellenceinbreeding.org/>
-  Ferrão et al., 2021. Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons from Blueberry Breeding. doi: 10.3389/fpls.2021.676326
-  Ferrão et al., 2017. Genomic Selection - State of Art. Book: Genetic Improvement of Tropical Species. In: Hugo Campos and Peter D. S Caligari (Ed.). Springer, 2017
-  Isik et al., 2017. Genetic Data Analysis for Plant and Animal Breeding. Book link

Introduction

○○○○

Breeder Equation

○○○○○

Accuracy

○○○○○○○○○○

Final Consideration

○○●

Thank you !!!

Question ??

About Me:

SCAN ME

