

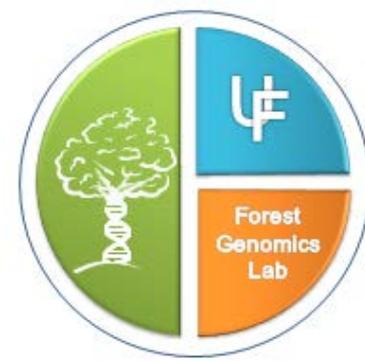
What is Next in the Application of Genome-Wide Information to Tree Breeding?

Matias Kirst

**Professor of Quantitative Genetics and Genomics
Coordinator Plant Molecular and Cellular Biology Graduate Program
Co-Director Cooperative Forest Genetics Research Program
University of Florida Genetics Institute
School of Forest Resources and Conservation**

SFTIC

Melbourne, Florida – June 20 2017



Will we fail (again)?

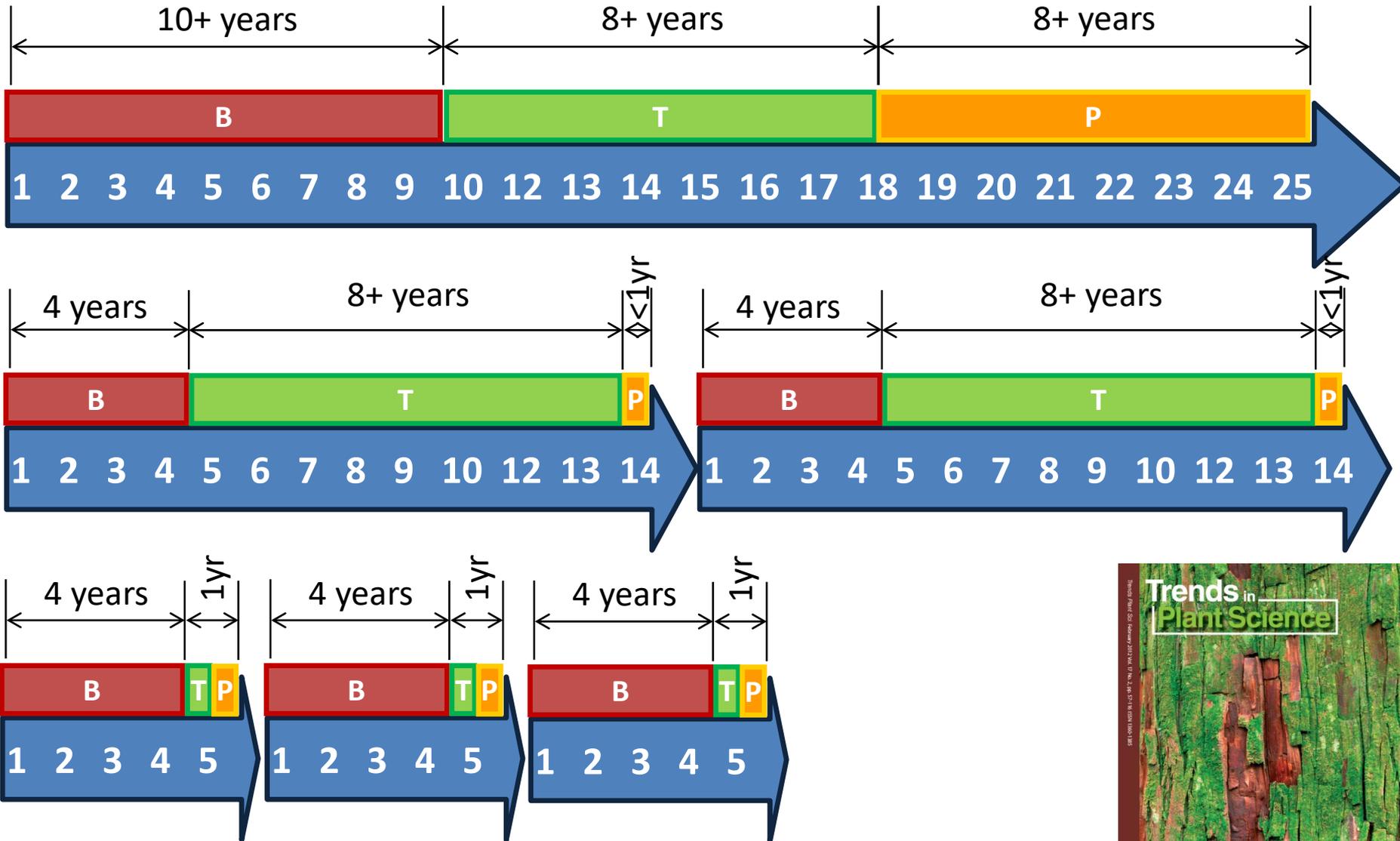
Matias Kirst

**Professor of Quantitative Genetics and Genomics
Coordinator Plant Molecular and Cellular Biology Graduate Program
Co-Director Cooperative Forest Genetics Research Program
University of Florida Genetics Institute
School of Forest Resources and Conservation**

SFTIC

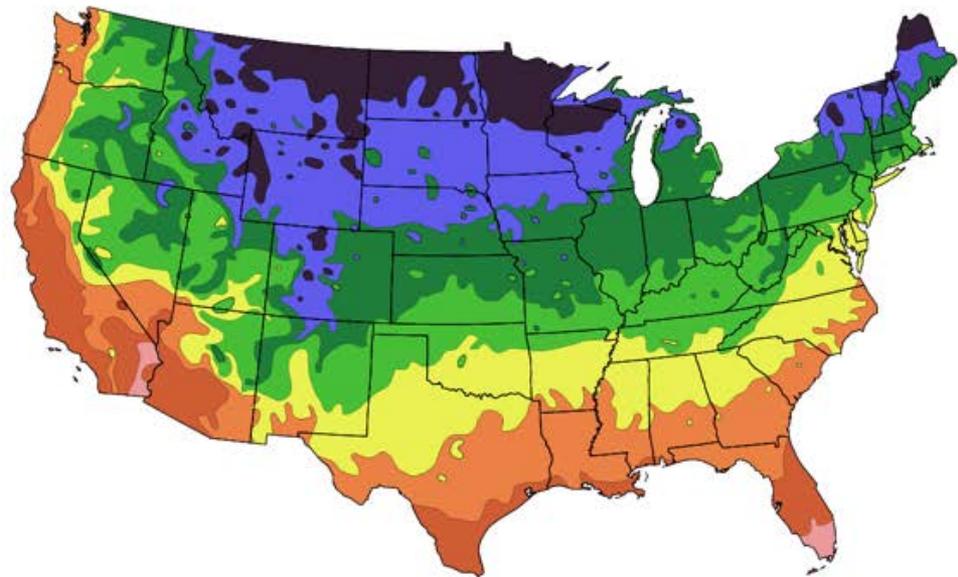
Melbourne, Florida – June 20 2017

Genomics incorporated into pine breeding



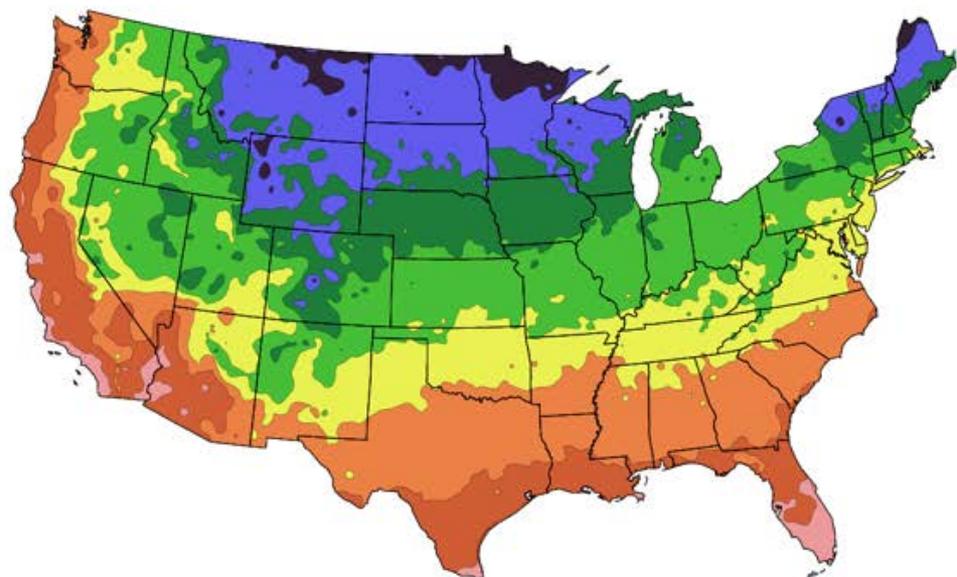
Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Mugnozza GS (2012) Accelerating the domestication of forest trees in a changing world. Trends in Plant Sciences.

1990 Map



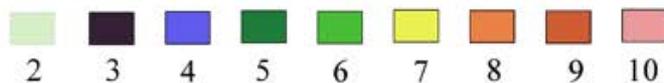
After USDA Plant Hardiness Zone Map, USDA Miscellaneous
Publication No. 1475, Issued January 1990

2006 Map



National Arbor Day Foundation Plant Hardiness Zone Map
published in 2006.

Zone

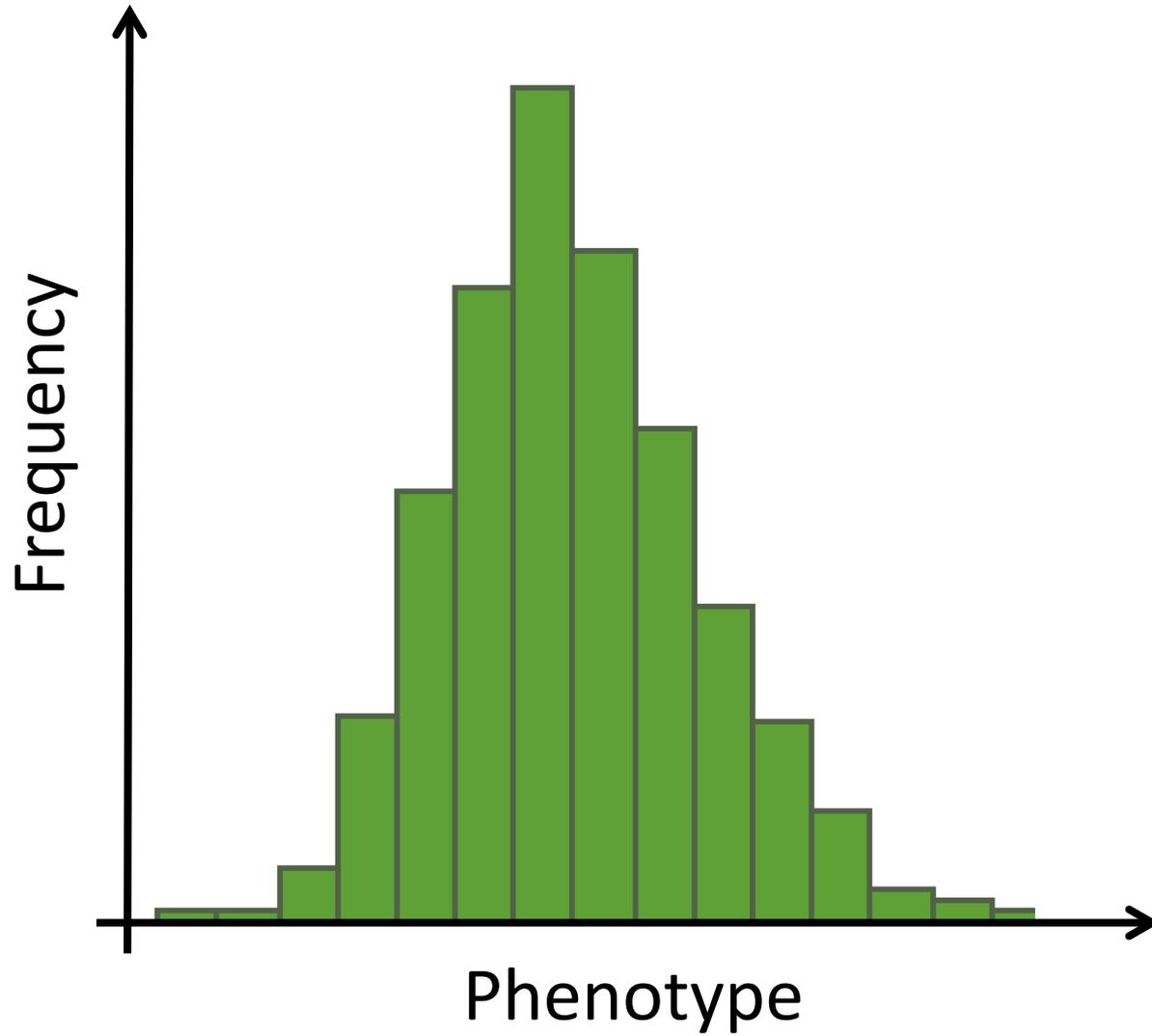


© 2006 by The National Arbor Day Foundation®





No. 34. Complete destruction of chestnut trees in a nearly pure stand. Many of the trunks have lost their bark. Scene in Forest Park, near Brooklyn, New York.—*Photograph by Prof. Collins. (1912)*



1990's

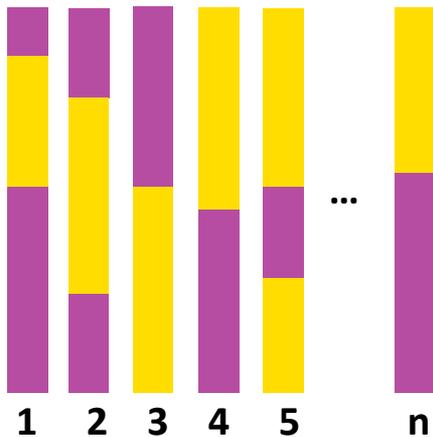
QTL

Parents



P1 P2

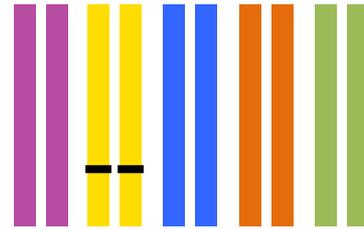
Current population



2010's

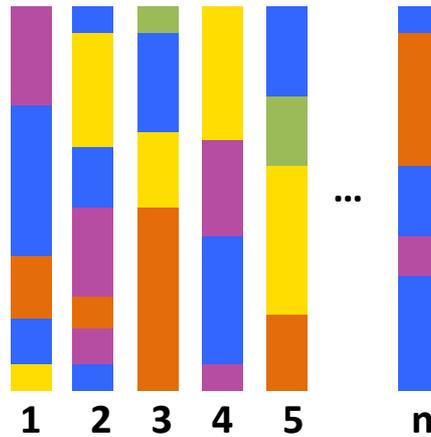
Genome-wide selection

Breeding parents



P1 P2 P3 P4 P5

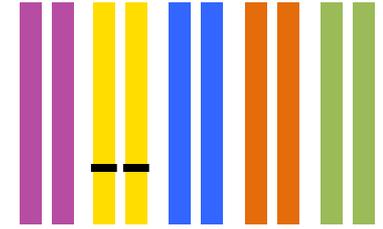
Current population



2000's

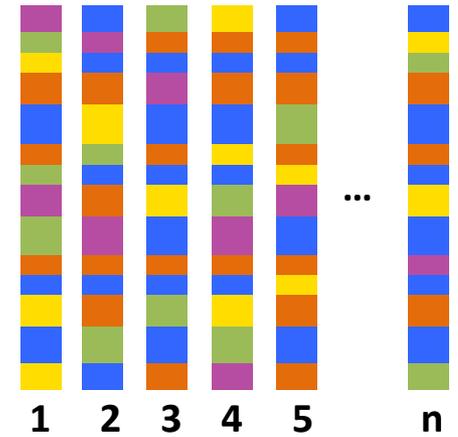
GWAS

Ancestral population

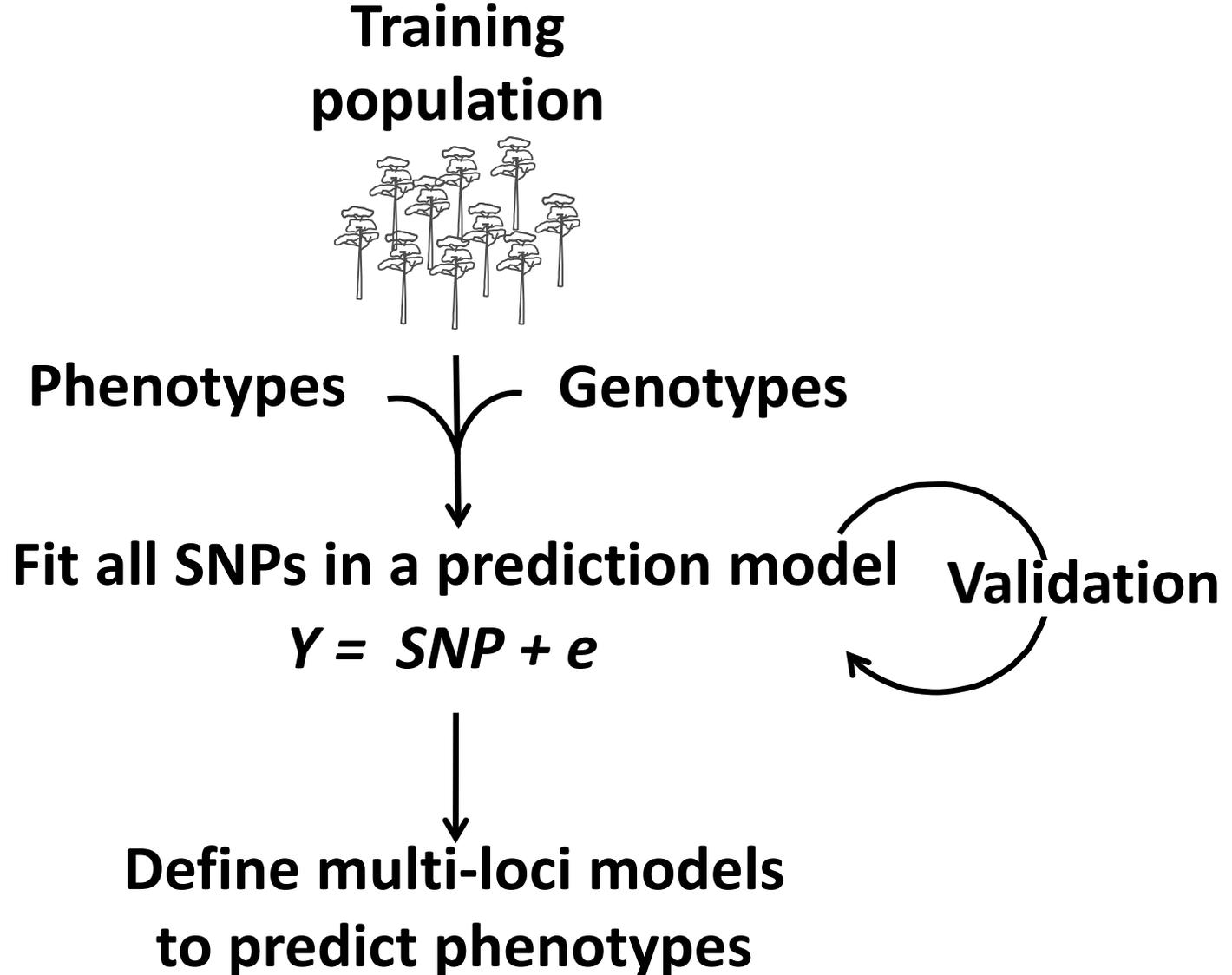


P1 P2 P3 P4 P5

Current population

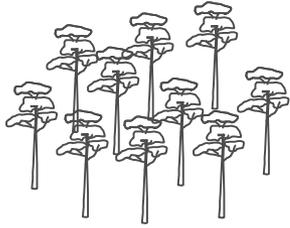


Genome-Wide Selection



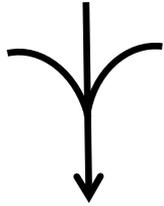
Genome-Wide Selection

**Training
population**



Phenotypes

Genotypes



Fit all SNPs in a prediction model

$$Y = SNP + e$$



**Define multi-loci models
to predict phenotypes**

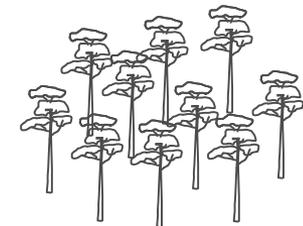
**Application in early MAS
Generation 3+**



**Application in early MAS
Generation 2**



**Application in early MAS
Generation 1**



**Training
population**

GWS

Genome-Wide Selection



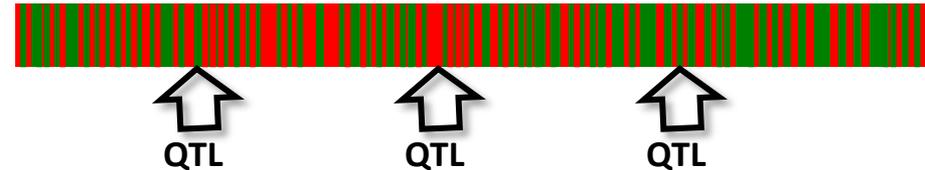
Lower $N_e \sim$ lower allelic range

Higher LD = lower resolution

Application to breeding

GWAS

Genome-Wide Association Studies



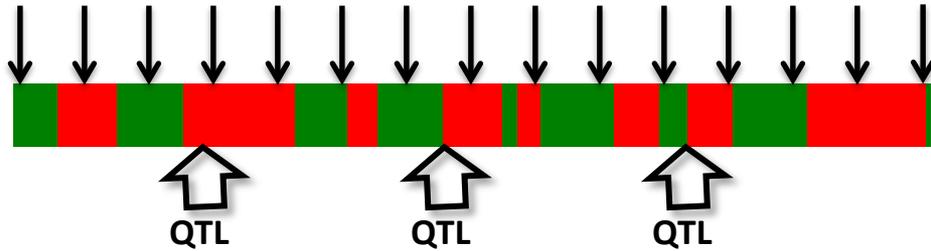
Higher $N_e \sim$ higher allelic range

Lower LD = higher resolution

Application to discovery of
gene function / functional
genomics

GWS

Genome-Wide Selection



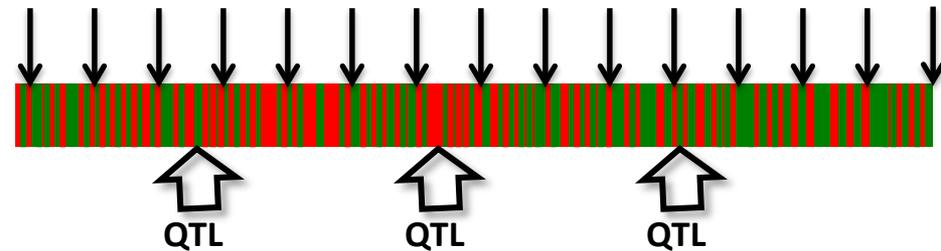
Lower $N_e \sim$ lower allelic range

Higher LD = lower resolution

Application to breeding

GWAS

Genome-Wide Association Studies



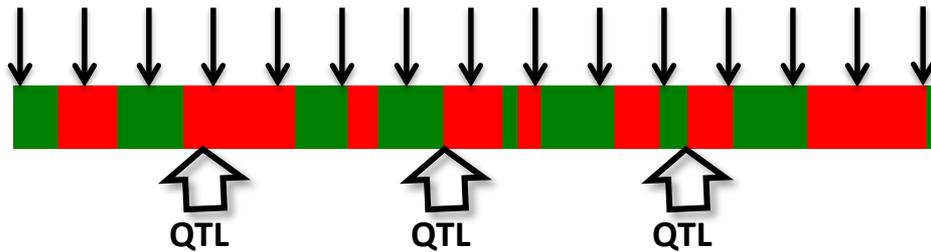
Higher $N_e \sim$ higher allelic range

Lower LD = higher resolution

Application to discovery of
gene function / functional
genomics

GWS

Genome-Wide Selection



Lower $N_e \sim$ lower allelic range

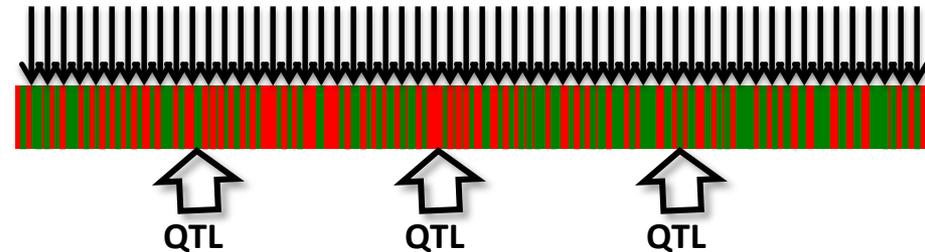
Higher LD = lower resolution

Application to breeding

Markers needed \approx 10-20/cM

GWAS

Genome-Wide Association Studies



Higher $N_e \sim$ higher allelic range

Lower LD = higher resolution

Application to discovery of gene function / functional genomics

Markers needed \approx 1M+

ORIGINAL PAPER

Genomic selection in fo

Dario Grattapaglia • Marcos D. V. Resende

Tree Genetics & Genomes (2011) 7:747–758
DOI 10.1007/s11295-011-0371-9

ORIGINAL PAPER

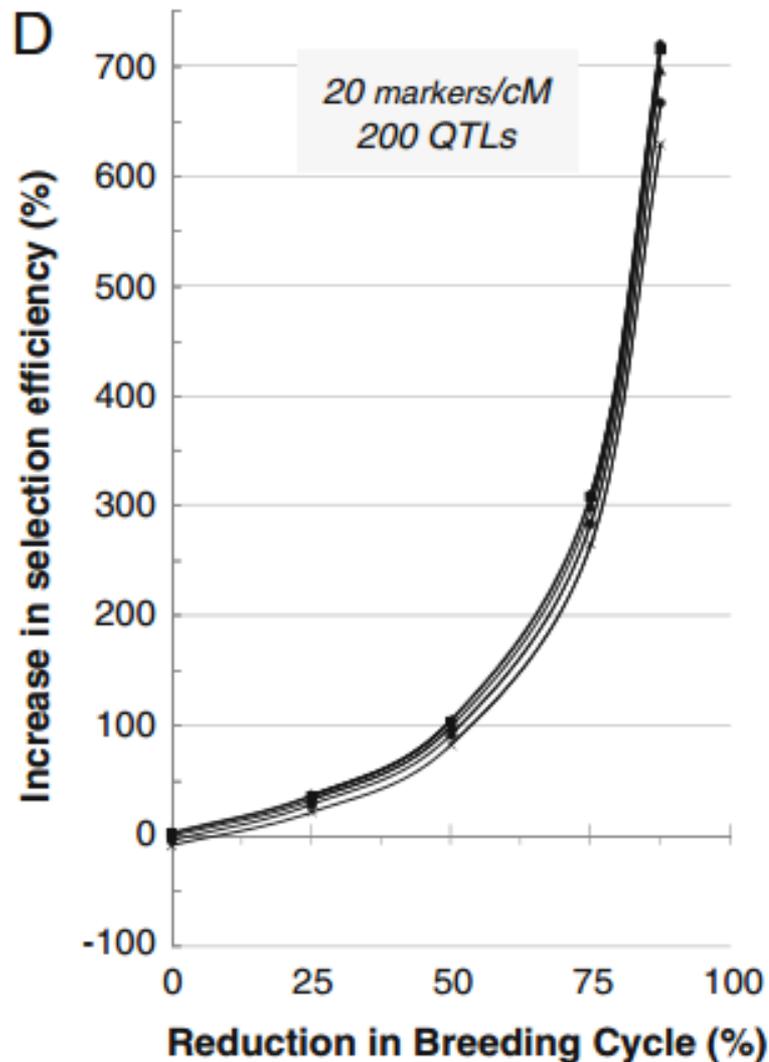
Prospects for genomic a simulation study of

Hiroyoshi Iwata • Takeshi Hayashi • Y



Accelerating the dom
accuracy of predictio

M. F. R. Resende Jr^{1,3}, P. Muñoz^{2,3},
M. D. V. Resende^{7,8} and M. Kirst^{3,4}



Methods in a Standard ly Pine (*Pinus taeda* L.)

V. Resende,^{5,***} D. J. Garrick,^{1†} R. L. Fernando,^{1†}
T. A. Martin,¹ G. F. Peter,^{1,††} and M. Kirst^{1,††,2}

Genetics, Vol. 190, 1503–1510 April 2012 1503

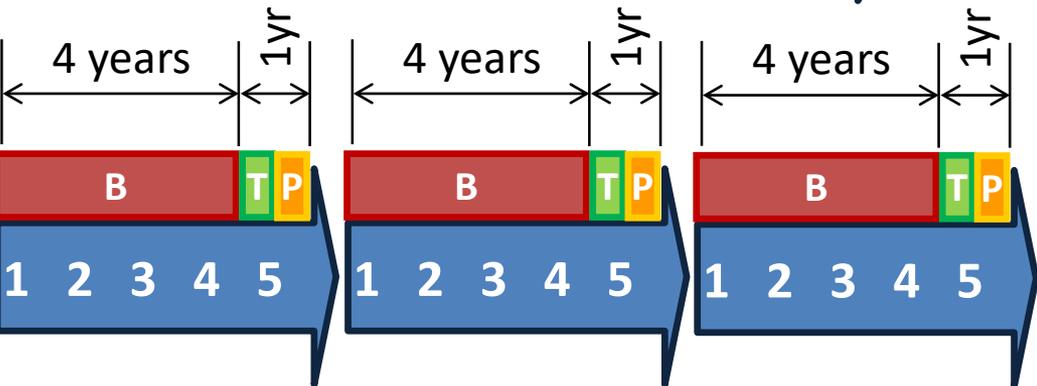
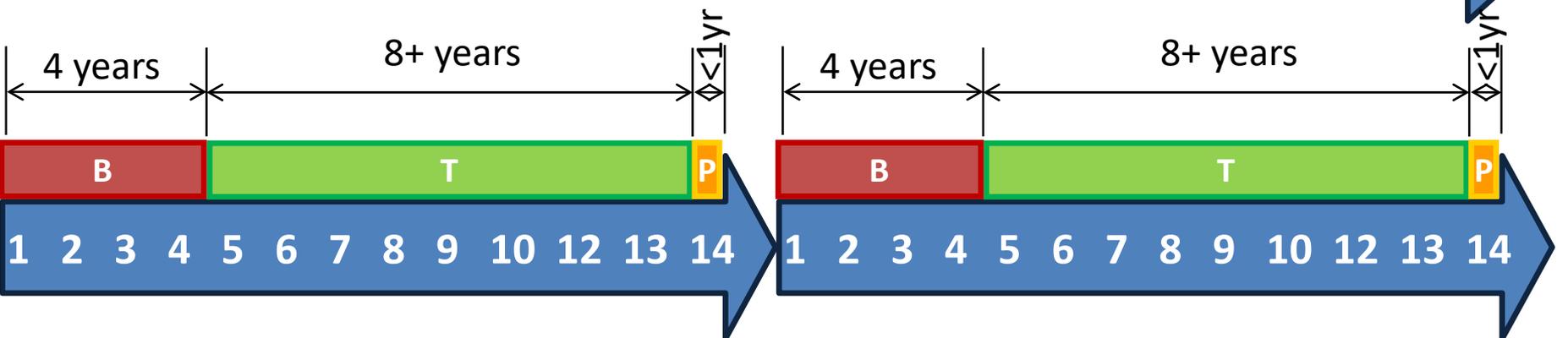


Open Access

acies within and between
breeding groups in

dré Rainville³ and Jean Bousquet²

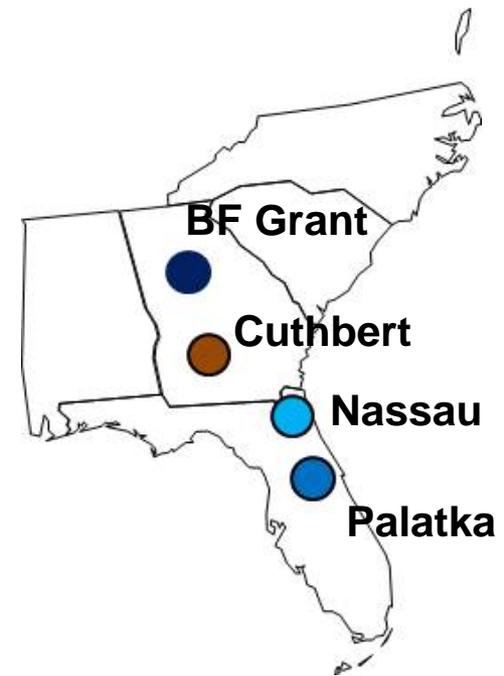
Genomics incorporated into pine breeding



Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Mugnozza GS (2012) Accelerating the domestication of forest trees in a changing world. Trends in Plant Sciences.

Genomic Selection in Conifers

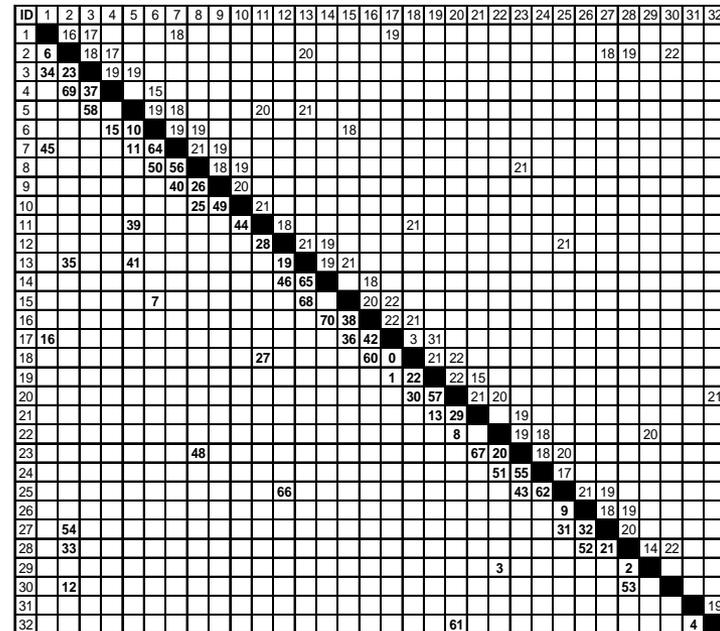
- "Comparing CLonal Lines on Experimental Sites = CCLONES"
- 900 replicated loblolly pine clones developed from 62 elite full-sib families
- $N_e \approx 40$
- Four field locations
- ~ 3 SNP markers/cM



Márcio Resende Jr.



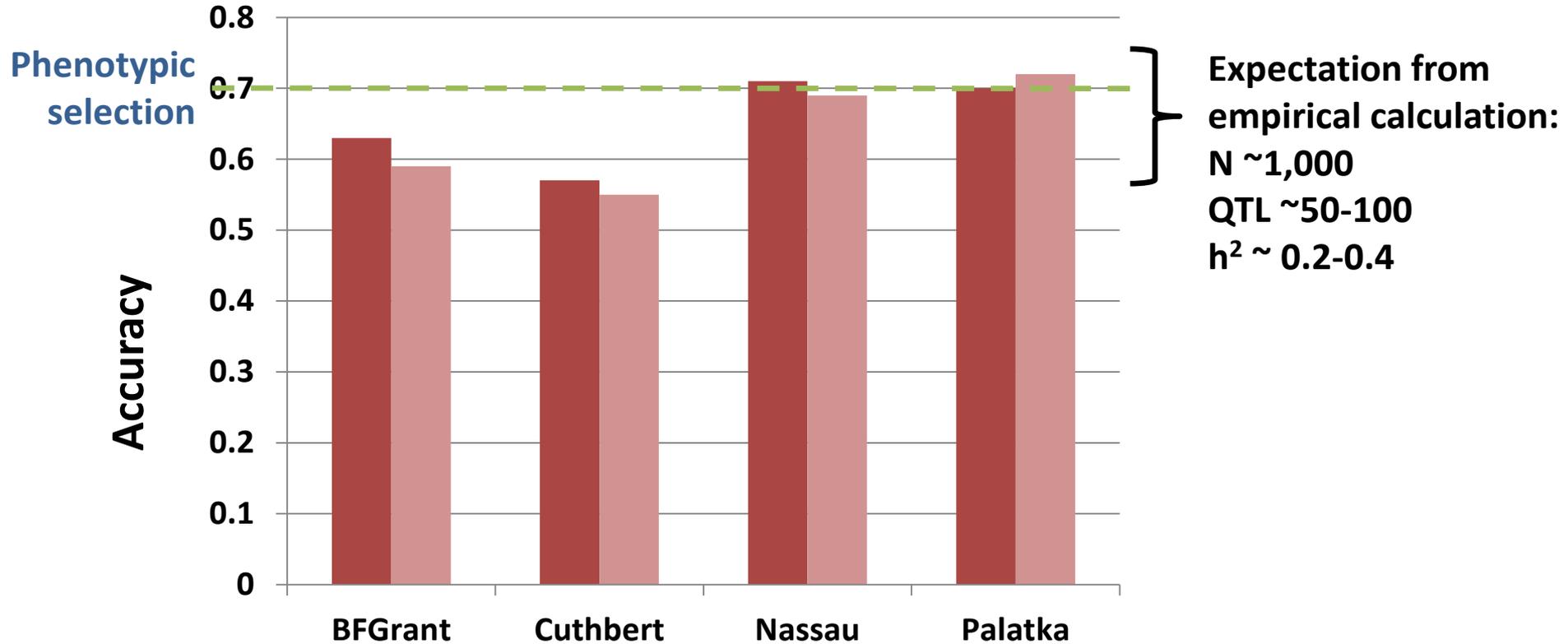
Patricio Munoz



de Almeida Filho et al. *Heredity*. 2016 Jul;117(1):33-41.
Muñoz PR et al. *Genetics*. 2014 Dec;198(4):1759-68.
Resende MF Jr et al. *Genetics*. 2012 Apr;190(4):1503-10.
Resende MF Jr et al. *New Phytol*. 2012 Feb;193(3):617-24.

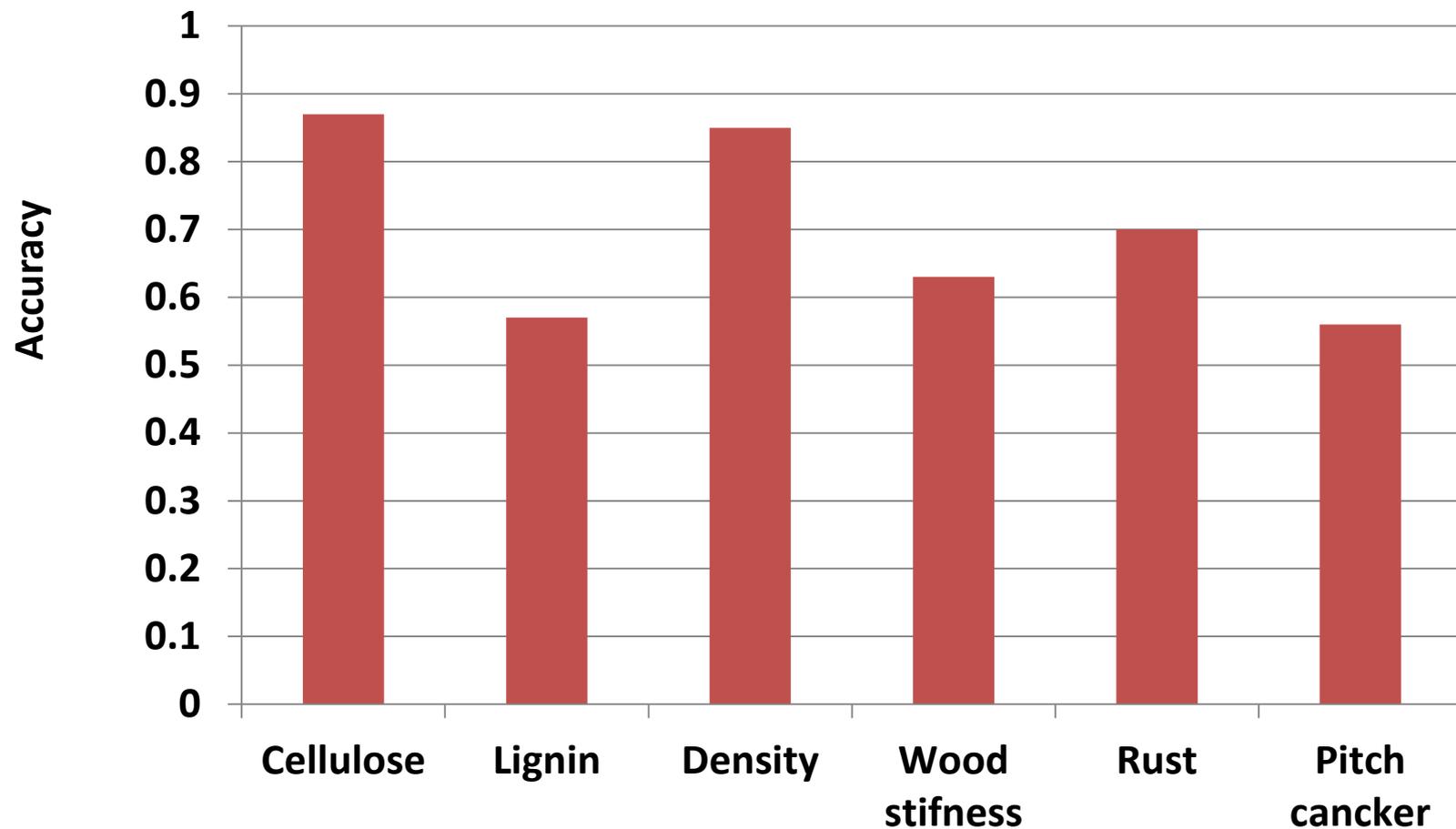
Genomic Selection Accuracies in CCLONES

DBH & Height



Genomic Selection Accuracies in CCLONES

Other trait accuracies

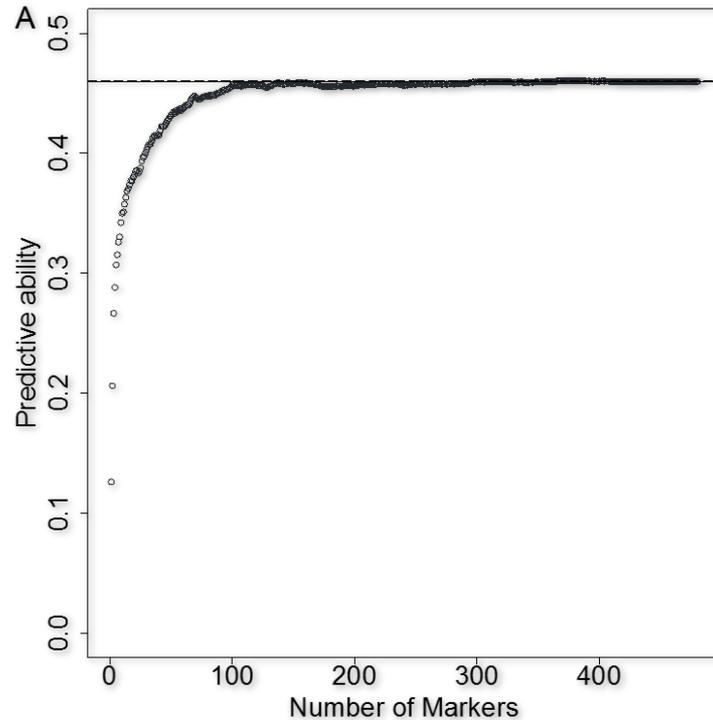


Genomic Selection Accuracies in CCLONES

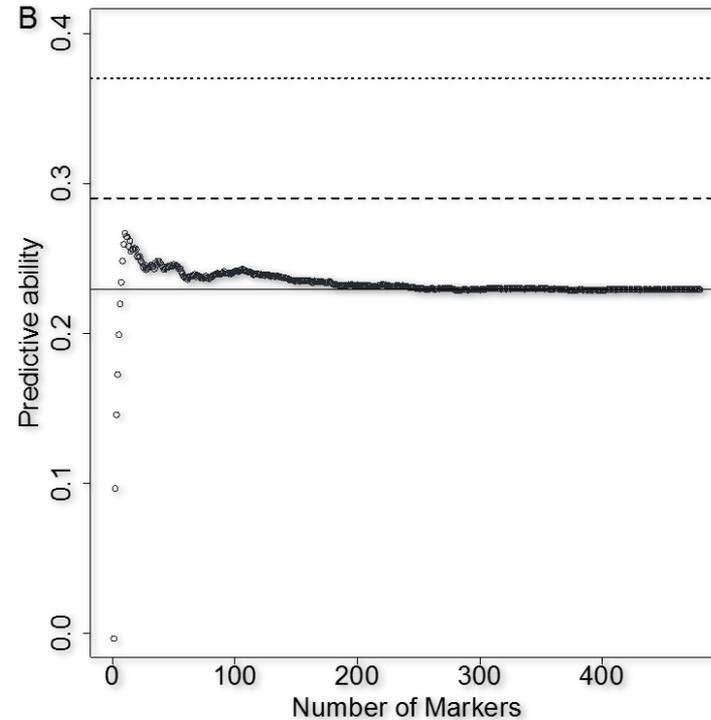
Trait category	Trait	Methods			
		RR-BLUP	BLASSO	Bayes A	Bayes Cπ
Growth	Height	0.39	0.38	0.38	0.38
	Diameter	0.46	0.46	0.46	0.46
Development	Crown width	0.48	0.46	0.47	0.47
	Branch diameter	0.27	0.25	0.27	0.27
	Branch angle	0.51	0.51	0.51	0.51
	Root number	0.24	0.26	0.25	0.24
Disease resistance	Rust	0.29	0.28	0.34	0.34
	Rust volume	0.23	0.24	0.28	0.29
Wood quality	Stiffness	0.43	0.39	0.42	0.42
	Lignin	0.17	0.17	0.17	0.17
	Late Wood	0.24	0.24	0.23	0.24
	Density	0.20	0.22	0.23	0.22
	C5C6	0.26	0.25	0.25	0.25

Genomic Selection Accuracies in CCLONES

Diameter



Rust volume

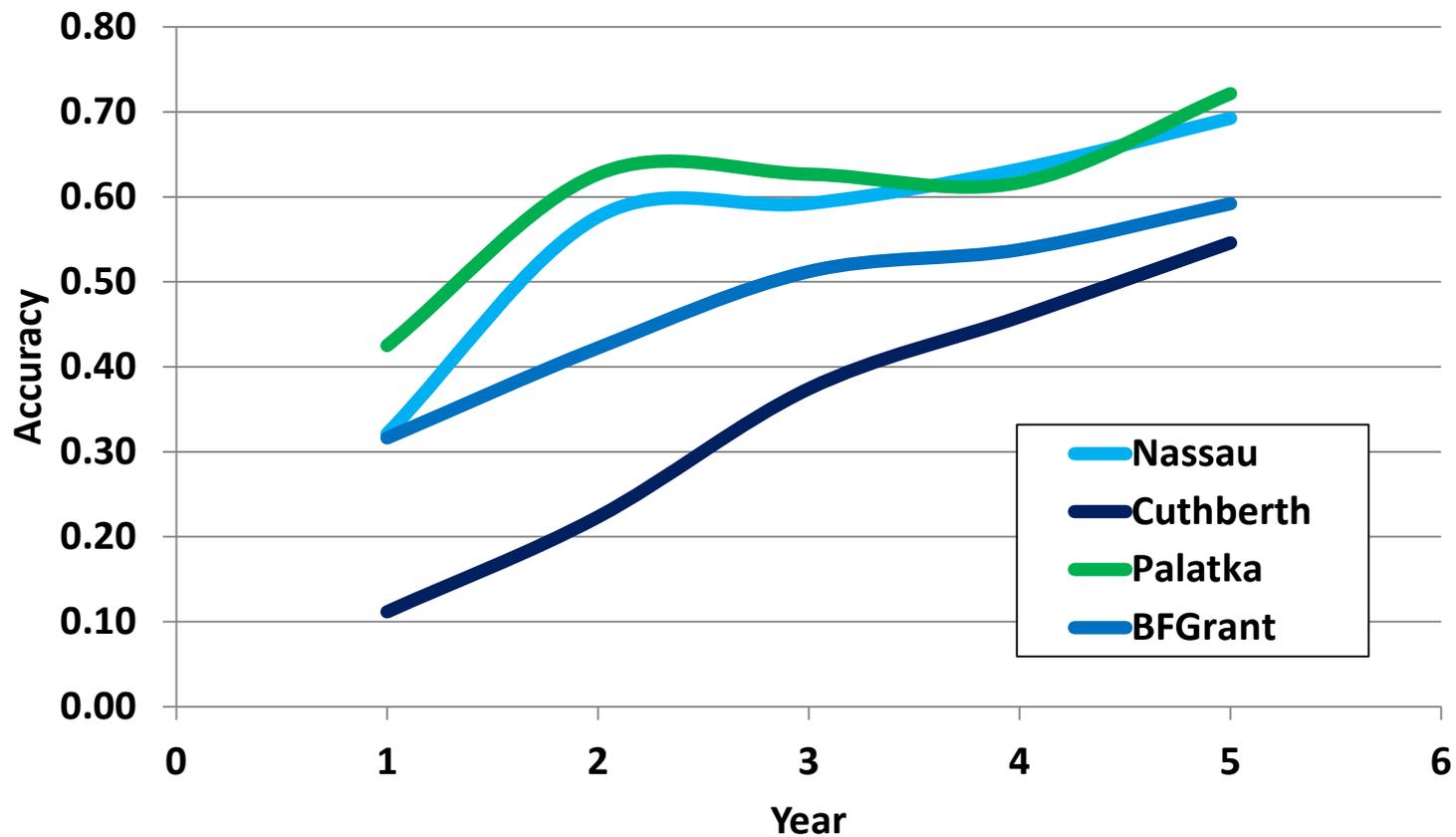


Trait	Methods				
	RR-BLUP	BLASSO	Bayes A	Bayes C π	RR-BLUP B
Rust	0.29	0.28	0.34	0.34	0.33
Rust volume	0.23	0.24	0.28	0.29	0.37

Genomic Selection Accuracies in CCLONES

How stable are prediction models across years?

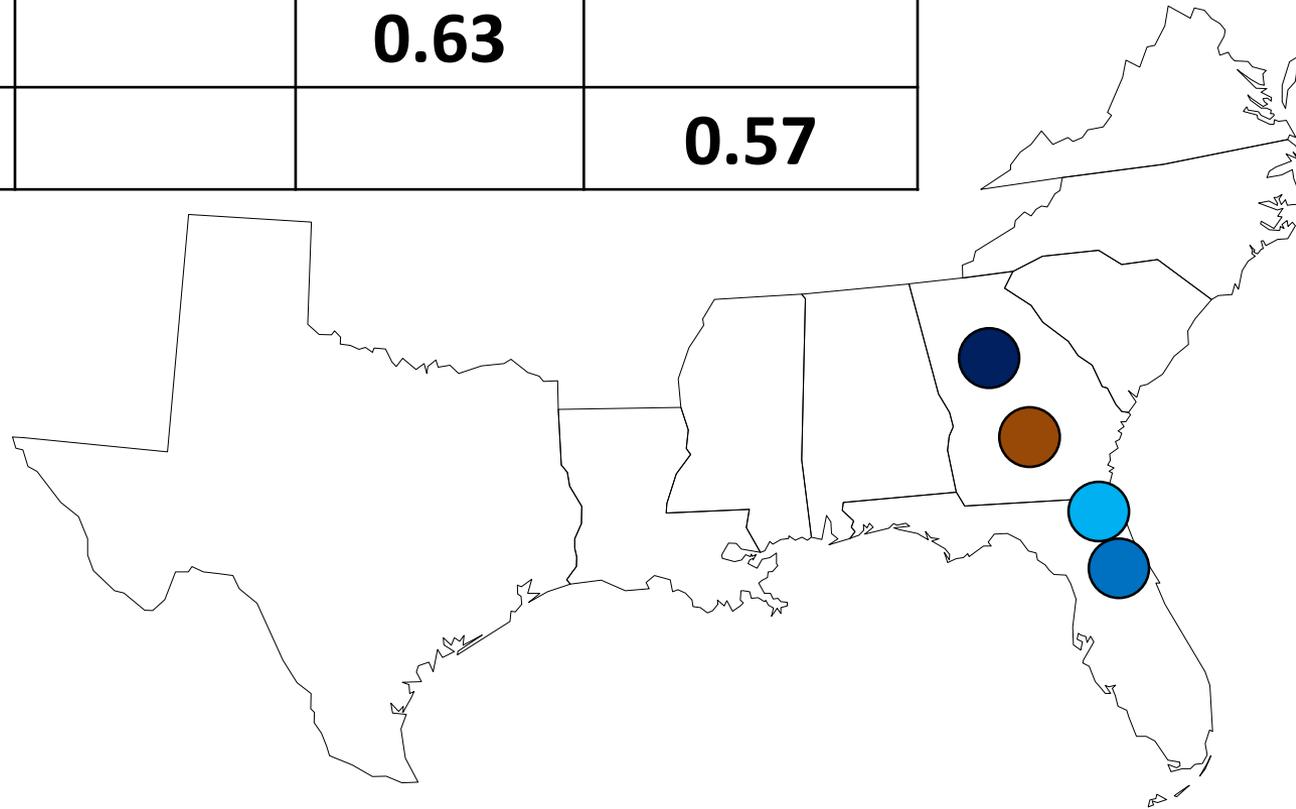
Height



Genomic Selection Accuracies in CCLONES

How stable are prediction models across sites?

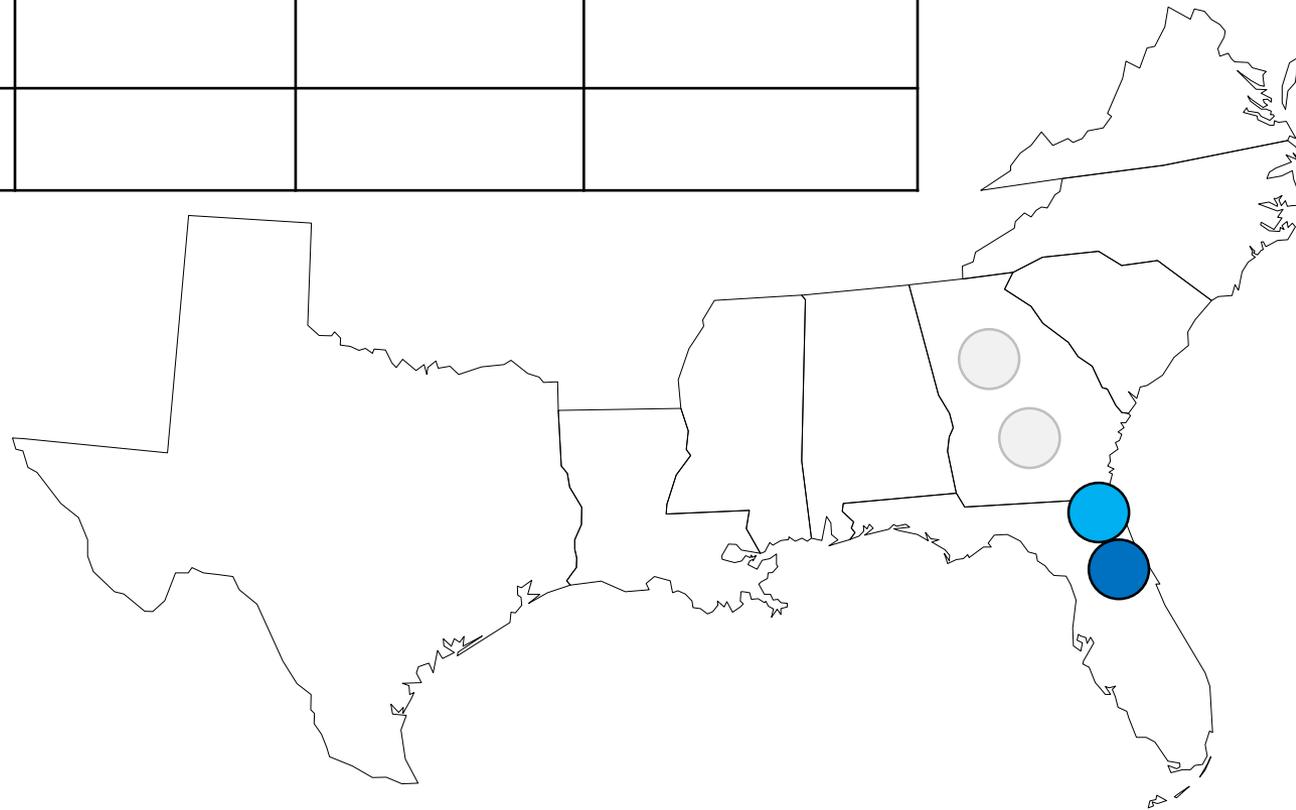
DBH	Nassau	Palatka	BFGGrant	Cuttberth
Nassau	0.71			
Palatka		0.70		
BFGGrant			0.63	
Cuttberth				0.57



Genomic Selection Accuracies in CCLONES

How stable are prediction models across sites?

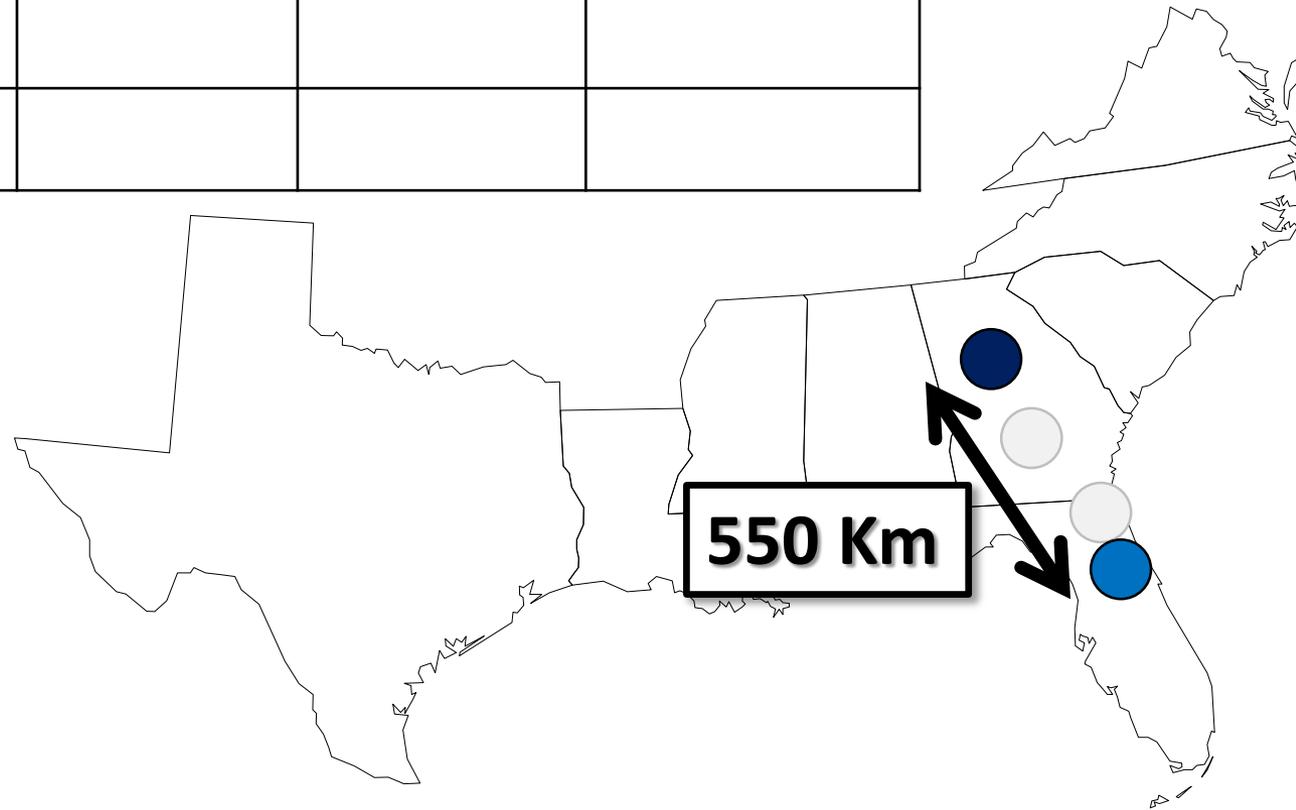
DBH	Nassau	Palatka	BFGGrant	Cuttberth
Nassau	0.71	0.44		
Palatka	0.48	0.70		
BFGGrant				
Cuttberth				



Genomic Selection Accuracies in CCLONES

How stable are prediction models across sites?

DBH	Nassau	Palatka	BFGGrant	Cuttberth
Nassau				
Palatka			0.11	
BFGGrant				
Cuttberth				



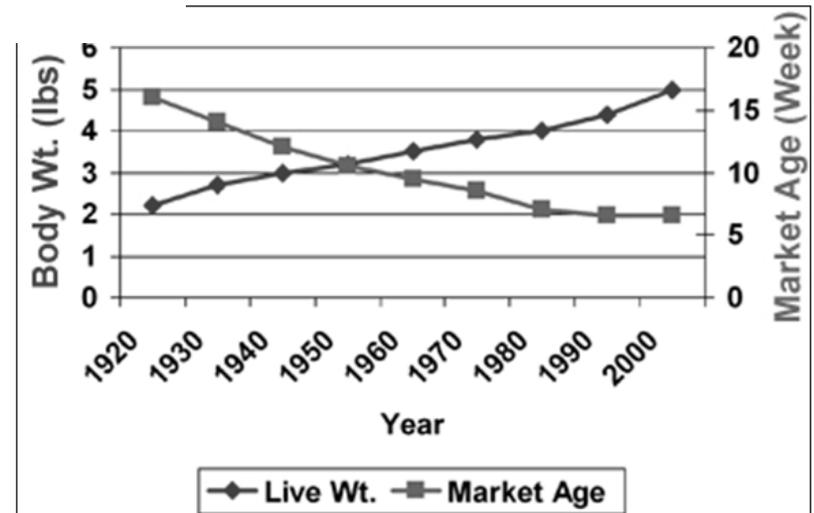
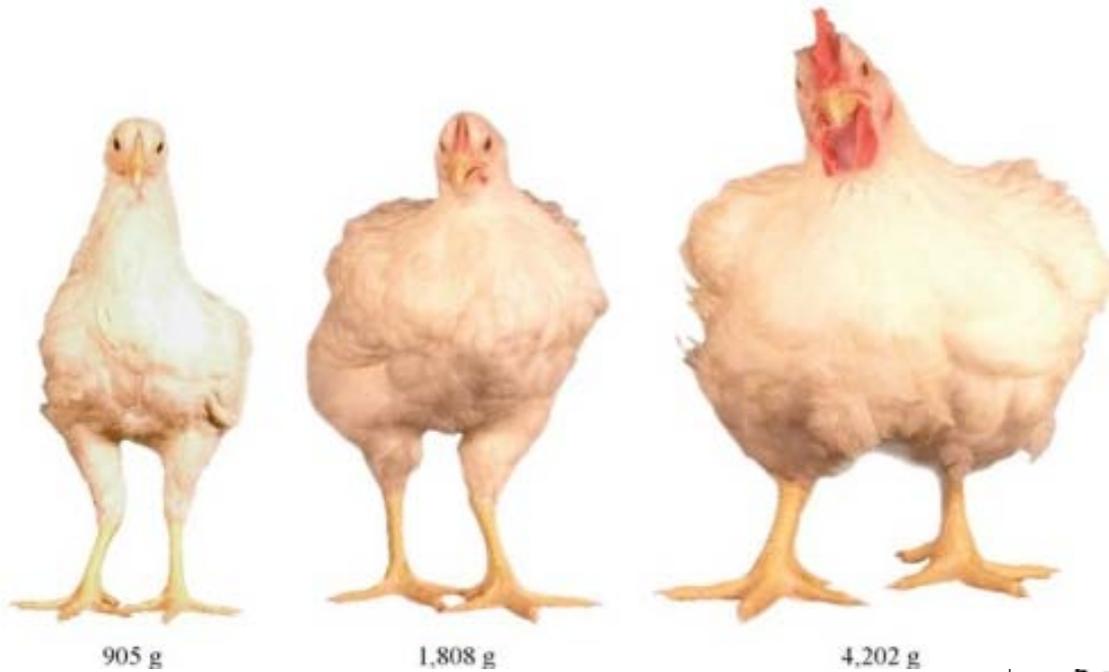
Genetic Gain Relative to Phenotypic Selection

- Increase in genetic gain using genomic selection relative to phenotypic selection.

Trait	Site	h (BLUP)	h (GS)	Efficiency	Increase relative to phenotypic selection (%)
DBH	B.F. Grant	0.79	0.75	1.90	90
	Cuthbert	0.75	0.73	1.95	95
	Nassau	0.85	0.65	1.53	53
	Palatka	0.81	0.67	1.65	65
HT	B.F. Grant	0.74	0.77	2.08	108
	Cuthbert	0.68	0.74	2.18	118
	Nassau	0.80	0.64	1.60	60
	Palatka	0.85	0.67	1.58	58

Will this work (next generation) in forestry?

Of course, it worked with...



Unfortunately, breeding trees is not the same as
breeding chicken...

What are the threats to the application of genomics/genome-wide selection in forestry?

- **Genotype by environment interaction**

Prediction models apply to narrow regions

- **Low multiplier**

Many suppliers, horizontal market

- **High cost of genotyping**

Scale and affordability

- **Low heritabilities**

Better training populations needed

- **Slow breeding**

Slow translation of science to application

What are the threats to the application of genomics/genome-wide selection in forestry?

- **Genotype by environment interaction**

Prediction models apply to narrow regions

- **Low multiplier**

Many suppliers, horizontal market

- **High cost of genotyping**

Scale and affordability

- **Low heritabilities**

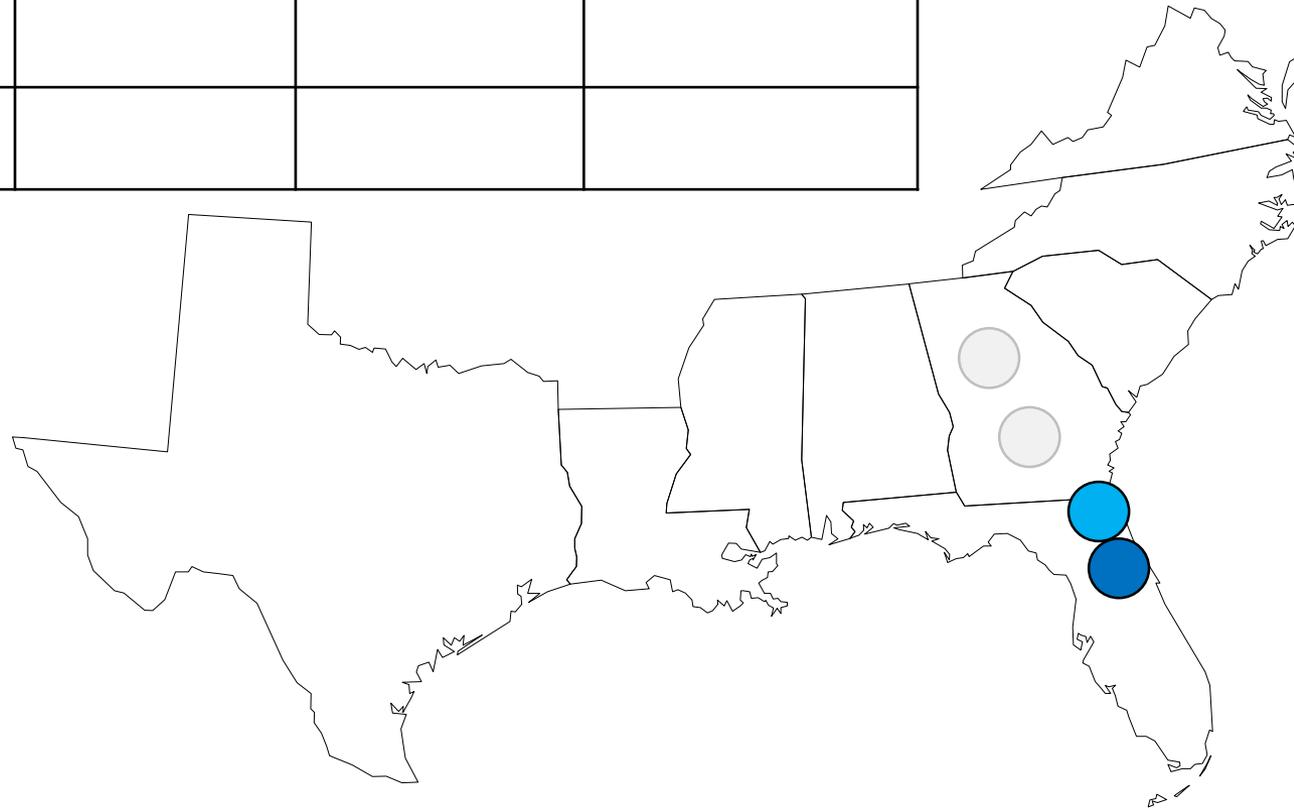
Better training populations needed

- **Slow breeding**

Slow translation of science to application

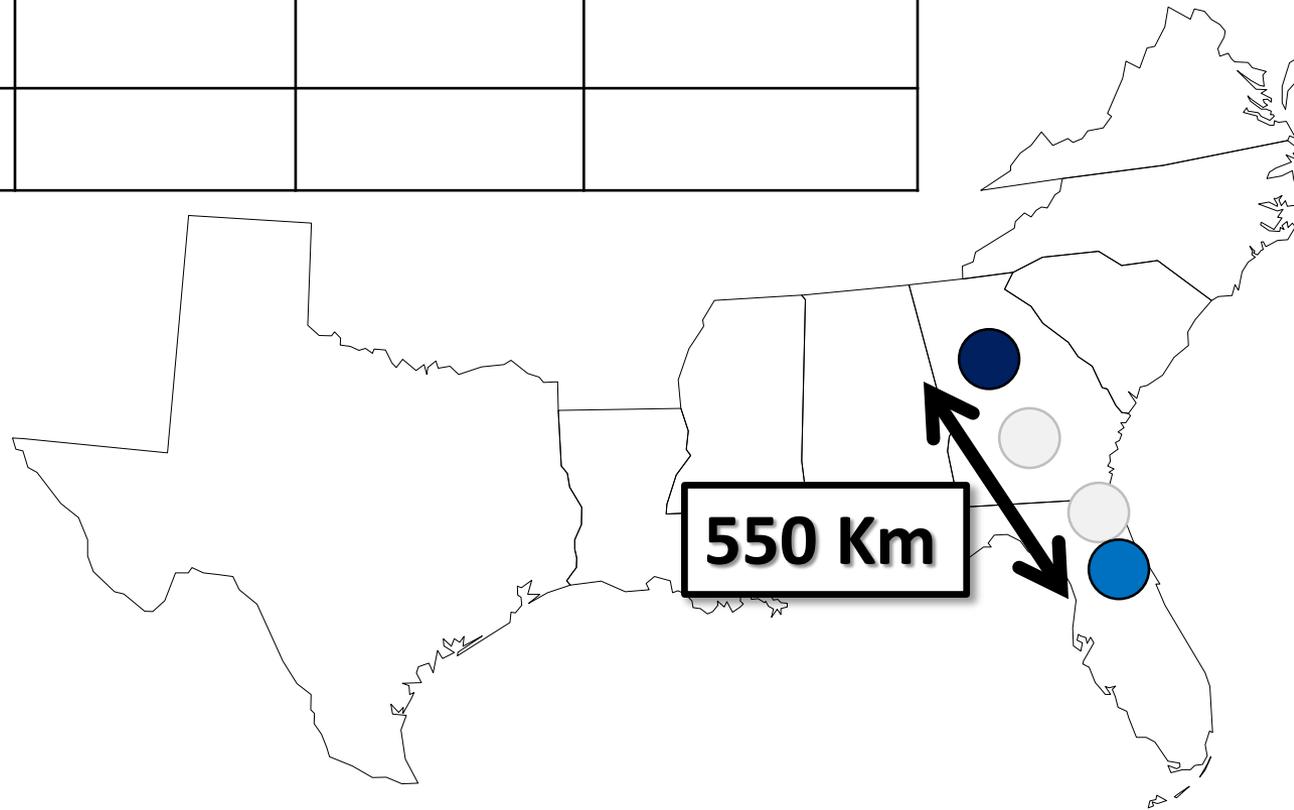
Genotype by environment interaction

DBH	Nassau	Palatka	BFGGrant	Cuttberth
Nassau	0.71	0.44		
Palatka	0.48	0.70		
BFGGrant				
Cuttberth				



Genotype by environment interaction

DBH	Nassau	Palatka	BFGGrant	Cuttberth
Nassau				
Palatka			0.11	
BFGGrant				
Cuttberth				





What are the threats to the application of genomics/genome-wide selection in forestry?

- **Genotype by environment interaction**

Prediction models apply to narrow regions

- **Low multiplier**

Many suppliers, horizontal market

- **High cost of genotyping**

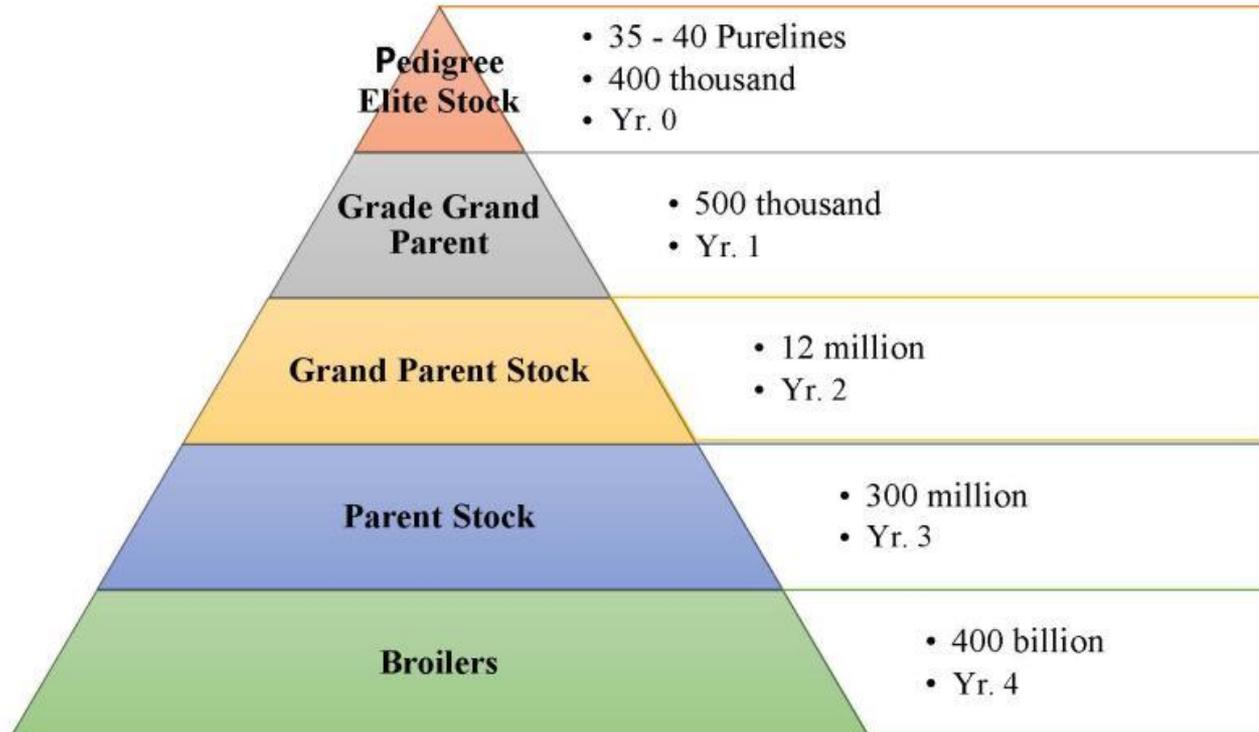
Scale and affordability

- **Low heritabilities**

Better training populations needed

- **Slow breeding**

Slow translation of science to application



What are the threats to the application of genomics/genome-wide selection in forestry?

- **Genotype by environment interaction**

Prediction models apply to narrow regions

- **Low multiplier**

Many suppliers, horizontal market

- **High cost of genotyping**

Scale and affordability

- **Low heritabilities**

Better training populations needed

- **Slow breeding**

Slow translation of science to application

High cost of genotyping – imputation or other solutions

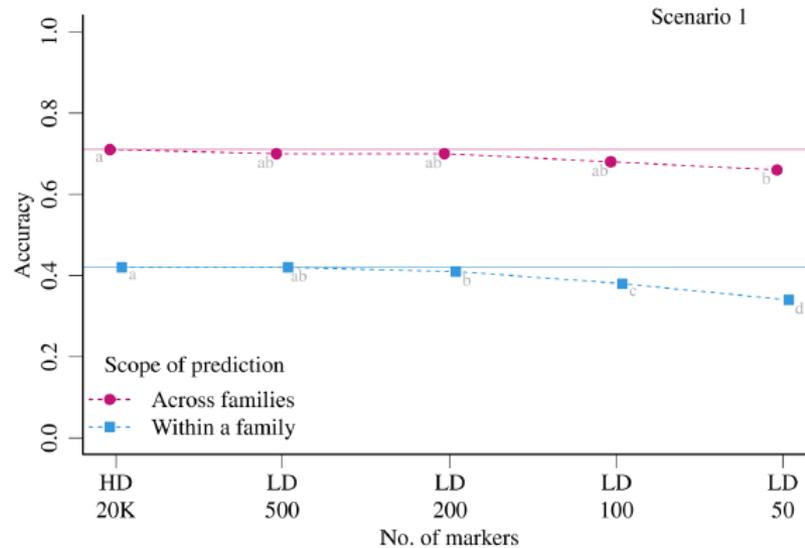
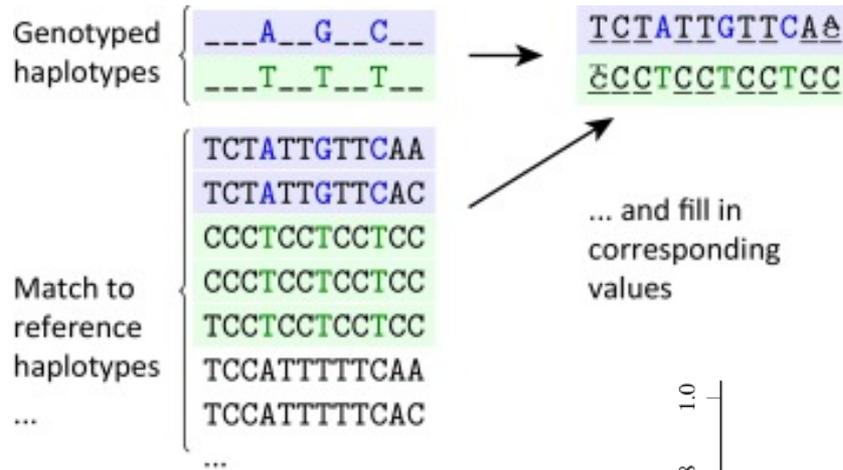


Fig. 2. Prediction accuracy across families and within a family against the number of markers used in the prediction set with 200 individuals; 2000 training individuals had the true high-density genotypes (HD); LD, low-density genotypes imputed to high density; letters denote significant difference within the scope of prediction at $p \leq 0.01$ according to the Tukey's multiple comparison test.

What are the threats to the application of genomics/genome-wide selection in forestry?

- **Genotype by environment interaction**

Prediction models apply to narrow regions

- **Low multiplier**

Many suppliers, horizontal market

- **High cost of genotyping**

Scale and affordability

- **Low heritabilities**

Better training populations needed

- **Slow breeding**

Slow translation of science to application

Are these factors really limiting?



Cooperative
Forest
Genetics
Research
Program



United States
Department of
Agriculture

National Institute
of Food and
Agriculture

Funded by the joint USDA/DOE Plant
Genomics for Bioenergy program (2013-
67009-21200)

Species	Cycle	Current Populations	Status
Slash	3	Main - 300 Elite - 60	3 rd cycle test planted 2011-12
Longleaf	1	Wild selections - 1000	Not active
Loblolly	2	Main - 150	2 nd cycle test planted 2012-13
Sand	1	Wild selections - 143	Not active
Hybrids	1	60 - F1	Non active

Time (yrs) ↓ 0

B
36

E
36

M
36

H
36

S
36

500 Selns
topgrafted
8 Sublines
CP Tests

C
36

F
36

I
36

L
36

W
36

Company 1

Company 2

Company 3

Company 4

Company 5

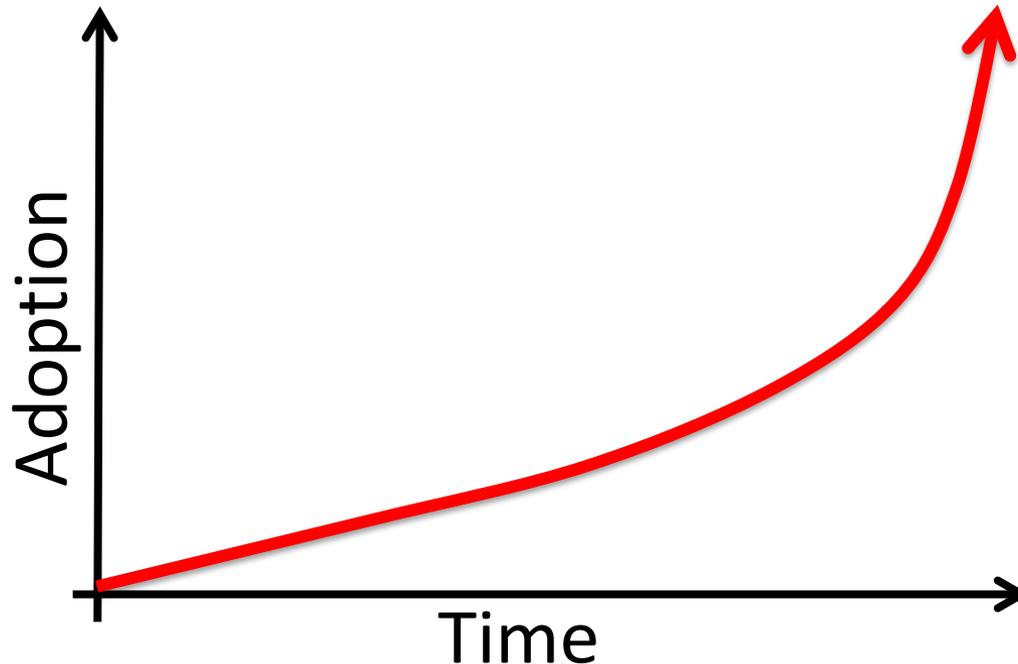
Company 6

Company 7



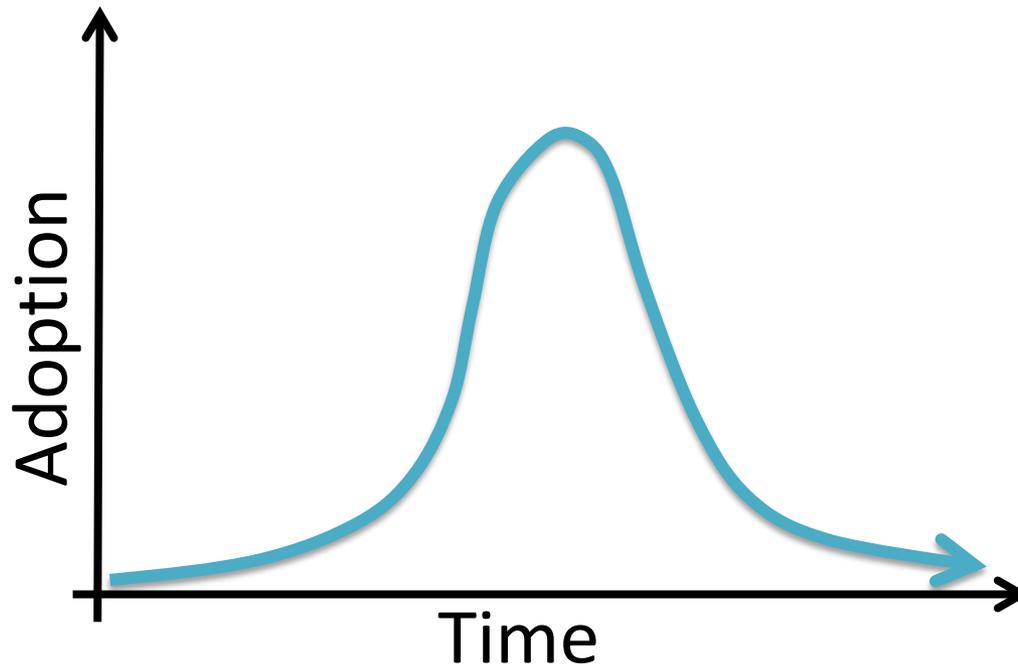
Summary

- Historical data in other systems suggests that GWS is applicable to forestry
- Preliminary data in conifers and other tree species support that expectation
- However, the uniqueness of tree breeding will impose limitations that may slow the application of GWS
- Which scenario will we find?

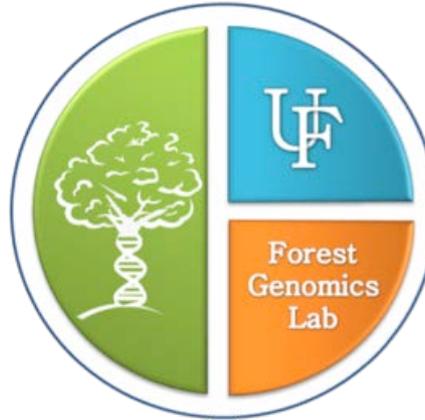


Summary

- Historical data in other systems suggests that GWS is applicable to forestry
- Preliminary data in conifers and other tree species support that expectation
- However, the uniqueness of tree breeding will impose limitations that may slow the application of GWS
- Which scenario will we find?



Acknowledgements



Chris Dervinis

Rodrigo Santos (UF - PMCB)

Janeo E. de Almeida Filho (UFV)

João F.R. Guimarães (UFV)

Gabi Nunes Silva (UFV)

Patrício Munoz (Hort. Sciences)

Esteban Rios (Agronomy)

Marcio Resende (RAPiD now UF)



Cooperative
Forest
Genetics
Research
Program

Salvador Gezan
Gary Peter
Greg Powell
Coop members



United States
Department of
Agriculture

National Institute
of Food and
Agriculture

This project was funded by the joint
USDA/DOE Plant Genomics for Bioenergy
program (2013-67009-21200)

