Tools for Polyploids 2024
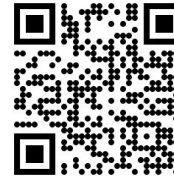
Genomic Prediction in an autotetraploid and outcrossing crop: lessons from blueberry

stay healthy eat blue

Felipe Ferrão

Research Scientist

✉ lferrao@ufl.edu

🐦 lfelipeferrao

San Diego, CA, USA
January, 2024

Background

Who am I?

2020 Research Assistant Scientist
Horticultural Science Department - IFAS

2017-2020 Postdoctoral Associate
Statistical Genetics at the Blueberry Breeding Program

2016 Visiting scholar
Genomic prediction in Coffea canephora using polygenic modeling

2013-2017 Ph.D in Genetics and Plant Breeding
Development and application of statistical genetic methods to genomic prediction in Coffea canephora

2011-2013 M.SC Genetics and Breeding
Microsatellite markers in diversity studies, genetic mapping and QTL analysis in Coffea canephora

2006-2011 Bachelor in Biological Sciences
Department of Statistics
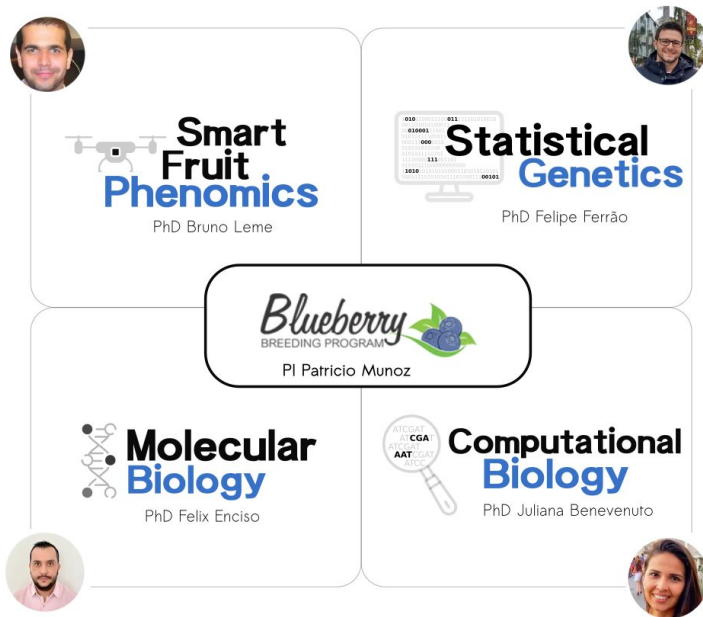Department of Biology

my **education**

Background

The Blueberry Breeding Program

Blueberry Breeder
Associate Professor
Horticultural Science Department
University of Florida
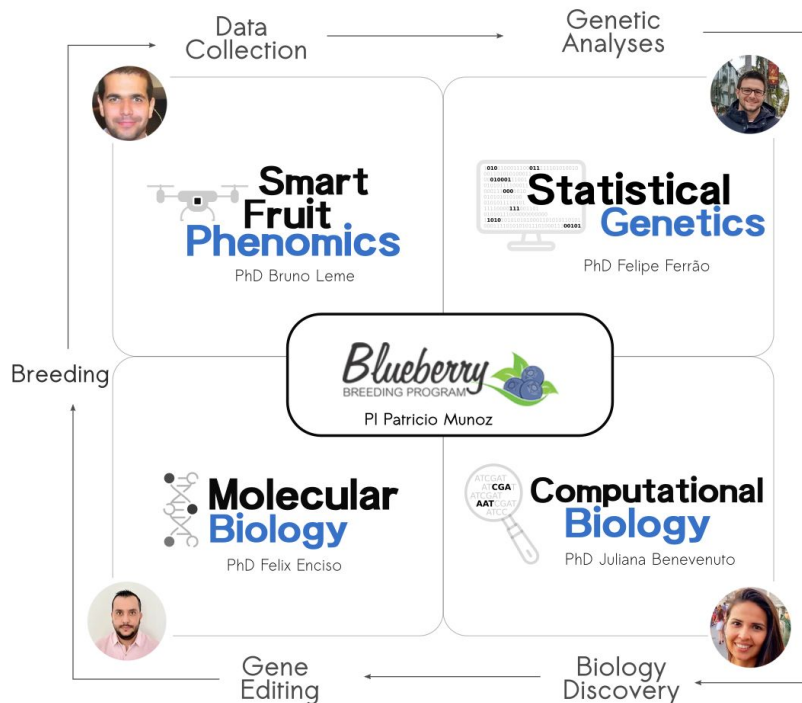
Background

The Blueberry Breeding Program

Blueberry Breeder
Associate Professor
Horticultural Science Department
University of Florida

Phenomics
- Computer vision
- Robots
- Mobile technologies

Statistical Genetics
- Genomic Selection
- Genome-wide association
- Statistical Learning

Molecular Biology
- Transformation
- CRISPR
- Tissue culture

Computational Biology
- Variant discovery
- Genome assembly
- Pan-genome
- RNAseq

Data Collection

Genetic Analyses

Smart Fruit **Phenomics**
PhD Bruno Leme

Statistical Genetics
PhD Felipe Ferrão

Blueberry
BREEDING PROGRAM
PI Patricio Munoz

Breeding

Molecular Biology
PhD Felix Enciso

Computational Biology
PhD Juliana Benevenuto

Gene Editing

Biology Discovery

Background
My research line

Goal

Explore problems at the interface of Statistics and Genetics.

Developing novel methods and software, or learn something new compared with existing approaches.

Background

My research line

Goal

Explore problems at the interface of Statistics and Genetics.

Developing novel methods and software, or learn something new compared with existing approaches.

Motivation

Why genomic selection in blueberry?

Motivation

Importance

- Blueberry is the second most important soft fruit
- Per capita consumption has increased 97% in the past 10 years
- Reason:
  - Delicious !!
  - Health benefits

Motivation

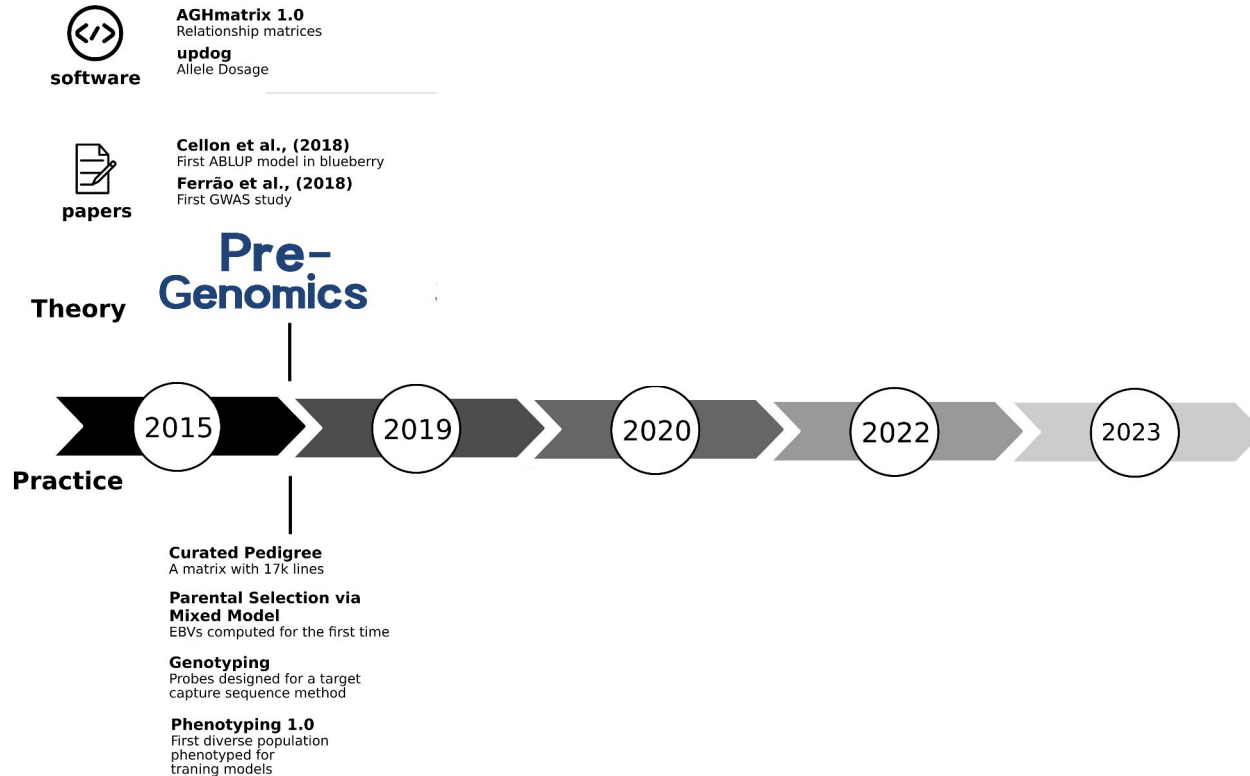Why genomic selection in blueberry?

Importance

- Blueberry is the second most important soft fruit
- Per capita consumption has increased 97% in the past 10 years
- Reason:
  - Delicious !!
  - Health benefits

University of Florida

- UF plays an important role on cultivar development
- Released more than 60 cultivars (from 1970 to 2023)
- PLUS ~45 releases in collaboration or directly from our crosses
- Physolosofy: **breeding** & **research** running side-by-side
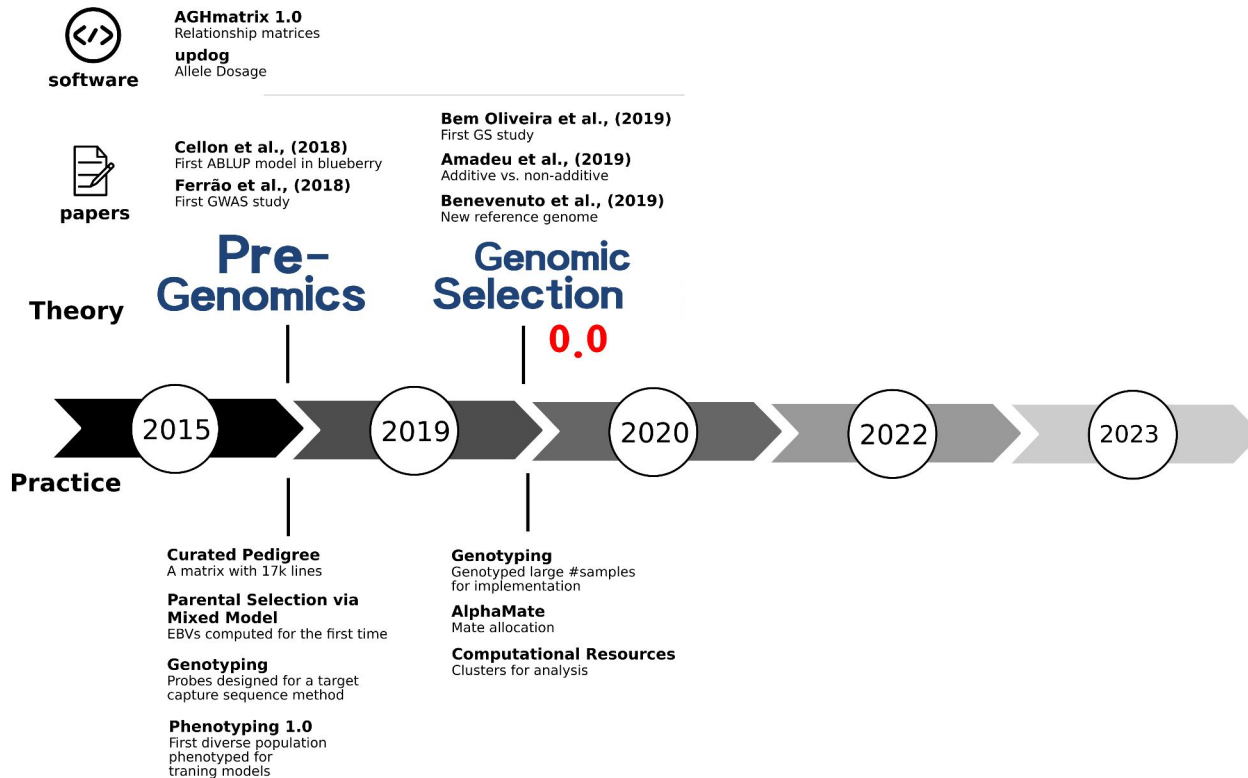
www.blueberrybreeding.com

Timeline
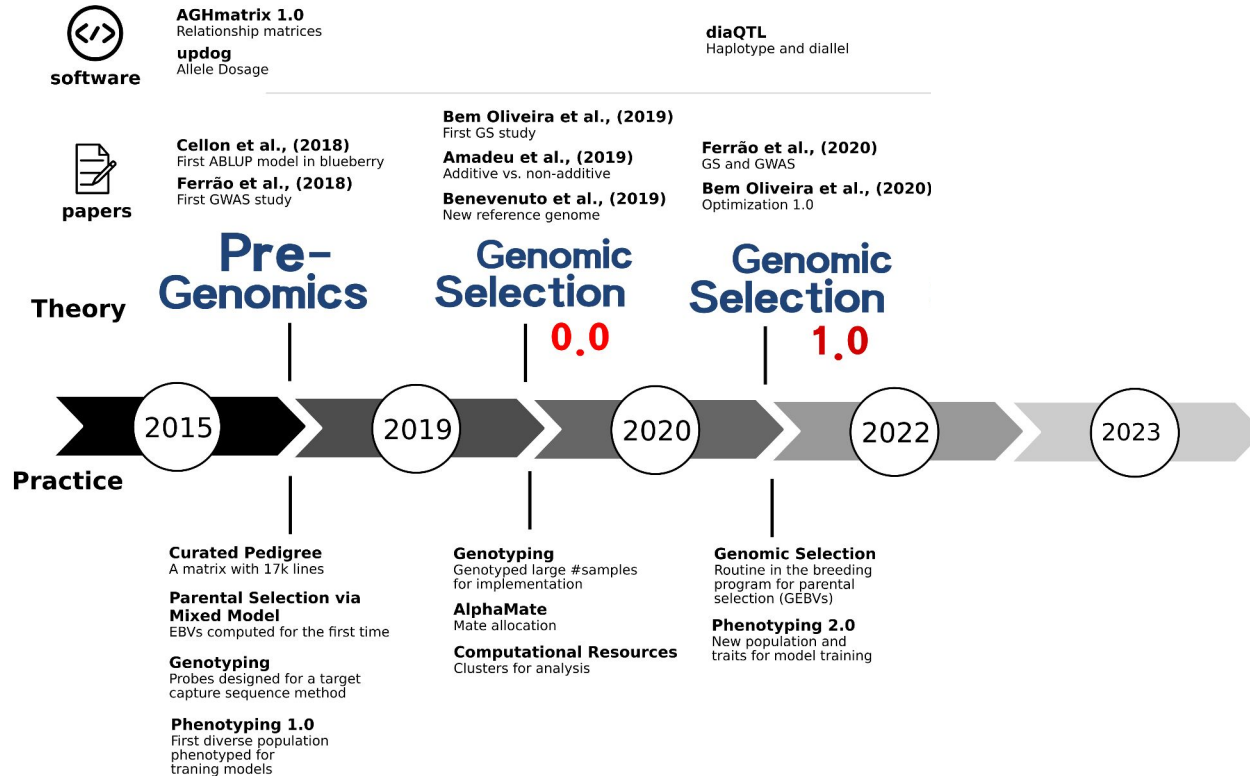
How and when do we start implementing GS?

software

AGHmatrix 1.0
Relationship matrices

updog
Allele Dosage

papers

Cellon et al., (2018)
First ABLUP model in blueberry

Ferrão et al., (2018)
First GWAS study

Theory

Pre-Genomics

Practice

2015 — 2019 — 2020 — 2022 — 2023

Curated Pedigree
A matrix with 17k lines

Parental Selection via Mixed Model
EBVs computed for the first time

Genotyping
Probes designed for a target capture sequence method

Phenotyping 1.0
First diverse population phenotyped for traning models

Timeline

How and when do we start implementing GS?

software

AGHmatrix 1.0
Relationship matrices

updog
Allele Dosage

papers

Cellon et al., (2018)
First ABLUP model in blueberry

Ferrão et al., (2018)
First GWAS study

Bem Oliveira et al., (2019)
First GS study

Amadeu et al., (2019)
Additive vs. non-additive

Benevenuto et al., (2019)
New reference genome

Theory

Pre-Genomics

Genomic Selection

0.0

Practice

| 2015 | 2019 | 2020 | 2022 | 2023 |

Curated Pedigree
A matrix with 17k lines

Parental Selection via Mixed Model
EBVs computed for the first time

Genotyping
Probes designed for a target capture sequence method

Phenotyping 1.0
First diverse population phenotyped for traning models

Genotyping
Genotyped large #samples for implementation

AlphaMate
Mate allocation

Computational Resources
Clusters for analysis

University of Florida
Felipe Ferrão

11

Timeline

How and when do we start implementing GS?

University of Florida
Felipe Ferrão

Timeline

How and when do we start implementing GS?

software

AGHmatrix 1.0
Relationship matrices

updog
Allele Dosage

diaQTL
Haplotype and diallel

berryCV
Fruit phenotyping

AGHmatrix 2.0
Relationship Matrices

SimpleMating
Mate allocation

papers

Cellon et al., (2018)
First ABLUP model in blueberry

Ferrão et al., (2018)
First GWAS study

Bem Oliveira et al., (2019)
First GS study

Amadeu et al., (2019)
Additive vs. non-additive

Benevenuto et al., (2019)
New reference genome

Ferrão et al., (2020)
GS and GWAS

Bem Oliveira et al., (2020)
Optimization 1.0

Ferrrão et al., (2021)
GS overview

Colantonio et al.,(2020)
Metabolomic selection

Theory

Pre-Genomics

Genomic Selection 0.0

Genomic Selection 1.0

Genomic Selection 2.0

2015 — 2019 — 2020 — 2022 — 2023

Practice

Curated Pedigree
A matrix with 17k lines

Parental Selection via Mixed Model
EBVs computed for the first time

Genotyping
Probes designed for a target capture sequence method

Phenotyping 1.0
First diverse population phenotyped for traning models

Genotyping
Genotyped large #samples for implementation

AlphaMate
Mate allocation

Computational Resources
Clusters for analysis

Genomic Selection
Routine in the breeding program for parental selection (GEBVs)

Phenotyping 2.0
New population and traits for model training

Multi-omics
Use of metabolites for flavor prediction

SimpleMating
Internal software for mate allocation

Genomic Selection
Breeding values for parental selection Genetic values for clonal selection

University of Florida
Felipe Ferrão

Genomic Selection 1.0 🫐

Past and present in blueberry

Genomic Selection 1.0

Questions:

- Is genomic selection better than phenotypic selection?
- What are the best predictive models?
- What is the importance of better genomic resources?
- Can we unify prediction and discovery in a single framework?

Genomic Selection 1.0

Questions:

- **Is genomic selection better than phenotypic selection?**
- What are the best predictive models?
- What is the importance of better genomic resources?
- Can we unify prediction and discovery in a single framework?

Molecular Breeding
Felipe Ferrão
Breeding Program

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Background

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	**Single Plant Selection**
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Background

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	**Single Plant Selection**
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Challenges

- Time–consuming (15 years –> cultivar)
- High level of inbreeding depression
- Autotetraploid (2n=4x)

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Molecular Breeding
Felipe Ferrão
Breeding Program

Background

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	**Single Plant Selection**
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Genomic Selection

Selecting crosses in early stages

Skip stages in a breeding cycle

Challenges
- Time–consuming (15 years –> cultivar)
- High level of inbreeding depression
- Autotetraploid (2n=4x)

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

frontiers
in Plant Science

Genomic Selection in an Outcrossing Autotetraploid Fruit Crop: Lessons From Blueberry Breeding

Luís Felipe V. Ferrão[1], Rodrigo R. Amadeu[1], Juliana Benevenuto[1], Ivone de Bem Oliveira[1,2] and Patricio R. Munoz[1]*

[1] Blueberry Breeding and Genomics Lab, Horticultural Sciences Department, University of Florida, Gainesville, FL, United States, [2] Hortifrut North America, Inc., Estero, FL, United States

OPEN ACCESS

Blueberry (Vaccinium corymbosum and hybrids) is a specialty crop with expanding production and consumption worldwide. The blueberry breeding program at the University of Florida (UF) has greatly contributed to expanding production areas by developing low-chilling cultivars better adapted to subtropical and Mediterranean climates of the globe. The breeding program has historically focused on recurrent phenotypic selection. As an autopolyploid, outcrossing, perennial, long juvenile phase crop, blueberry breeding cycles are costly and time consuming, which results in low genetic gains per unit of time. Motivated by applying molecular markers for a more accurate selection in the early stages of breeding, we performed pioneering genomic selection studies and optimization for its implementation in the blueberry breeding program. We have also addressed some complexities of sequence-based genotyping and model parametrization for an autopolyploid crop, providing empirical contributions

G3
Genes | Genomes | Genetics

Genomic Prediction of Autotetraploids; Influence of Relationship Matrices, Allele Dosage, and Continuous Genotyping Calls in Phenotype Prediction

Ivone de Bem Oliveira,*,[†] Marcio F. R. Resende, Jr.,[‡] Luis Felipe V. Ferrão,* Rodrigo R. Amadeu,* Jeffrey B. Endelman,[§] Matias Kirst,** Alexandre S. G. Coelho,[†] and Patricio R. Munoz*,[1]

*Blueberry Breeding and Genomics Lab, and [‡]Sweet Corn Genomics and Breeding, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, [†]Plant Genetics and Genomics Lab, Agronomy College, Federal University of Goias, GO, Brazil, 74690-900, [§]Department of Horticulture, University of Wisconsin, Madison, WI 53706, and **Forest Genomics Lab, School of Forestry Resources and Conservation, University of Florida, Gainesville, FL 32610

ORCID IDs: 0000-0003-3723-9747 (I.d.B.O.); 0000-0002-2367-0766 (M.F.R.R.); 0000-0002-9655-4838 (L.F.V.F.); 0000-0001-5127-4448 (R.R.A.); 0000-0003-0957-4337 (J.B.E.); 0000-0001-8973-9351 (P.R.M.)

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Challenges on Quantitative Genetics analyses applied to polyploid analyses

- More genotypic classes –> allele dosage

- Multisiomic segregation –> relationship matrix

- More complex gene actions –> multiple levels of non–additive effects

Genomic Selection 1.0
Is genomic selection better than phenotypic selection?

Blueberry
Genetic Discovery
& Prediction

Challenges on Quantitative Genetics analyses applied to polyploid analyses

- More genotypic classes –> allele dosage
- Multisiomic segregation –> relationship matrix
- More complex gene actions –> multiple levels of non–additive effects

updog R package
- ~20000 downloads
- Inference on allele dosage,
- Accounting by common features
  of NGS data

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Blueberry
Genetic Discovery & Prediction

Challenges on Quantitative Genetics analyses applied to polyploid analyses

- More genotypic classes –> allele dosage
- Multisiomic segregation –> relationship matrix
- More complex gene actions –> multiple levels of non–additive effects

updog R package

- ~20000 downloads
- Inference on allele dosage,
- Accounting by common features of NGS data

AGHmatrix R package

- Pedigree (**A**), genomic (**G**) and hybrid matrices (**H**)
- Any ploidy level
- Fast and user friendly

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Challenges on Quantitative Genetics analyses applied to polyploid analyses

- More genotypic classes –> allele dosage
- Multisiomic segregation –> relationship matrix
- More complex gene actions –> multiple levels of non–additive effects
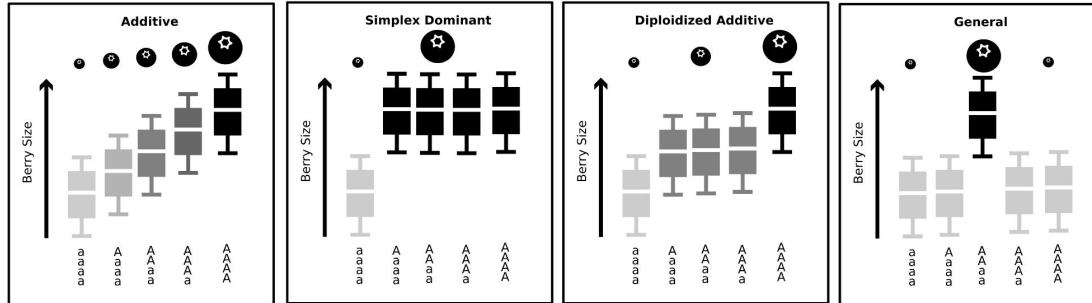
Additive vs. non-additive

- Multiple gene actions
- Different dominance levels

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Back to 2017 …

- **Population:** large population (~2000 ind) representing our breeding collection
- **Phenotype:** fruit quality traits (firmness, size, brix, acidity and weight)
- **Models:** ABLUP (pedigree), G2 (GBLUP assuming diploid markers), G4 (GBLUP accounting for allele dosage)

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Back to 2017 ...

- **Population:** large population (~2000 ind) representing our breeding collection
- **Phenotype:** fruit quality traits (firmness, size, brix, acidity and weight)
- **Models:** ABLUP (pedigree), G2 (GBLUP assuming diploid markers), G4 (GBLUP accounting for allele dosage)

Trait	Matrix	Accuracy	Method
Firmness	A	0.375	Pedigree
Firmness	G2	0.415	GBLUP (2n=2x)
Firmness	G4	**0.426**	GBLUP (2n=4x)
Size	A	0.386	Pedigree
Size	G2	0.400	GBLUP (2n=2x)
Size	G4	**0.432**	GBLUP (2n=4x)

Genomic Selection 1.0

Is genomic selection better than phenotypic selection?

Molecular Breeding
Felipe Ferrão
Breeding Program

Back to 2017 ...

- **Population:** large population (~2000 ind) representing our breeding collection
- **Phenotype:** fruit quality traits (firmness, size, brix, acidity and weight)
- **Models:** ABLUP (pedigree), G2 (GBLUP assuming diploid markers), G4 (GBLUP accounting for allele dosage)

Trait	Matrix	Accuracy	Method
Firmness	A	0.375	Pedigree
Firmness	G2	0.415	GBLUP (2n=2x)
Firmness	G4	**0.426**	GBLUP (2n=4x)
Size	A	0.386	Pedigree
Size	G2	0.400	GBLUP (2n=2x)
Size	G4	**0.432**	GBLUP (2n=4x)

Take-home message

(i) Genomic > Pedigree
(ii) Use polyploid methods

Genomic Selection 1.0

Questions:

- Is genomic selection better than phenotypic selection?
- **What are the best predictive models?**
- What is the importance of better genomic resources?
- Can we unify prediction and discovery in a single framework?
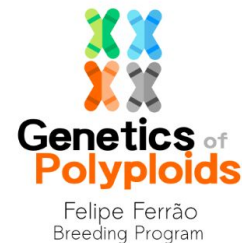
Genomic Selection 1.0

What are the best predictive models?

Impact of Dominance Effects on Autotetraploid Genomic Prediction

Rodrigo R. Amadeu, Luis Felipe V. Ferrão, Ivone de Bem Oliveira, Juliana Benevenuto, Jeffrey B. Endelman, and Patricio R. Munoz*

ABSTRACT
Many commercially important plants are autopolyploid. As a result of the multiple chromosome sets in their genomes, higher orders of allele interactions can occur, implying

R.R. Amadeu, L.F.V. Ferrão, I.D.B. Oliveira, J. Benevenuto, and P.R. Munoz, Blueberry Breeding and Genomics Lab, Horticultural Sciences Dep., Univ. of Florida, Gainesville, FL 32611; J.B. Endelman, Dep. of Horticulture, Univ. of Wisconsin, Madison, WI 53706. Received 28 Feb. 2019. Accepted 22 May 2019. *Corresponding author (p.munoz@ufl.edu). Assigned to Associate Editor Carlos Messina.

Exploring Deep Learning for Complex Trait Genomic Prediction in Polyploid Outcrossing Species

Laura M. Zingaretti[1]*, Salvador Alejandro Gezan[2], Luis Felipe V. Ferrão[3], Luis F. Osorio[4], Amparo Monfort[1,5], Patricio R. Muñoz[3], Vance M. Whitaker[4] and Miguel Pérez-Enciso[1,6]*

[1] Centre for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Campus UAB, Barcelona, Spain, [2] School of Forest Resources and Conservation, University of Florida, Gainesville, FL, United States, [3] Blueberry Breeding and Genomics Lab, Horticultural Sciences Department, University of Florida, Gainesville, FL, United States, [4] IFAS Gulf Coast Research and Education Center, University of Florida, Wimauma, FL, United States, [5] Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Barcelona, Spain, [6] ICREA, Passeig de Lluís Companys 23, Barcelona, Spain

Genomic Selection 1.0
What are the best predictive models?

Background

- Polyploid is complex !
- Most of prediction models are based on linear methods and additive gene actions
- Can we use more elaborate models?
  - Bayesian alphabet + mixed models
  - Different gene actions
  - Deep learning methods that can incorporate non–linearity

Genomic Selection 1.0

What are the best predictive models?

Bayesian vs. Mixed Models vs. Gene actions

- Blueberry (5 traits) vs. Potato (2 traits)

University of Florida
Felipe Ferrão
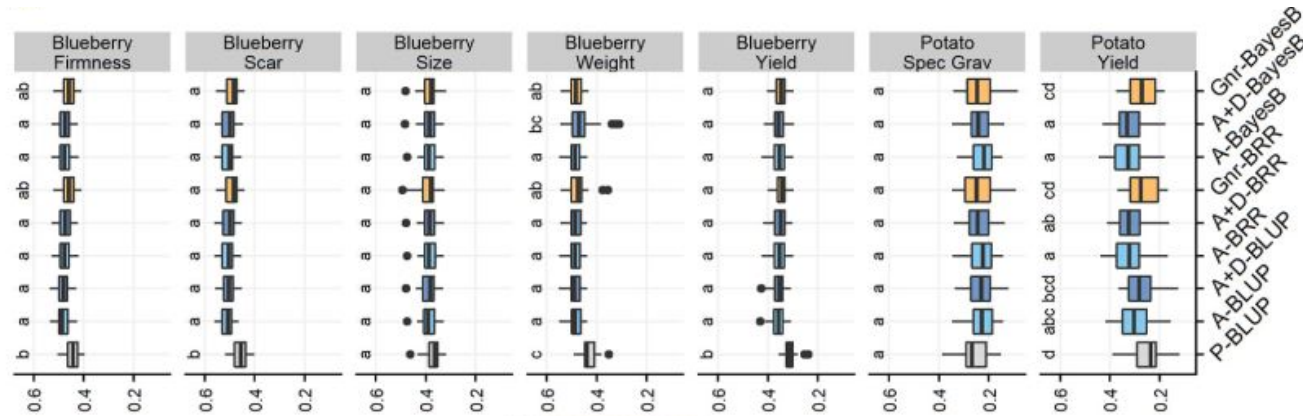
Genomic Selection 1.0

What are the best predictive models?

Bayesian vs. Mixed Models vs. Gene actions

- Blueberry (5 traits) vs. Potato (2 traits)

Take-home message

On the relevance of additive models

University of Florida
Felipe Ferrão

Genomic Selection 1.0

What are the best predictive models?

Bayesian vs. Mixed Models vs. Deep Learning

Genomic Selection 1.0

What are the best predictive models?

Bayesian vs. Mixed Models vs. Deep Learning

B Blueberry
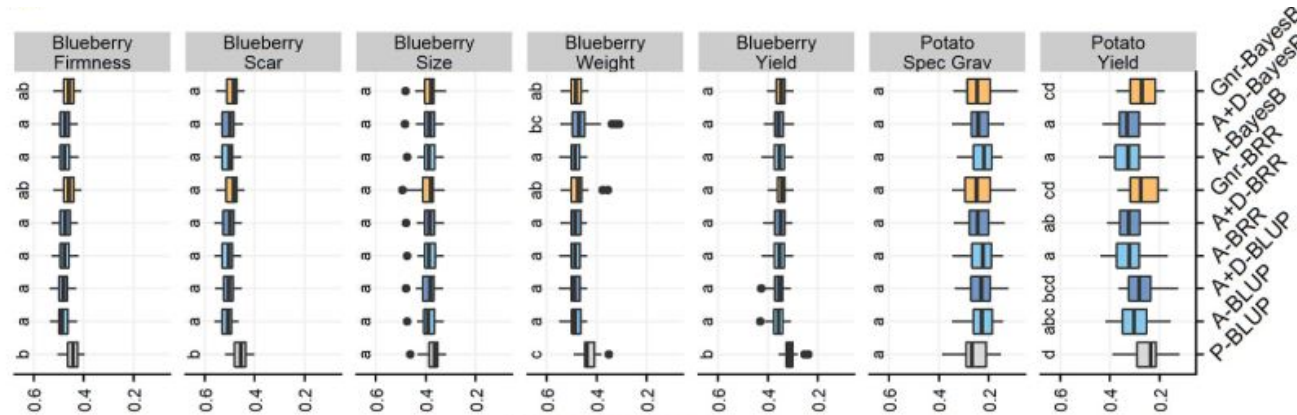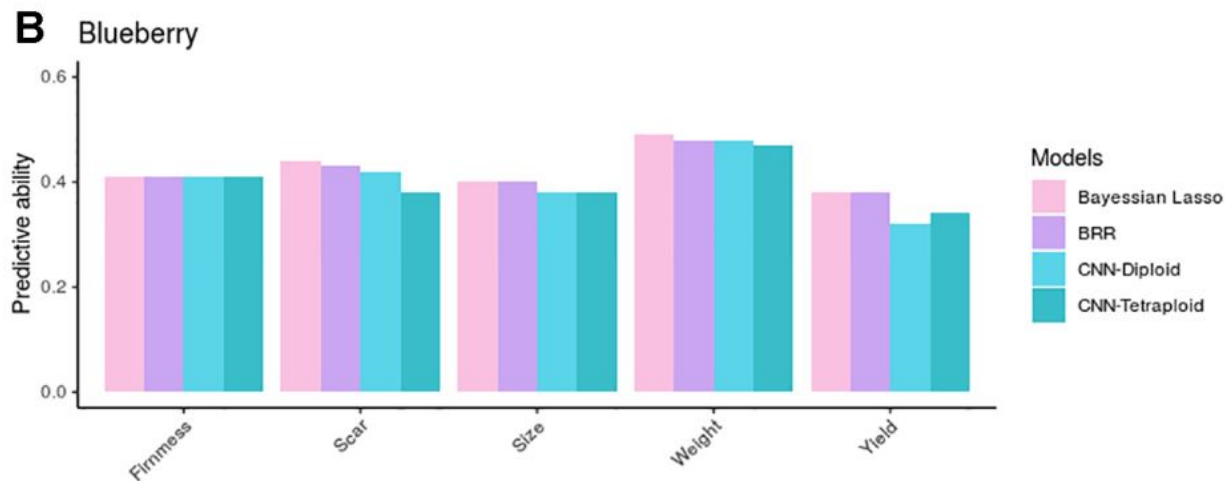
Models
- Bayessian Lasso
- BRR
- CNN-Diploid
- CNN-Tetraploid
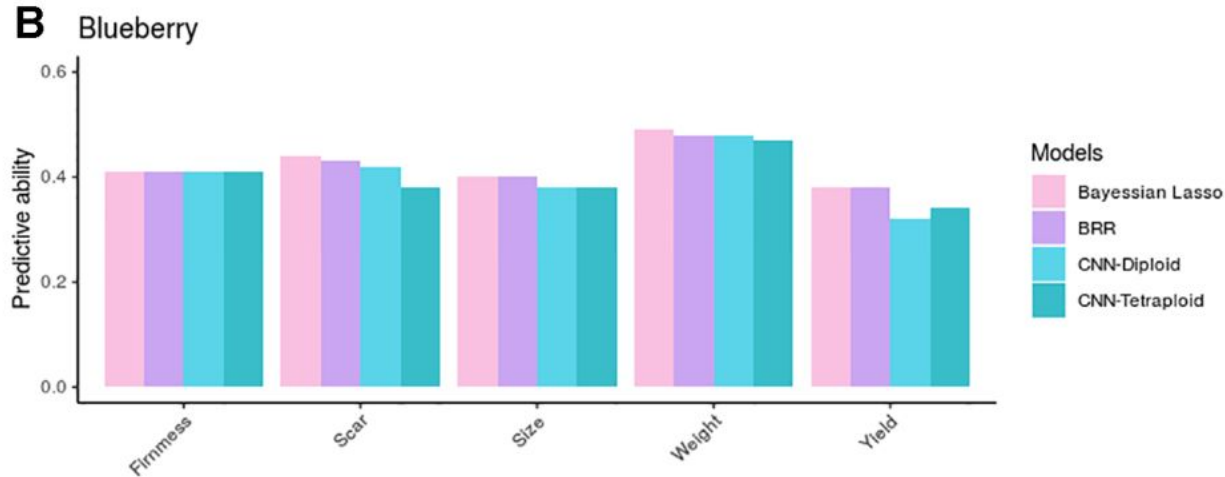
Take home message

On the relevance of additive models[2] !!!

Genomic Selection 1.0

What are the best predictive models?

More recently …

- New paper using a transformer DL architecture
- Similar architecture used by ChatGPT
- The authors tested in multiple data set …
- … and made some strong claims !!
- Ok, let's test it !!

GPFN: Prior-Data Fitted Networks for Genomic Prediction

Jordan Ubbens[1,✉], Ian Stavness[1,2], and Andrew G. Sharpe[1]

[1]Global Institute for Food Security (GIFS), University of Saskatchewan, Saskatoon, SK S7N 0W9, Canada
[2]Department of Computer Science, University of Saskatchewan, Saskatoon, SK S7N 0W9, Canada

Genomic Prediction (GP) methods predict the breeding value of unphenotyped individuals in order to select parental candidates in breeding populations. Among models for GP, classical linear models have remained consistently popular, while more complex nonlinear methods such as deep neural networks have shown comparable accuracy at best. In this work we propose the Genomic Prior-Data Fitted Network (GPFN), a new paradigm for GP. GPFNs perform amortized Bayesian inference by drawing hundreds of thousands or millions of synthetic breeding populations during the prior fitting phase. This allows GPFNs to be deployed without requiring any training or tuning, providing predictions in a single inference pass. On three populations of crop plants across two different crop species, GPFNs perform significantly better than the linear baseline on 13 out of 16 traits. On a challenging between-families NAM prediction task, the GPFN performs significantly better in 3 locations while only falling behind in one. GPFNs represent a completely new direction for the field of genomic prediction, and have the potential to unlock levels of selection accuracy not possible with existing methods.

least as well as newer, more modern methods while being simpler, faster, and requiring less tuning (Azodi et al., 2019; Abdollahi-Arpanahi et al., 2020; Zingaretti et al., 2020; Ubbens et al., 2021; John et al., 2022; Ray et al., 2023).

In this work, we develop a new approach to GP based on amortized Bayesian inference, which we term *Genomic Prior-Data Fitted Networks* (GPFNs). Unlike existing methods based on neural networks, a GPFN is not trained on the end user's dataset, does not require any tuning, and in fact is never exposed to any real data at all prior to inference. As a proof of concept, we show that the GPFN is the first method which is able to consistently and significantly outperform classical methods in several datasets of crop plants. An implementation as well as several trained GPFNs for various population types are provided at https://github.com/jubbens/gpfn. We propose the GPFN approach as a major new direction for genomic prediction.

Genomic Selection 1.0

What are the best predictive models?

Trait	#data points	GPFN	GBLUP (2x)	GBLUP (4x)
pH	~3000	0.31	0.32	0.35
Brix	~1000	0.28	0.27	0.28
Eucalyptol	~1000	0.42	0.37	0.40
Yield	~250	0.33	0.36	–
Maturation	~250	0.11	0.21	–
Vigor	~250	0.08	0.13	–

Camila Azevedo. Deep Learning for Genomic Prediction in Blueberry (work in progress). 2023

University of Florida
Felipe Ferrão

Genomic Selection 1.0

Questions:

- Is genomic selection better than phenotypic selection?
- What are the best predictive models?
- **What is the importance of better genomic resources?**
- Can we unify prediction and discovery in a single framework?

Molecular Breeding
Felipe Ferrão
Breeding Program

Genomic Selection 1.0

What is the importance of better genomic resources?

COMMENTARY

How can a high-quality genome assembly help plant breeders?

Juliana Benevenuto [iD], Luís Felipe V. Ferrão, Rodrigo R. Amadeu and Patricio Munoz [iD]*

Blueberry Breeding and Genomics Laboratory, Horticultural Sciences Department, University of Florida, Gainesville, 2550 Hull Road, FL, USA

University of Florida
Felipe Ferrão

Genomic Selection 1.0

What is the importance of better genomic resources?

Background

- Back to 2019, <u>no genome reference</u> for blueberry.
- Available only a poor draft
- Challenges:
    - Design our genotyping platforms
    - Poor gene annotation for gene mining
    - Complex to design markers for marker assisted selection (MAS)

Genomic Selection 1.0

What is the importance of better genomic resources?

Background
- Back to 2019, <u>no genome reference</u> for blueberry.
- Available only a poor draft
- Challenges:
  - Design our genotyping platforms
  - Poor gene annotation for gene mining
  - Complex to design markers for marker assisted selection (MAS)

Good news:
- Started a collaboration with the Michigan State (Patrick Edger)
- Access to a high–quality and haplotype–phased reference genome (Colle et al., 2019)

University of Florida
Felipe Ferrão

Genomic Selection 1.0

What is the importance of better genomic resources?

Poor Genome

High-quality genome

GWAS
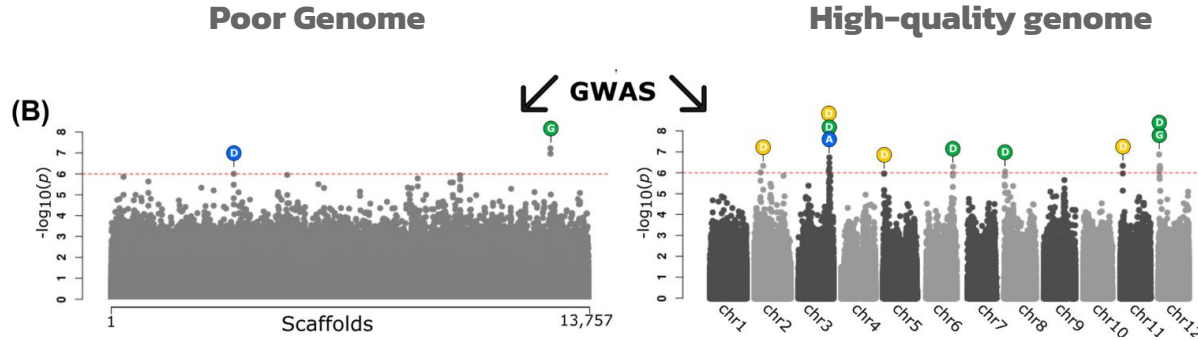
More associations

Traits: ● scar size ● pH ● firmness ● volatile
Gene action: Ⓐ additive Ⓓ dominant Ⓖ general

Genomic Selection 1.0

What is the importance of better genomic resources?
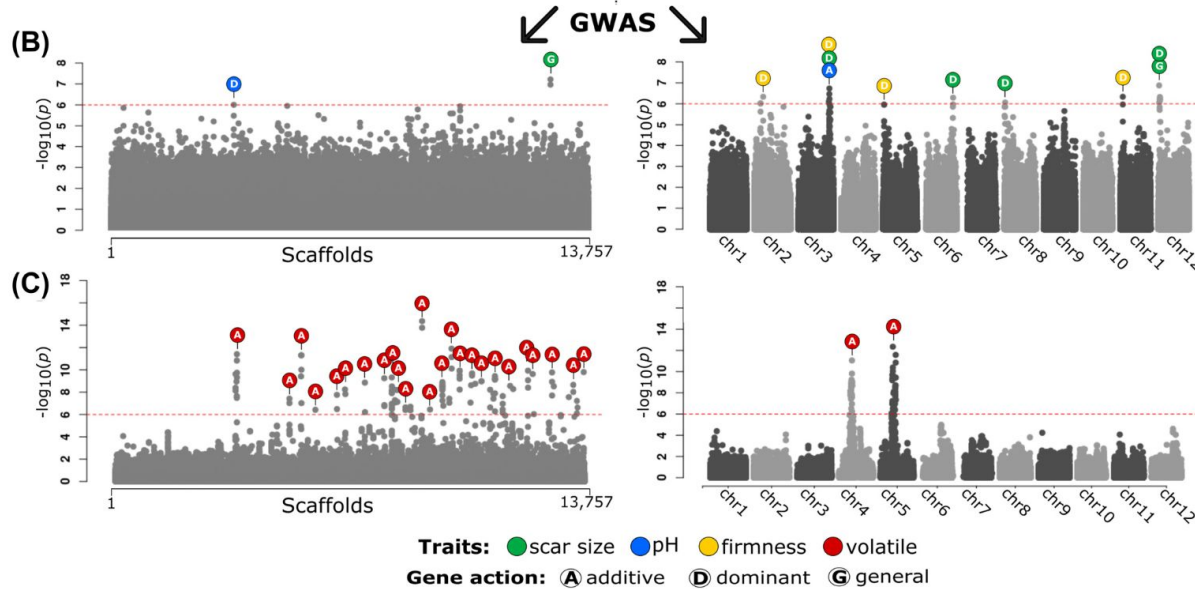
More associations

Better understanding of the genetic architecture

Genomic Selection 1.0

What is the importance of better genomic resources?

Reduced the number of probes (from 30k to 10k) only using bioinformatic

Genomic Selection 1.0

What is the importance of better genomic resources?

Molecular Breeding
Felipe Ferrão
Breeding Program

Reduced the number of probes (from 30k to 10k) only using bioinformatic
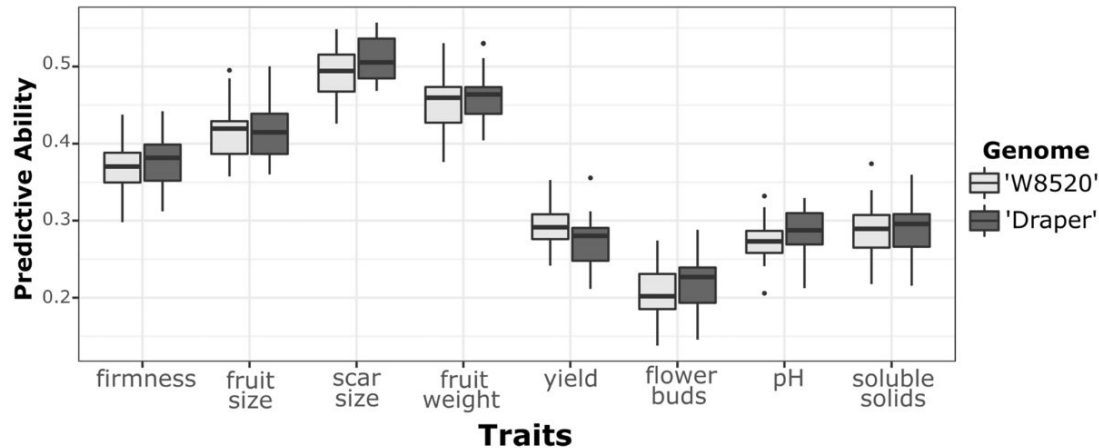
Take home message

Good genomic resources can help plant breeders

Genomic Selection 1.0

Questions:

- Is genomic selection better than phenotypic selection?
- What are the best predictive models?
- What is the importance of better genomic resources?
- **Can we unify prediction and discovery in a single framework?**

Metabolomics for **Flavor**
Felipe Ferrão
Breeding Program

Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?
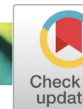
Genome-wide association of volatiles reveals candidate loci for blueberry flavor

Luís Felipe V. Ferrão[1]* iD, Timothy S. Johnson[2]* iD, Juliana Benevenuto[1] iD, Patrick P. Edger[3] iD, Thomas A. Colquhoun[2] iD and Patricio R. Munoz[1] iD

[1]Blueberry Breeding and Genomics Lab, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA; [2]Environmental Horticulture Department, Plant Innovation Center, University of Florida, Gainesville, FL 32611, USA; [3]Department of Horticulture, University of Michigan, Michigan State University, East Lansing, MI 48824, USA

Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?

What is Flavor?

- Flavor is the sum of inputs from multiple senses that inform our brain what we are eating
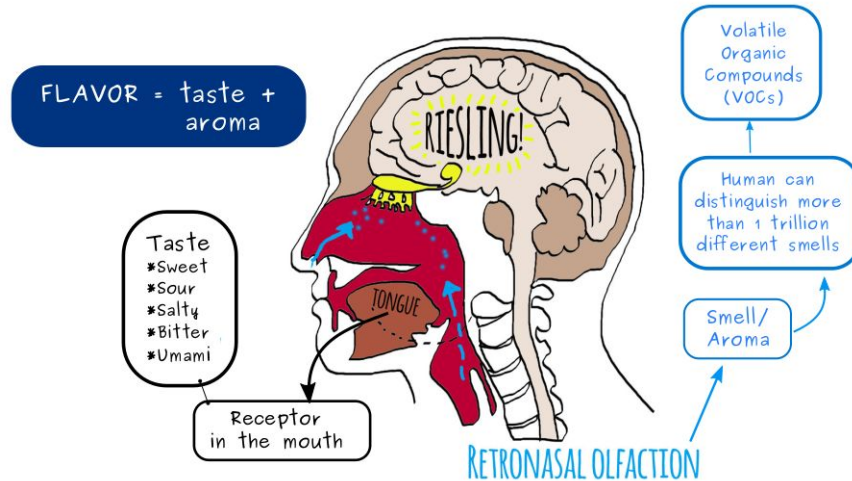
Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?

Evidences based on the GWAS analysis:

- Significant SNPs converging to a tower-like structure
- Single markers explaining large portions of the phenotypic variation
- Hypothesis: VOCs are traits with simple genetic architecture

2-undecanone

University of Florida
Felipe Ferrão

Metabolomics
for Flavor
Felipe Ferrão
Breeding Program
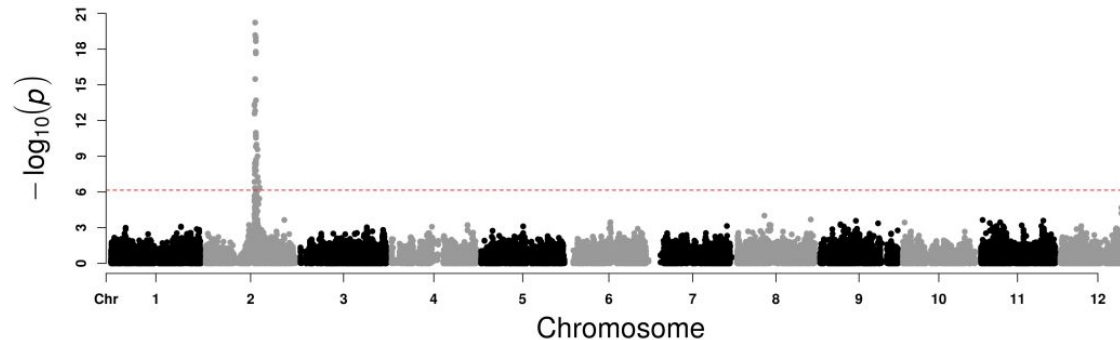
Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?
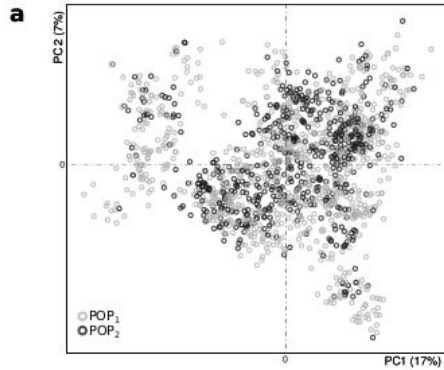
*Original Population
*GWAS
*90 full-sib families

*New Population
*Validation
*Genetic related to POP1

Genomic Selection 1.0
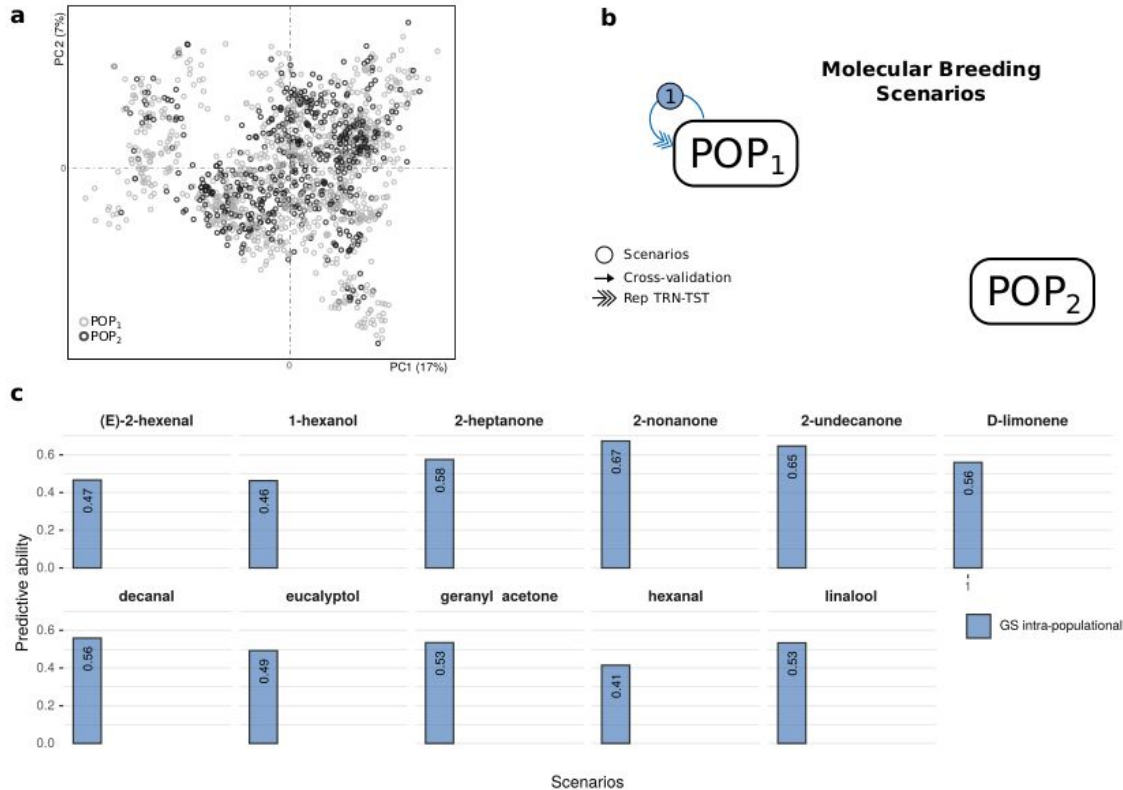
Can we unify prediction and discovery in a single framework?

Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?
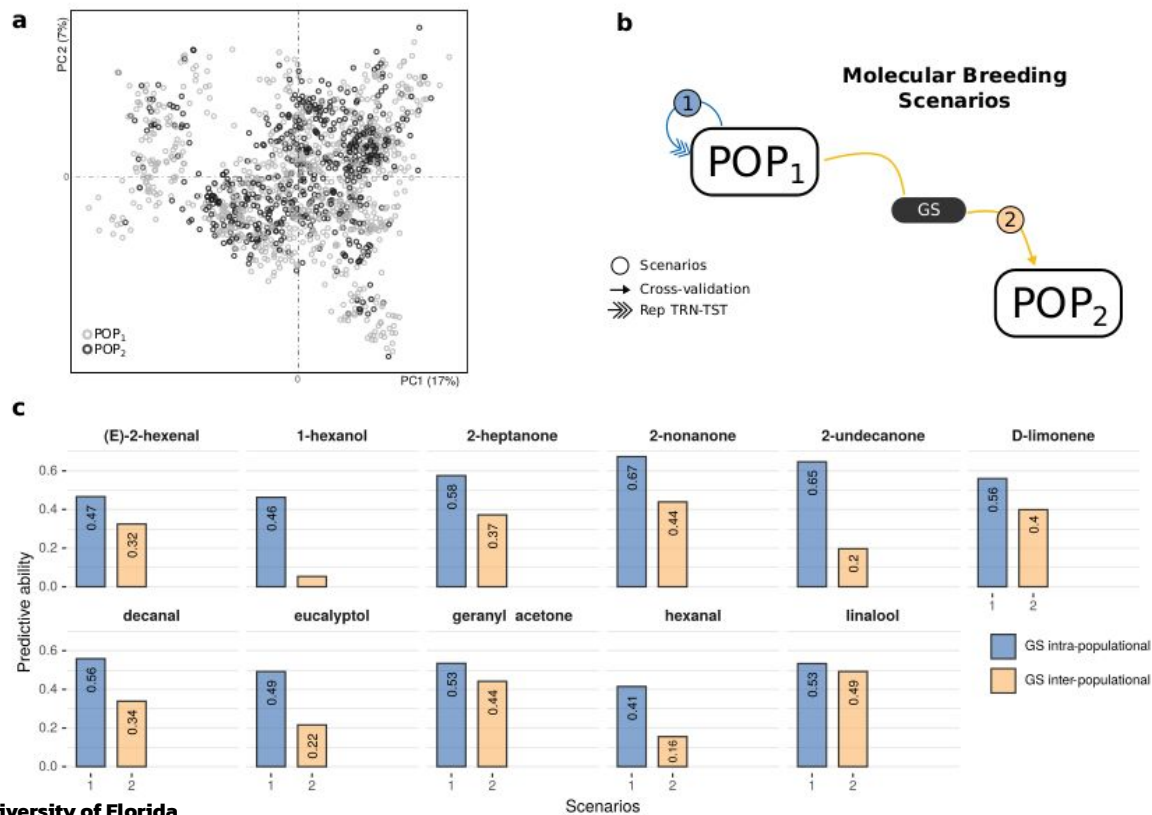
Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?
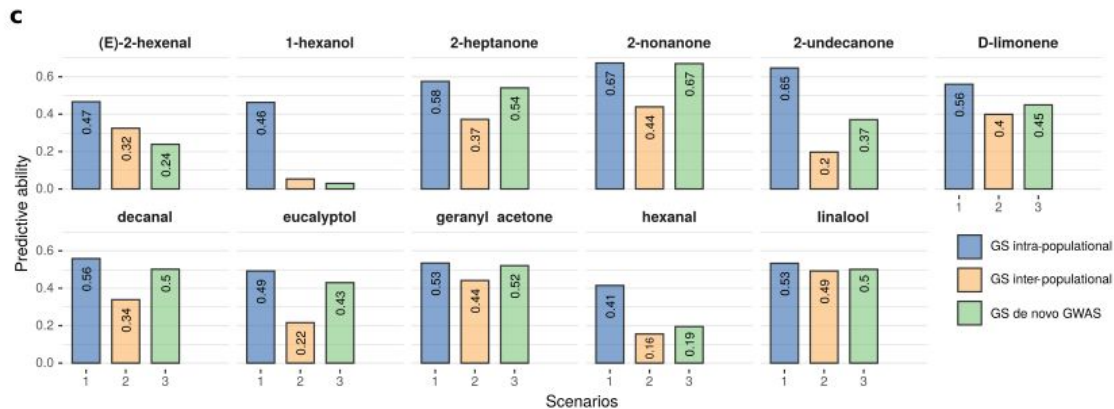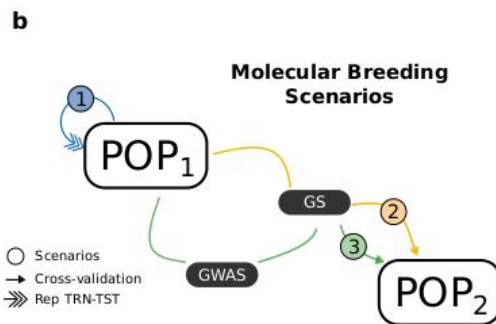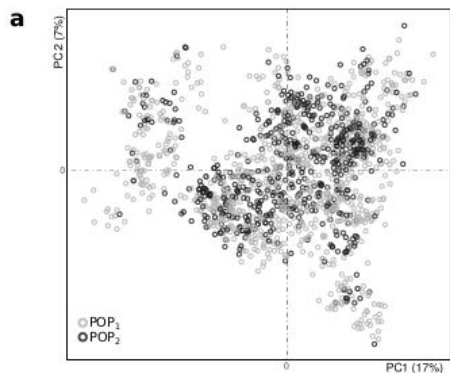
Genomic Selection 1.0

Can we unify prediction and discovery in a single framework?
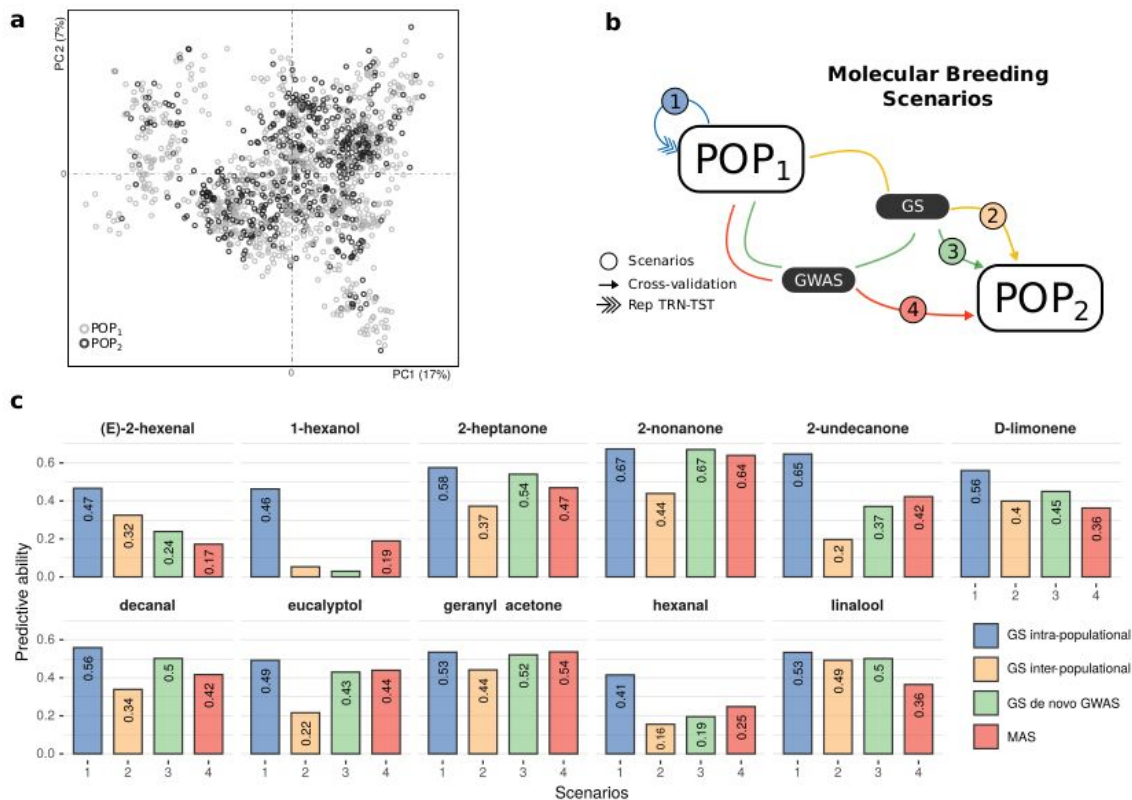
Take home message

Prior biological information can improve prediction ability

Genomic Selection 2.0 🫐

Present and future in blueberry

What is the future?

New methods

New traits

Metabolomic Selection

Artificial Intelligence

Genomic Selection 2.0

Questions

- Can we reduce the number of markers and optimize our training set?
- Is multi–omic predictions a good alternative?
- What is the impact of phenomics on fruit quality prediction?
- How AI can shape the future of modern breeding programs?

Phenomic Selection
Paul Adunola
PhD project

Metabolomics for Flavor
Felipe Ferrão
Breeding Program

Optimization
Simulation • Mate Allocation
Camila Azevedo
Breeding Program

MAS
marker assisted selection
Jim-hee
Breeding Program

Artificial Intelligence
Bruno Leme
Breeding Program

Computer Vision
Bruno Leme
Breeding Program

Genomic Selection 2.0

Questions

- **Can we reduce the number of markers and optimize our training se**
- Is multi–omic predictions a good alternative?
- What is the impact of phenomics on fruit quality prediction?
- How AI can shape the future of modern breeding programs?

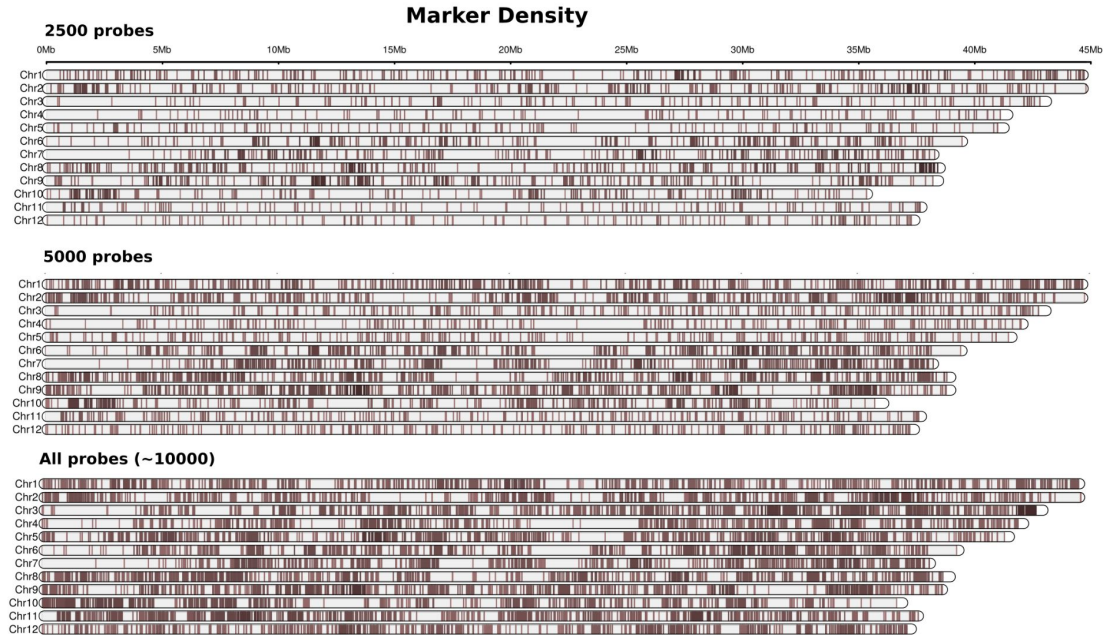Optimization
Simulation • Mate Allocation

Camila Azevedo
Breeding Program

Genomic Selection 2.0

How low can we go?

Optimization
Simulation • Mate Allocation

Camila Azevedo
Breeding Program
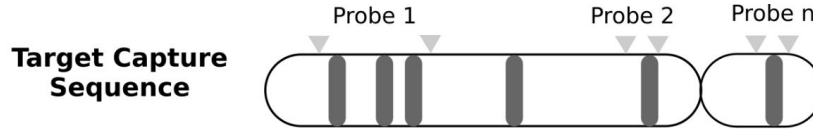
Target Capture Sequencing

- 30k probes originally designed in 2013 (Ferrão et al., 2018)
- 10k probes redefined after the reference genome (Benevenuto et al., 2019)
- Probes are targeting genomic regions aligning to the 4 homologous, and well distributed
- Different #SNPs per probe, and quality parameters
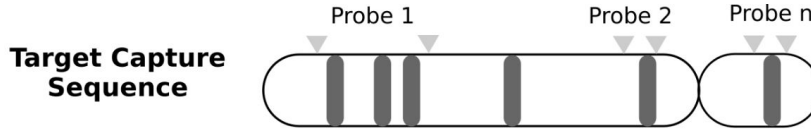
Adunola et al., 2023. Optimizing Genomic Selection in Blueberry. In preparation

58

Genomic Selection 2.0

How low can we go?
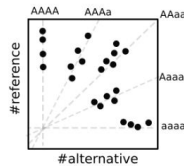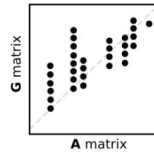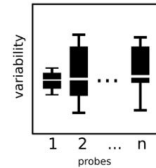
- Data–driven methods to select a final number of probes
- It includes quality and genetic information
- We used a selection index to weight all these information

Adunola et al., 2023. Optimizing Genomic Selection in Blueberry. In preparation

Genomic Selection 2.0

How low can we go?

Marker Density

2500 probes

5000 probes

All probes (~10000)

Camila Azevedo
Breeding Program

Adunola et al., 2023. Optimizing Genomic Selection in Blueberry. In preparation

Genomic Selection 2.0

How low can we go?

Can we keep reducing the number of probes?

# probes	Brix	Firmness	TTA
All (10k)	0.31	0.50	0.42
2500	0.32	0.48	0.40
5000	0.32	0.49	0.41

- All prediction abilities computed in cross–validation
- Probes are regions in the genome where SNPs will be mapped
- 2500 probes (~25% of the total costs) results in good predictive ability for multiple traits

University of Florida
Felipe Ferrão

Genomic Selection 2.0

How low can we go?

What about the TRN population ?

Is my old model still calibrated?

TRN

TST

Population: ~2000 ind
5 Fruit Quality Traits
Model trained: 2014 & 2015
Blueberry Breeding Program

Population: ~1000 ind
5 Fruit Quality Traits
Model trained: 2020 & 202
VacCap Population

University of Florida
Felipe Ferrão

Genomic Selection 2.0
How low can we go?

Optimization
Simulation • Mate Allocation

Camila Azevedo
Breeding Program

What about the TRN population ?

Is my old model still calibrated?

TRN

Population: ~2000 ind
5 Fruit Quality Traits
Model trained: 2014 & 2015
Blueberry Breeding Program

TST

Population: ~1000 ind
5 Fruit Quality Traits
Model trained: 2020 & 202
VacCap Population

TRN

How many new ind should I add?
How to select them?

Genomic Selection 2.0

How low can we go?

Predictive ability optimized using different training population designs

University of Florida
Felipe Ferrão

Adunola et al., 2023. Optimizing Genomic Selection in Blueberry. In preparation

Genomic Selection 2.0

How low can we go?

Predictive ability optimized using different training population designs

Take home message

We can reduce number of markers and optimize our TRN population to maximize the accuracy

Genomic Selection 2.0

Questions

- Can we reduce the number of markers and optimize our training set
- **Is multi-omic predictions a good alternative?**
- What is the impact of phenomics on fruit quality prediction?
- How AI can shape the future of modern breeding programs?

Metabolomics for Flavor

Felipe Ferrão
Breeding Program

Genomic Selection 2.0
Metabolomic Selection

Metabolomic selection for enhanced fruit flavor

Vincent Colantonio[a,1], Luis Felipe V. Ferrão[a,1], Denise M. Tieman[a], Nikolay Bliznyuk[b,c,d], Charles Sims[e], Harry J. Klee[a,2], Patricio Munoz[a,2], and Marcio F. R. Resende Jr.[a,2]

[a]Horticultural Sciences Department, University of Florida, Gainesville, FL 32611; [b]Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL 32611; [c]Department of Biostatistics, University of Florida, Gainesville, FL 32611; [d]Department of Statistics, University of Florida, Gainesville, FL 32611; and [e]Food Science and Human Nutrition Department, University of Florida, Gainesville, FL 32611

Although they are staple foods in cuisines globally, many commercial fruit varieties have become progressively less flavorful over time. Due to the cost and difficulty associated with flavor phenotyping, breeding programs have long been challenged in selecting for this complex trait. To address this issue, we leveraged targeted metabolomics of diverse tomato and blueberry accessions and their corresponding consumer panel ratings to create statistical and machine learning models that can predict sensory perceptions of fruit flavor. Using these models, a breeding program can assess flavor ratings for a large number of genotypes, previously limited by the low throughput of consumer sensory panels. The ability to

program. The difficulties associated with accurate flavor phenotyping have contributed to the lack of selection for fruit flavor and thereby contributed to the widespread consumer belief that commercial fruit flavor has declined (6, 7). Cheap and scalable flavor selection methods would greatly benefit the breeding process.

The main driver of fruit flavor perception is its chemical composition. Fruits contain a diverse array of sugars, acids, and volatiles whose concentrations are driven by genetic and environmental effects. Sugars and acids are largely perceived by taste receptors on the tongue and the volatiles by receptors
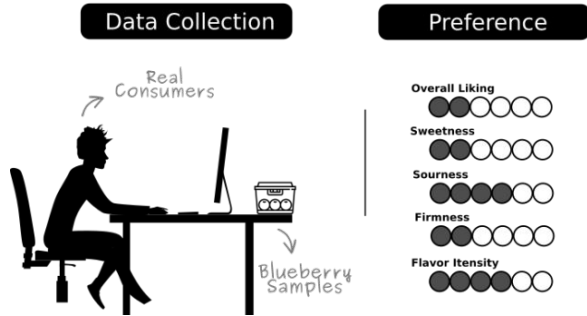
Genomic Selection 2.0
Metabolomic Selection

How to evaluate flavor?

Metabolomics for **Flavor**

Felipe Ferrão
Breeding Program

1 Sensory Panels

Data Collection | Preference
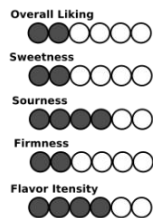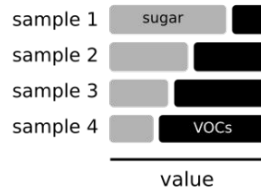
- Most accurate way to evaluate flavor preference is by providing consumers with a sample set of diverse food and quantify their opinion
- Limitations:
  - Expensive
  - Time-consuming
  - Low throughput

Genomic Selection 2.0

Metabolomic Selection

Genomic Selection 2.0
Metabolomic Selection

Metabolomics
for Flavor

Felipe Ferrão
Breeding Program
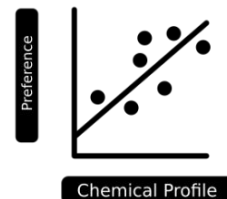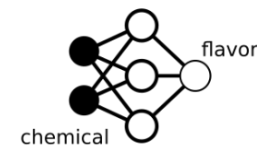
1) Organized Historical Data
Tomato and Blueberry UF Breeding Programs

|	Tomato	Blueberry
	209 samples	244 samples
	34 Taste Panels (2010-2015)	53 Taste Panels (2012-2017)
	2844 panelists	3763 panelists
	68 VOCs + 6 Sugars/Acids	49 VOCs + 6 Sugars/Acids
	Liking, Texture, Sweetnes, Sourness, Intensity, Salty, Bitter, Umami	Liking, Texture, Sweetness, Sourness, Intensity

Sensory ~ f(metabolites)

Can we predict flavor preference?

- We tested different machine learning and statistical approaches for prediction
- Historical sensory and chemical data
- Tomato and blueberry as our biological models
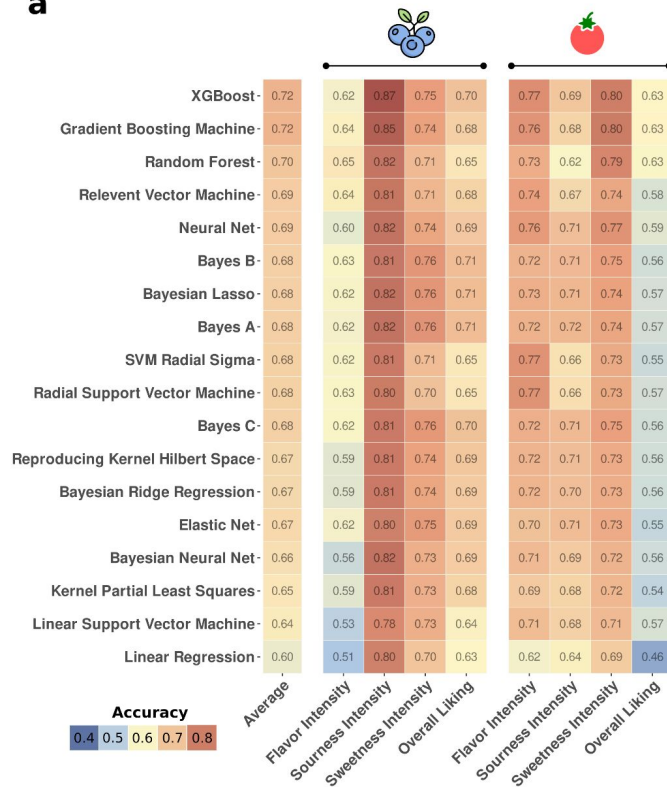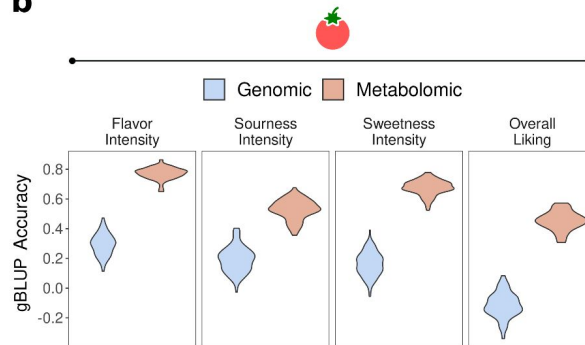
Genomic Selection 2.0
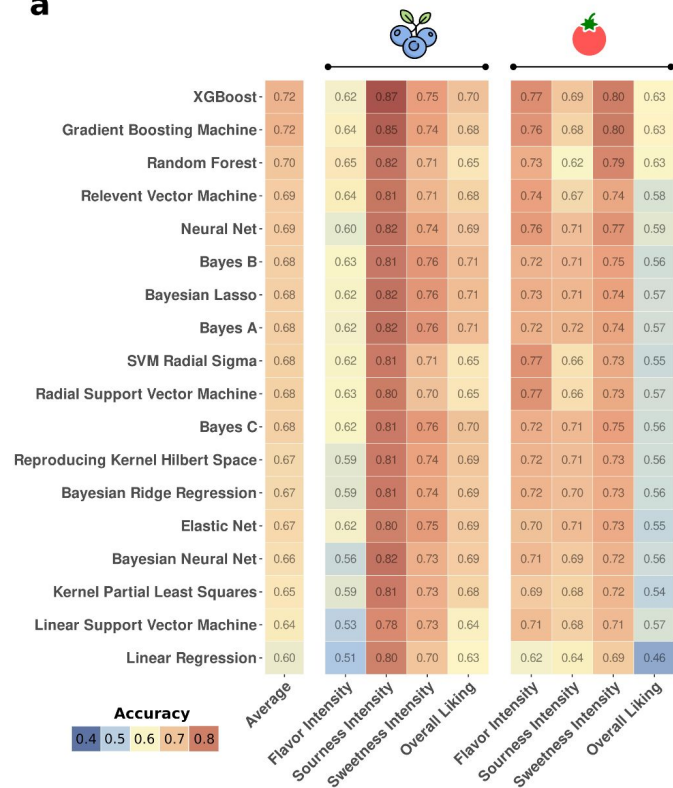
Genomic Selection 2.0

a

b

Metabolomic selection >> Genomic Selection

Felipe Ferrão
Breeding Program

Genomic Selection 2.0

Genomic Selection 2.0
Metabolomic Selection

2022 Blueberry Example

- 1060 blueberry samples
- 60 volatiles, 5 FQ traits and ~50K snps
- Sensory: aroma (1–5) and liking (1–5)
- Two breeders evaluated sensory traits
- Multi–Kernel mixed model

Cazorso., 2024. A multi–omic roadmap for flavor prediction in blueberries. In preparation

Genomic Selection 2.0
Metabolomic Selection

2022 Blueberry Example

- 1060 blueberry samples
- 60 volatiles, 5 FQ traits and ~50K snps
- Sensory: aroma (1–5) and liking (1–5)
- Two breeders evaluated sensory traits
- Multi–Kernel mixed model

Better predictive abilities when multiple source of information are combined

Felipe Ferrão
Breeding Program

Genomic Selection 2.0
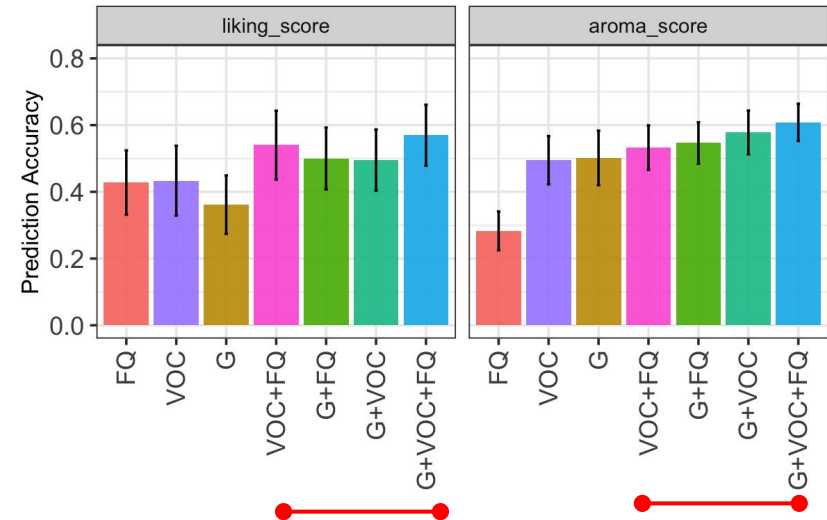## Metabolomic Selection

Felipe Ferrão
Breeding Program

2022 Blueberry Example

- 1060 blueberry samples
- 60 volatiles, 5 FQ traits and ~50K snps
- Sensory: aroma (1–5) and liking (1–5)
- Two breeders evaluated sensory traits
- Multi–Kernel mixed model

Take home message

Multi–omic data is a valid tool for predicting complex traits
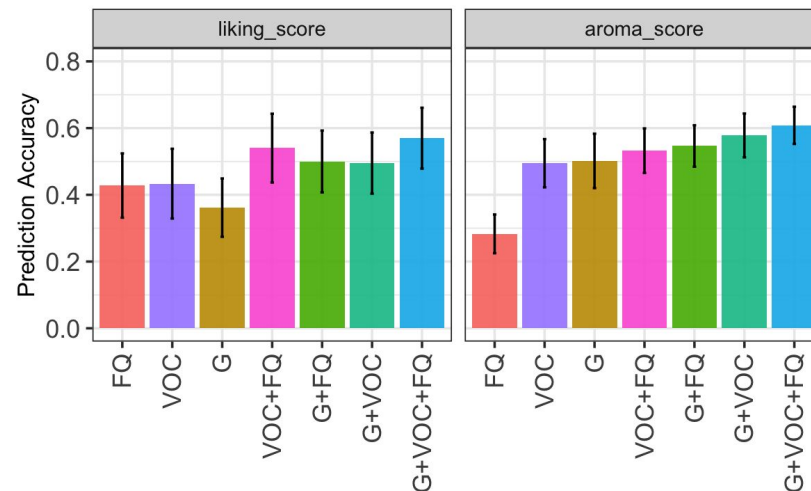
Cazorso., 2024. A multi–omic roadmap for flavor prediction in blueberries. In preparation

University of Florida
Felipe Ferrão

Genomic Selection 2.0

Questions

- Can we reduce the number of markers and optimize our training set?
- Is multi–omic predictions a good alternative?
- **What is the impact of phenomics on fruit quality prediction?**
- How AI can shape the future of modern breeding programs?

Phenomic Selection

Paul Adunola
PhD project

Genomic Selection 2.0
Phenomic Selection

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	Single Plant Selection
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Genomic Selection 2.0
Phenomic Selection

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	Single Plant Selection
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Selection Criteria

Visual

GEBV

GEBV + metabolites + phenotypic data

GEBV + metabolites + phenotypic data + sensory

Genomic Selection 2.0
Phenomic Selection

Field year	Stage	# Plants	Goal
0	0	~150 crosses	Crossing + Seedlings
1	I	20.000	High–density nursery– Single Plant Selection
2–4	II	2.000	Single Plant Selection
5–9	III	200	Farm Condition – Experimental design
10–15	IV	20	Regional Yield Trial – Experimental design
16	V	1–2	Cultivar Release

Selection Criteria

Visual

GEBV

GEBV + metabolites + phenotypic data

GEBV + metabolites + phenotypic data + sensory

University of Florida
Felipe Ferrão

Genomic Selection 2.0
Phenomic Selection

Motivation Phenomic Selection

- Near–infrared spectroscopy (NIRS) is a non–destructive high–throughput method
- It is based on the absorption of electromagnetic radiation in the near–infrared region
- While NIR has been used to predict target traits, recent studies suggested phenomic selection as a low–cost and high–throughput method

Genomic Selection 2.0
Phenomic Selection

Motivation Phenomic Selection

- Near–infrared spectroscopy (NIRS) is a non–destructive high–throughput method
- It is based on the absorption of electromagnetic radiation in the near–infrared region
- While NIR has been used to predict target traits, recent studies suggested phenomic selection as a low–cost and high–throughput method
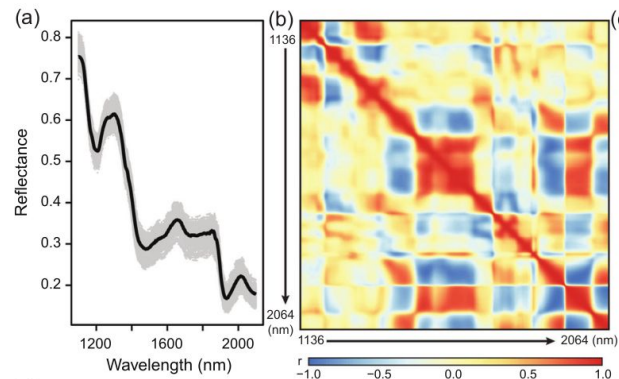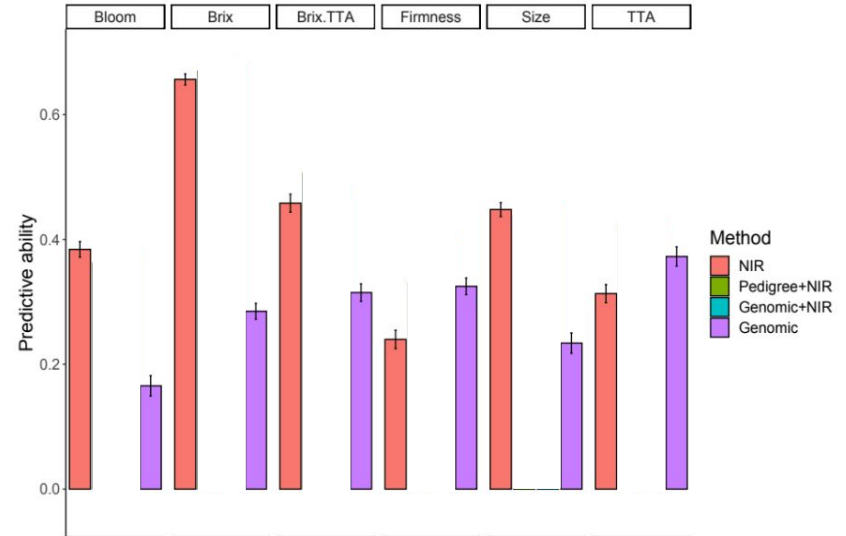
Example in soybean

DOI: 10.1002/ppj2.20027
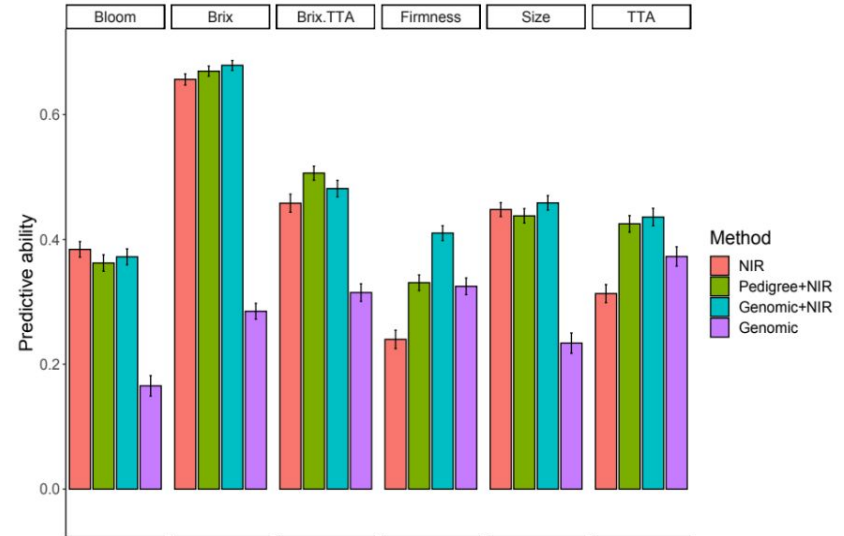
Genomic Selection 2.0
Phenomic Selection

Material and Methods

- Population Size: ~400 ind
- MicroNIR OnSite–W kit: 900–1700 wavelength
- Tissue: mature berries
- Model: P–GBLUP (multi kernel mixed model)

Adunola and Tavares et al., 2023. Phenomic Selection in Blueberry. In preparation

83

Genomic Selection 2.0

Phenomic Selection

Paul Adunola
PhD project

Material and Methods

- Population Size: ~400 ind
- MicroNIR OnSite–W kit: 900–1700 wavelength
- Tissue: mature berries
- Model: P–GBLUP (multi kernel mixed model)

Be careful...

- Very sensitive to GxE
- Very sensitive to tissue (berry vs. fruit vs. juice)
- Working in progress

University of Florida
Felipe Ferrão

Adunola and Tavares et al., 2023. Phenomic Selection in Blueberry. In preparation

Genomic Selection 2.0
Phenomic Selection

Paul Adunola
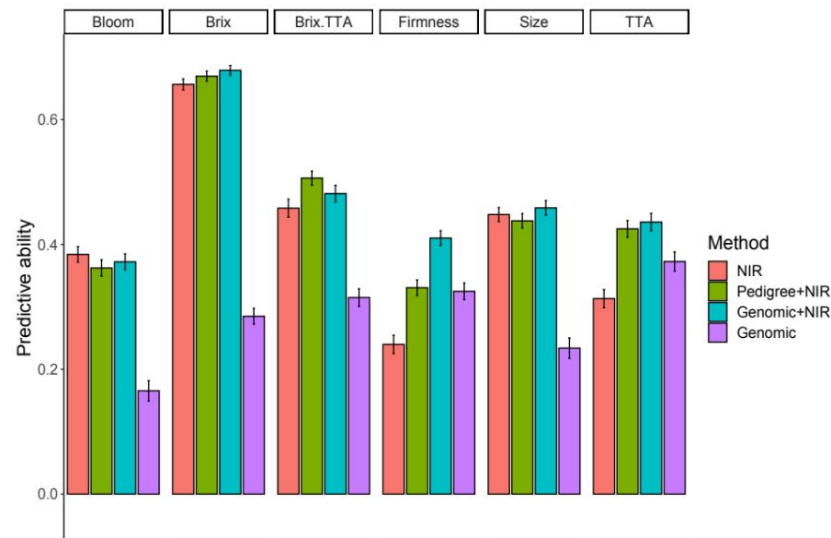PhD project

Material and Methods

- Population Size: ~400 ind
- MicroNIR OnSite–W kit: 900–1700 wavelength
- Tissue: mature berries
- Model: P–GBLUP (multi kernel mixed model)

Be careful...

- Very sensitive to GxE
- Very sensitive to tissue (berry vs. fruit vs. juice)
- Working in progress

Take home message

An alternative to genomic selection

University of Florida
Felipe Ferrão

Adunola and Tavares et al., 2023. Phenomic Selection in Blueberry. In preparation
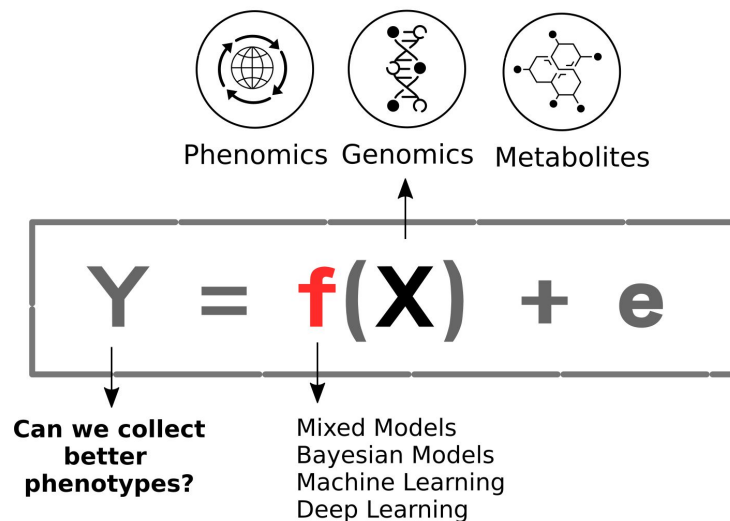
Genomic Selection 2.0

Questions

- Can we reduce the number of markers and optimize our training
- Is multi–omic predictions a good alternative?
- What is the impact of phenomics on fruit quality prediction?
- **How AI can shape the future of modern breeding programs?**

Artificial Intelligence
Bruno Leme
Breeding Program

Computer Vision
Bruno Leme
Breeding Program

Genomic Selection 2.0

Artificial Intelligence

Phenomics   Genomics   Metabolites

$$Y = f(X) + e$$

Can we collect better phenotypes?

Mixed Models
Bayesian Models
Machine Learning
Deep Learning
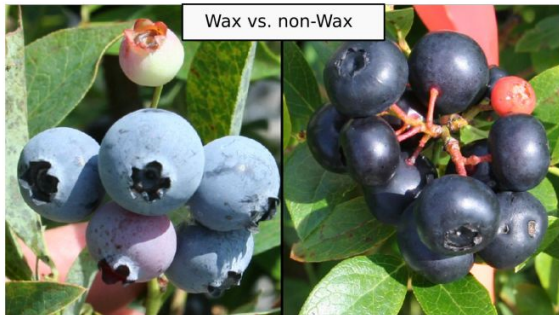
Genomic Selection 2.0

Artificial Intelligence

Motivation

Qi et al., 2019. doi.org/10.1186/s12870-019-2073-7

- The naturally occurring cuticular wax covering the fruit (a.k.a bloom) is what gives the blueberries their whitish looking
- Important for consumers, post–harvest and disease resistance
- Trait traditionally scored using visual scales

Genomic Selection 2.0
Artificial Intelligence

Artificial Intelligence
Bruno Leme
Breeding Program

Computer Vision
Bruno Leme
Breeding Program

Motivation

Wax vs. non-Wax

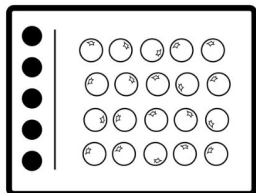Qi et al., 2019. doi.org/10.1186/s12870-019-2073-7

- The naturally occurring cuticular wax covering the fruit (a.k.a bloom) is what gives the blueberries their whitish looking
- Important for consumers, post–harvest and disease resistance
- Trait traditionally scored using visual scales
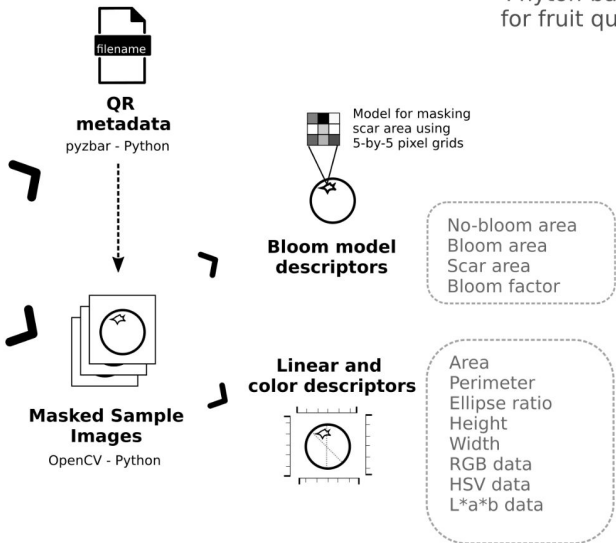
Genomic Selection 2.0
Artificial Intelligence

Ferrão et al., 2024. Computer vision tool integrated to genome–wide association identifies candidate genetic loci controlling waxy bloom in blueberries. In preparation
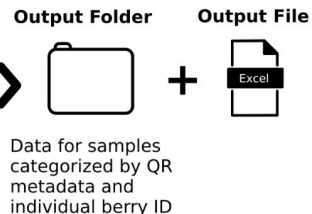
Genomic Selection 2.0
Artificial Intelligence

berryCV vs. Traditional Visual scores

- Across different time–points during the post–harvest
- Higher heritability values
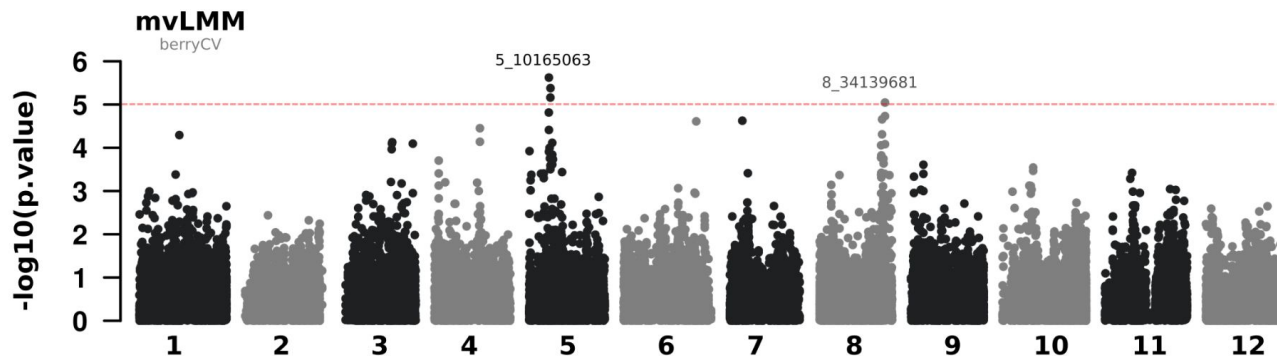
Ferrão et al., 2024. Computer vision tool integrated to genome–wide association identifies candidate genetic loci controlling waxy bloom in blueberries. In preparation

Genomic Selection 2.0
Artificial Intelligence

- Hits identified only using computer vision

University of Florida
Felipe Ferrão

Genomic Selection 2.0
Artificial Intelligence

GDSL lipolytic enzyme family
Genes expressed in the epidermal cells ,
also reported for tomato and pepper

salicylic acid methyl transferase
recent studies in blueberry have
reported a direct effect of salicylic acid
on the cuticular wax during storage

mvLMM
berryCV

5_10165063

8_34139681

- Hits identified only using computer vision

Ferrão et al., 2024. Computer vision tool integrated to genome–wide association identifies candidate
genetic loci controlling waxy bloom in blueberries. In preparation

Genomic Selection 2.0

Artificial Intelligence

Artificial Intelligence
Bruno Leme
Breeding Program

Computer Vision
Bruno Leme
Breeding Program
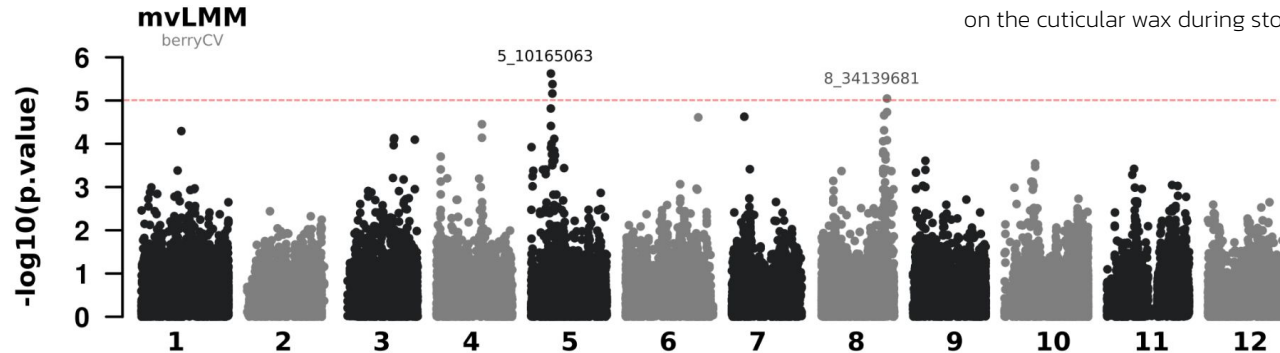

GDSL lipolytic enzyme family
Genes expressed in the epidermal cells,
also reported for tomato and pepper

salicylic acid methyl transferase
recent studies in blueberry have
reported a direct effect of salicylic acid
on the cuticular wax during storage
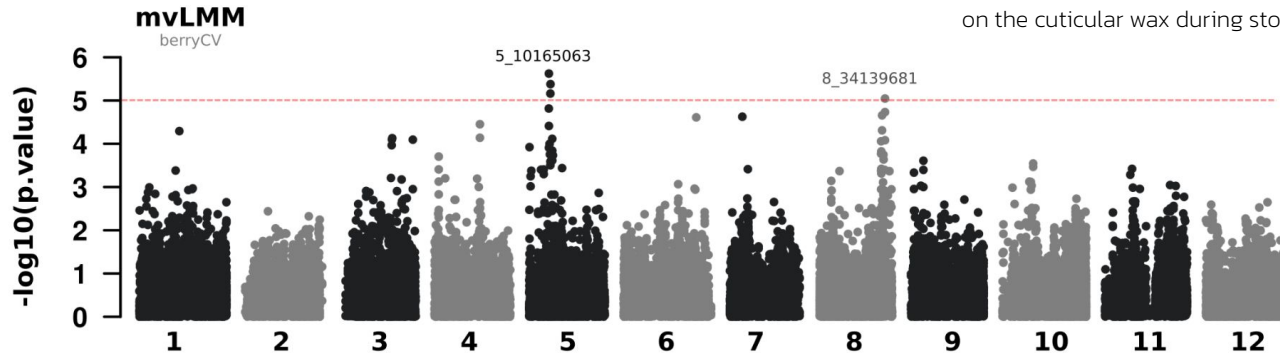
- Hits identified only using computer vision

Take home message

AI can improve data collection and shed
new light on the genetic architecture

Conclusions 🫐

FInal message

Conclusions
Final considerations

Other key tools used to support breeding decisions

1) To predict the future: stochastic **simulations**
2) To optimize the future: **mate allocation** for design crosses
3) To save time and money: Marker-Assisted Selection (**MAS**) for seedling selections

Conclusions
Final considerations

Other key tools used to support breeding decisions

1) To predict the future: stochastic **simulations**
2) To optimize the future: **mate allocation** for design crosses
3) To save time and money: Marker–Assisted Selection (**MAS**) for seedling selections
4) For biological validation: new protocols for **transformation** and **gene editing**
5) For an effective use of plant genetic resources: **pre-breeding** and **introgressions**
6) To speed up: testing a new generation of techniques for **speed-breeding**

Conclusions
Final considerations

Other key tools used to support breeding decisions

1) To predict the future: stochastic **simulations**
2) To optimize the future: **mate allocation** for design crosses
3) To save time and money: Marker–Assisted Selection (**MAS**) for seedling selections
4) For biological validation: new protocols for **transformation** and **gene editing**
5) For an effective use of plant genetic resources: **pre-breeding** and **introgressions**
6) To speed up: testing a new generation of techniques for **speed–breeding**
7) **Artificial Intelligence** is everywhere:
   a) Sensory Panels (**DeepFlavor**)
   b) Field data collection (**Computer Vision**)
   c) New methods for prediction

Conclusions
Final considerations

Seven main lessons from the last 7 years

- Genomic Selection works!

- Solid gains using data driven methods

- Prefer statistical genetics methods designed for polyploid systems

- "Simplicity is the Ultimate Sophistication"1: on the Relevance of Additive GBLUP Models

- GS can be optimized, after having a good understanding about the breeding pipeline

- Knowledge is power: decisions based on multi–layer (or omics) on information

- AI has an important play to role on data collection and analyses

1 Quote by Leonardo da Vinci

Conclusions
Final considerations

My personal opinion for practical implementation in polyploids:

- Genomic selection is a tool to assist breeding and not the other way around

- 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

- Genomic Selection is multidisciplinary topic

- Biological discovery and prediction can run side–by–side

Acknowledges

Patricio Munoz
PI Blueberry Breeding

Blueberry Lab

Juliana Bevenevenuto
Camila Azevedo
Werner Collante
Gonzalo Casorzo
Estefania Tavares
Paul Adunola
Juan Gimenez
Sava Glisic
Bruno Leme
Felix Enciso
Jim–hee

HOS departament

Marcio Resende Jr.
Denise Tieman
Harry Klee
Ali Sarkhosh
Lorenzo Rossi
Guilherme Locatelli
Jonathan Crane
Catalin Voiniciuc

Other UF departament

William Hammnond
Charlie Sims
Raquel Dias
Diego Jarquin

Other universities/institutions

David Gerard (The American University)
Matthew Stephens (UChicago)
Massimo Iorizzo (NCState)
Mary Ann Lila (NCState)
Antonio Augusto F. Garcia (USP)
Christophe Montagnon (R2D vision)
Incaper team (Brazil)
Embrapa team (Brazil)
Jeff Endelman (Wisconsin)
Rodrigo Amadeu (Bayer)
James Harynuk (University of Alberta)

Join us in PAG !

Health promotion
Estefania Tavarez
PhD project

Abstract #53796
Genome–Wide Association Analyses Reveal Candidate Genes Associated with Health Components in Blueberry

Genetics of Parthenocarpic fruits
Julie Cromie
PhD project

Talk: Plant Reproductive Genomics Workshop
Identification of markers and candidate genes for blueberry parthenocarpy through integration of GWAS and GS methodology

Phenomic Selection
Paul Adunola
PhD project

Abstract #53417
Fruit Phenomic–Assisted Selection: Assessment of the Potential of Near–Infrared Spectroscopy for Blueberry Breeding

Optimization
Simulation • Mate Allocation
Camila Azevedo
Breeding Program

Abstract #53463
Enhancing Blueberry Fruit Quality Traits Using Consumer–Centric Genomic Breeding

Non-Gaussian & Bayesian Analyses
Molecular Breeding
Patricio Munoz
PI Blueberry Breeding Program

Talk: Fruit and Nuts Workshop
Let's be Categorical: Different Genetic Statistical and Logistic Strategies to Increase Accuracy of Genomic Selection in an Apply Breeding Program

Disease Resistance
Antracnose • Botrytis
Rust • Ralstonia
Lushan Ghimire
PhD project

Abstract #53389
GWAS and Genomic Selection Strategies for Breeding Anthracnose Resistance in Southern Highbush Blueberries

Genome & QTL mapping
for wild species
Gabriel Matsumoto
PhD project

Abstract #52794
Chromosome–Level Assembly of Vaccinium stamineum: Unlocking the Genetic Basis of Anthocyanin Accumulation

Climate-Smart UF Coffee
Research & Production
Felipe Ferrão
Breeding Program

Talk: Coffee Workshop
Genomic–assisted breeding for climate–smart coffee cultivars

Thank you !!!

Question ??

About Me:

SCAN ME

Extra Questions 🫐

Genomic Selection 2.0

Questions

- What is the impact of genomic selection in the long term?
- Can we reduce the number of markers and optimize our training set?
- Is multi–omic predictions a good alternative?
- What is the impact of phenomics on fruit quality prediction?
- **Is marker–assisted selection an alternative for seedlings selections?**
- How AI can shape the future of modern breeding programs?

◉MAS
marker assisted
selection
Jim-hee
Breeding Program

Genomic Selection 2.0

Marker–assisted selection (MAS)

Motivation

- First breeding stages generates large volume of seedlings (~20.000)
- Moving all plants to the field is expensive and time–consuming for evaluation
- Genomic Selection is not an alternative at this level ($$$$)
- Can we use few markers, for key traits, to discard plants at the greenhous level?

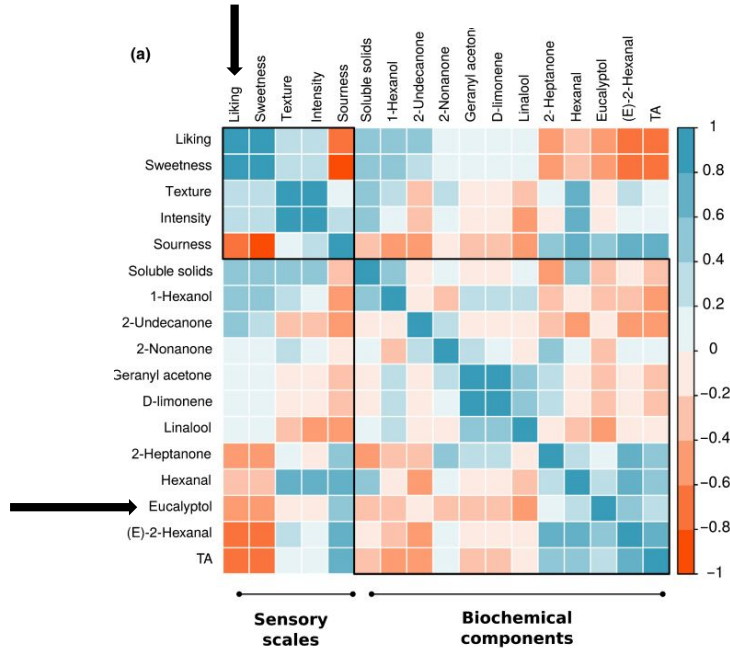# Genomic Selection 2.0
Marker–assisted selection (MAS)

Motivation
- First breeding stages generates large volume of seedlings (~20.000)
- Moving all plants to the field is expensive and time–consuming for evaluation
- Genomic Selection is not an alternative at this level ($$$$)
- Can we use few markers, for key traits, to discard plants at the greenhous level?

Goals
- Develop a rapid and cheap assay for DNA extraction
- Establish protocols for SNP genotyping using HRM and KASP
- Validate association between markers and phenotypes

Genomic Selection 2.0

Marker–assisted selection (MAS)

sensory~VOC

Eucalyptol negatively impact consumer preference

Ferrão et al., (2020)

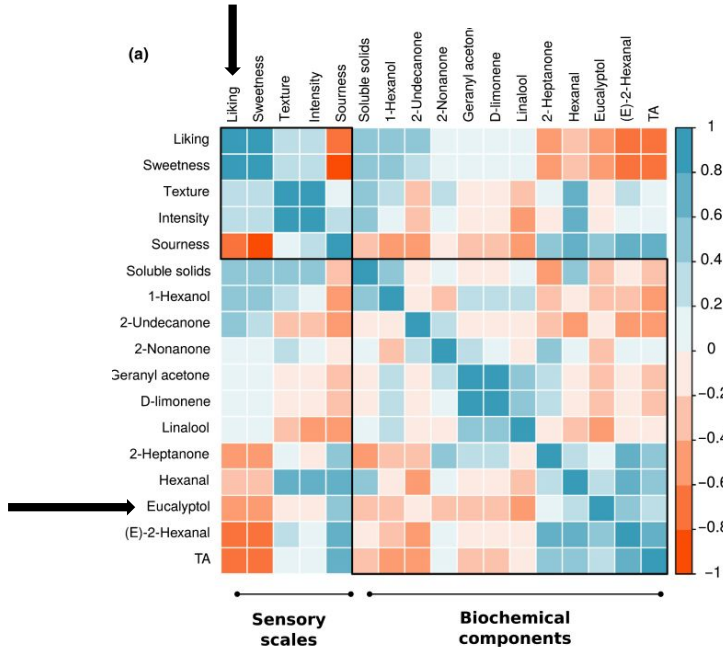https://doi.org/10.1111/nph.16459

Genomic Selection 2.0
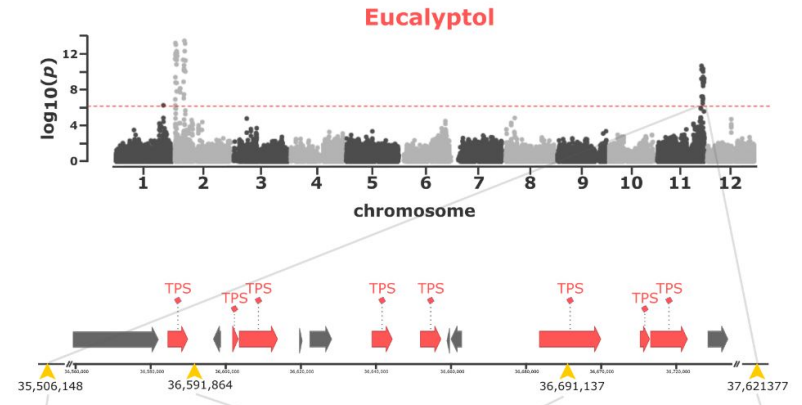Marker–assisted selection (MAS)

sensory~VOC

Eucalyptol negatively impact consumer preference

Genome-wide association analyses

Eucalyptol has a simple genetic architecture

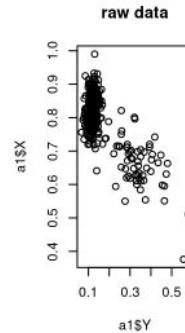**Ferrão et al., (2020)**
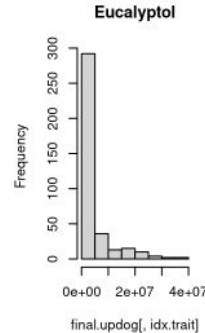
https://doi.org/10.1111/nph.16459

Genomic Selection 2.0

Marker–assisted selection (MAS)

KASP markers

- Target: Eucalyptol
- Population: 384 individuals
- Single marker explaining > 50% phenotypic variance

Genomic Selection 2.0
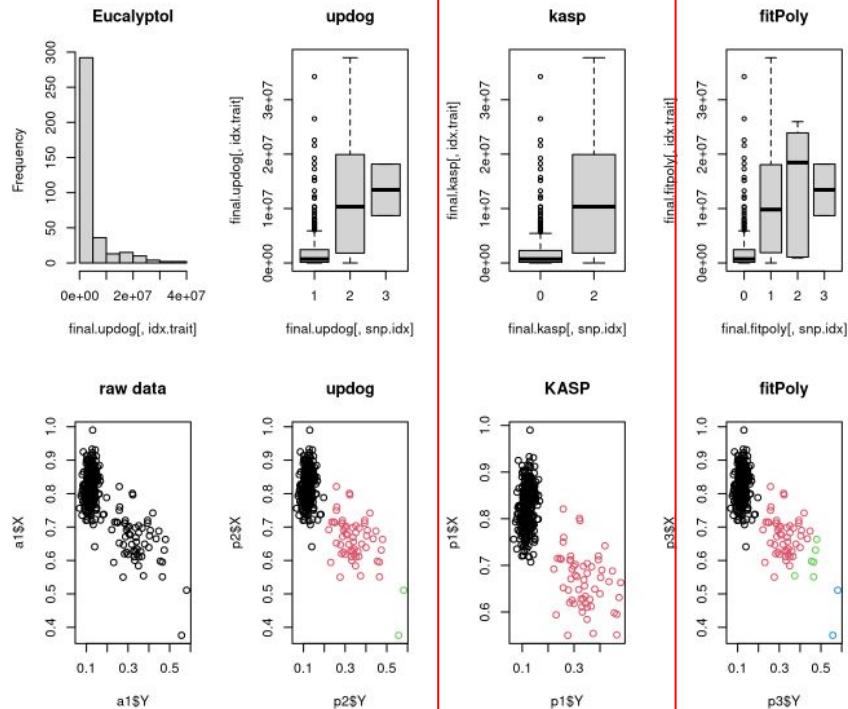
Marker–assisted selection (MAS)

KASP markers

- Target: Eucalyptol
- Population: 384 individuals
- Single marker explaining > 50% phenotypic variance

> **Multiplex vs. Duplex**
> Discriminate high vs. low eucalyptol

Genomic Selection 2.0

Marker-assisted selection (MAS)

Can we combine markers?

- Markers from two different chromosomes
- Marker1 + Marker2 = pseudo-haplotyple
- Pseudo-haplotypes 22 (duplex-duplex) and 23 (duplex-triplex) high levels of eucalyptol
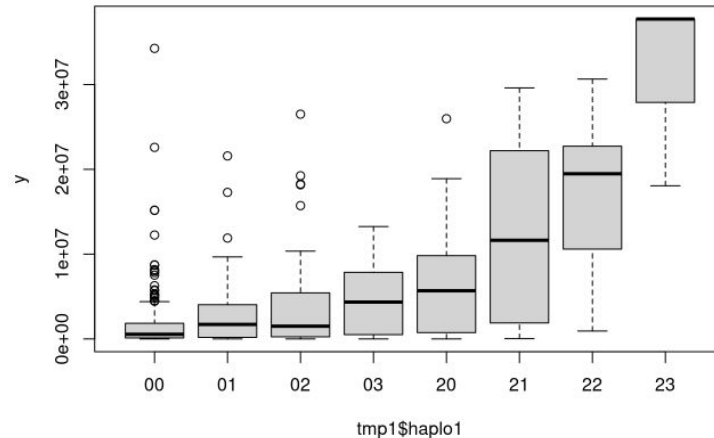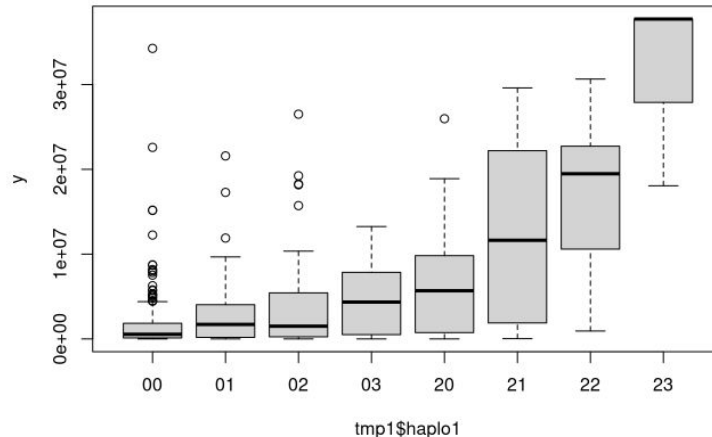
Genomic Selection 2.0

Marker–assisted selection (MAS)

Can we combine markers?

- Markers from two different chromosomes
- Marker1 + Marker2 = pseudo–haplotyple
- Pseudo–haplotypes 22 (duplex–duplex) and 23 (duplex–triplex) high levels of eucalyptol

Take home message

Cheap and low density marker platforms can be used to discard plants in the seedling stage

Genomic Selection 2.0
What is the impact of genomic selection in the long term?

Camila Azevedo
Breeding Program

Simulation

- We can use stochastic <u>oriented</u> simulation to project the future
- Oriented because we can use real information from the breeding program
- AlphaSim package, to simulate a complex trait (h2=0.30, 100 QTL)

Genomic Selection 2.0

What is the impact of genomic selection in the long term?

Simulation

- We can use stochastic <u>oriented</u> simulation to project the future
- Oriented because we can use real information from the breeding program
- AlphaSim package, to simulate a complex trait (h2=0.30, 100 QTL)

Stage I	▷	Stage II	▷	Stage III	▷	Stage IV	▷	Cultivar
20,000 plants		2,000 plants		200 plants		20 plants		

PARENTAL SELECTION

Different criteria for selection and mate allocation in the UF/IFAS breeding program

Selection Criteria

1. Massal Selection
2. EBV (A-BLUP)
3. GEBV (G-BLUP)

Mate allocation for GS

1. GEBV based on parental means (GS)
2. GEBV + SimpleMatting (GS -Opt)

Genomic Selection 2.0

What is the impact of genomic selection in the long term?

Camila Azevedo
Breeding Program

Simulation

- We can use stochastic <u>oriented</u> simulation to project the future
- Oriented because we can use real information from the breeding program
- AlphaSim package, to simulate a complex trait (h2=0.30, 100 QTL)

Example from AlphaMate

Stage I		Stage II		Stage III		Stage IV		Cultivar
20,000 plants		2,000 plants		200 plants		20 plants		

PARENTAL SELECTION

Different criteria for selection and mate allocation in the UF/IFAS breeding program
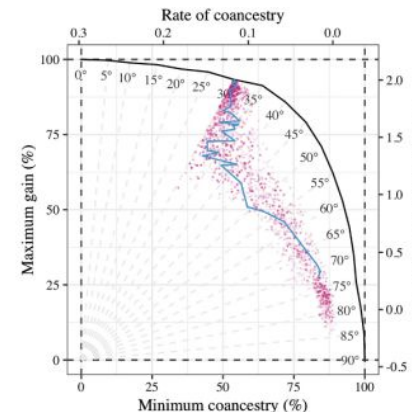
Selection Criteria

1. Massal Selection
2. EBV (A-BLUP)
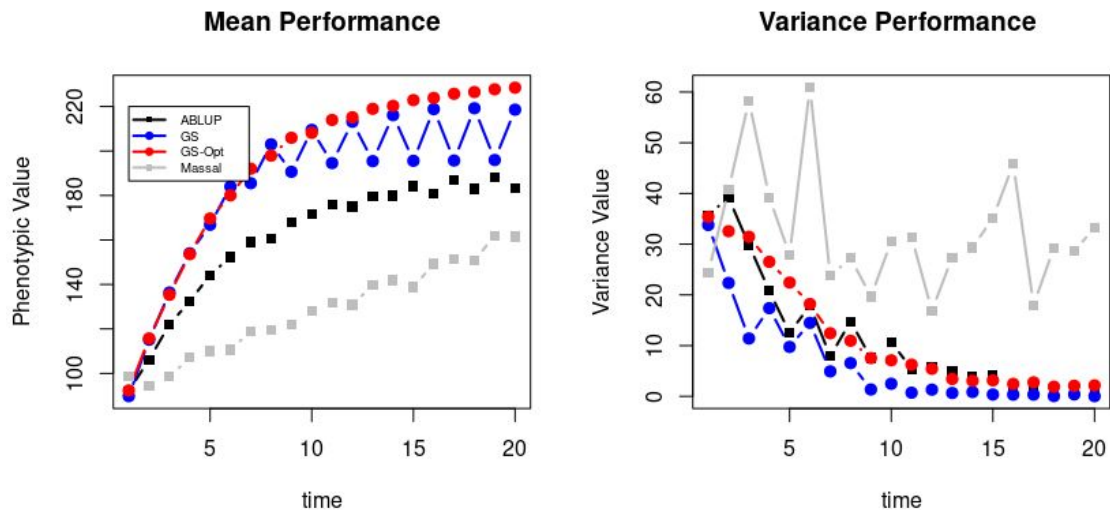3. GEBV (G-BLUP)

Mate allocation for GS

1. GEBV based on parental means (GS)
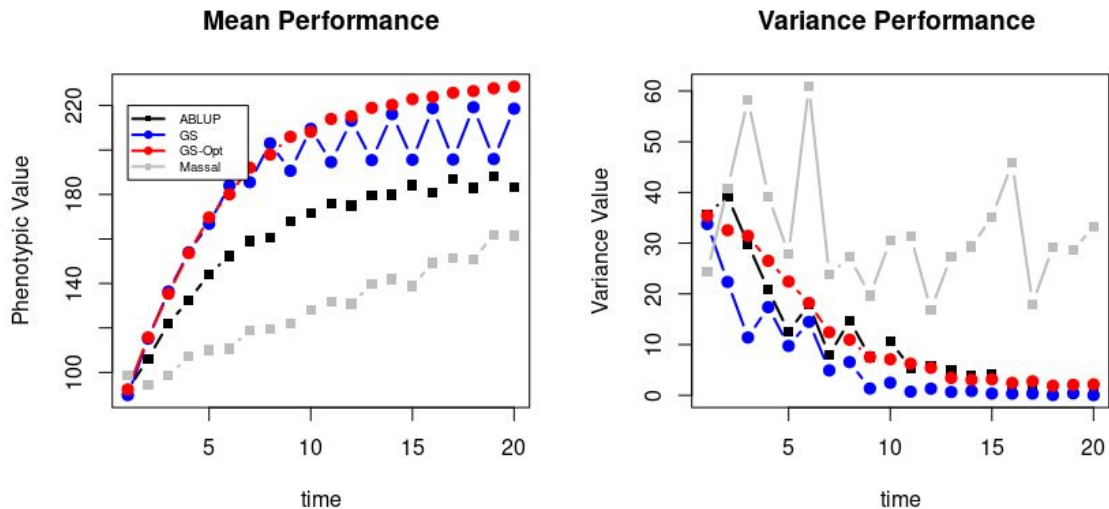2. GEBV + SimpleMatting (GS -Opt)

Genomic Selection 2.0

What is the impact of genomic selection in the long term?

Camila Azevedo
Breeding Program

Azevedo et al., 2023. Stochastic Simulation in Blueberry. In preparation

Genomic Selection 2.0

What is the impact of genomic selection in the long term?

Optimization
Simulation • Mate Allocation
Camila Azevedo
Breeding Program

Mean Performance

Variance Performance

Azevedo et al., 2023. Stochastic Simulation in Blueberry. In preparation

Take home message

GS and mate allocation can maximize the gains in the long term