

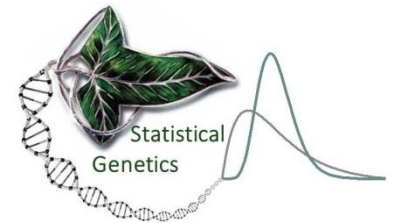
Incorporating genotype-by-environment and genomic selection information into oat breeding programs



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

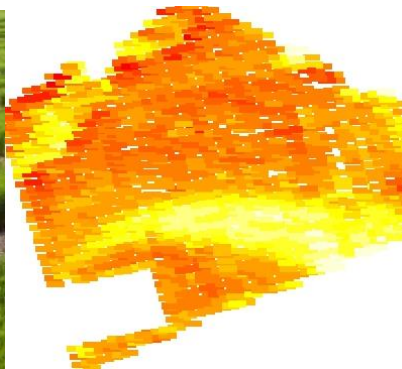
Lucía Gutiérrez
Quantitative Genetics and Cereals Breeding

Agronomy Department
University of Wisconsin - Madison
gutierrezcha@wisc.edu

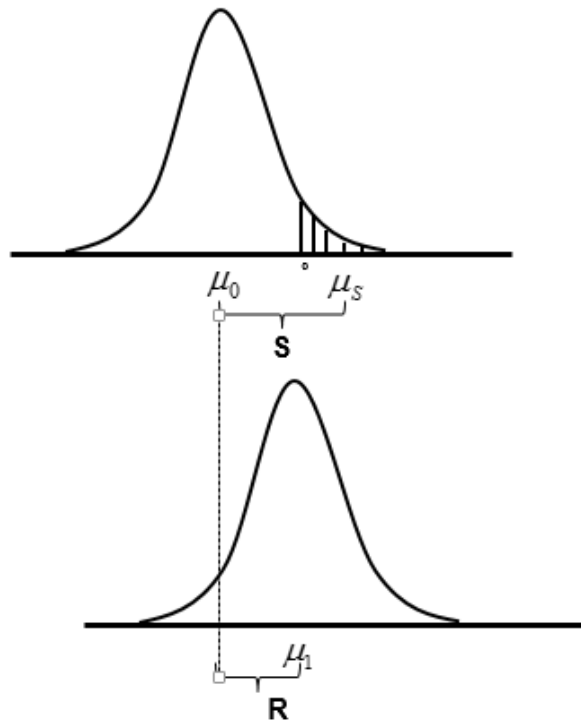


July 2016

International Oats Conference, St. Petersburg, Russia



Selection Response

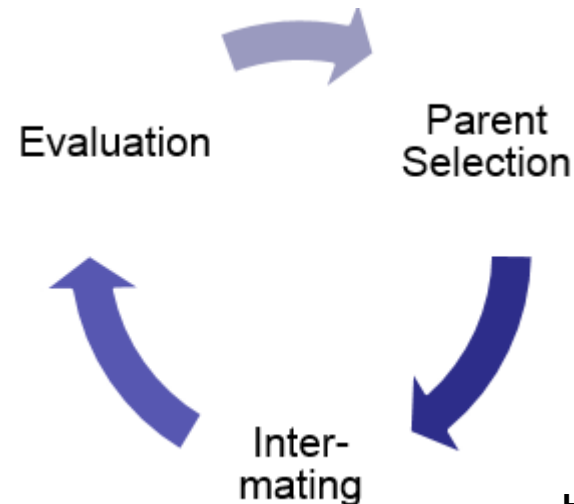


μ_0 = mean of the initial Random Mating population
 μ_s = mean of a selected individual s from R.M. population
 μ_1 = mean of a progeny of selected individual s
 c = truncation point
 S = selection differential (ΔX)
 R = response to selection (ΔY)

$$\Delta Y = b(\Delta X)$$

$$R = b_{x,y}S \quad b_{x,y} = \frac{R}{S} = h^2$$

$$\Delta G = \frac{CiV_A}{y\sigma_P}$$

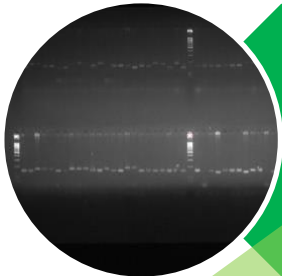


Plant Breeding



TRADITIONAL PLANT BREEDING

Relies mainly on plant
or population
phenotypic evaluations
and pedigree
information



MARKER ASSISTED SELECTION

Uses some genotypic
information to
improve selection

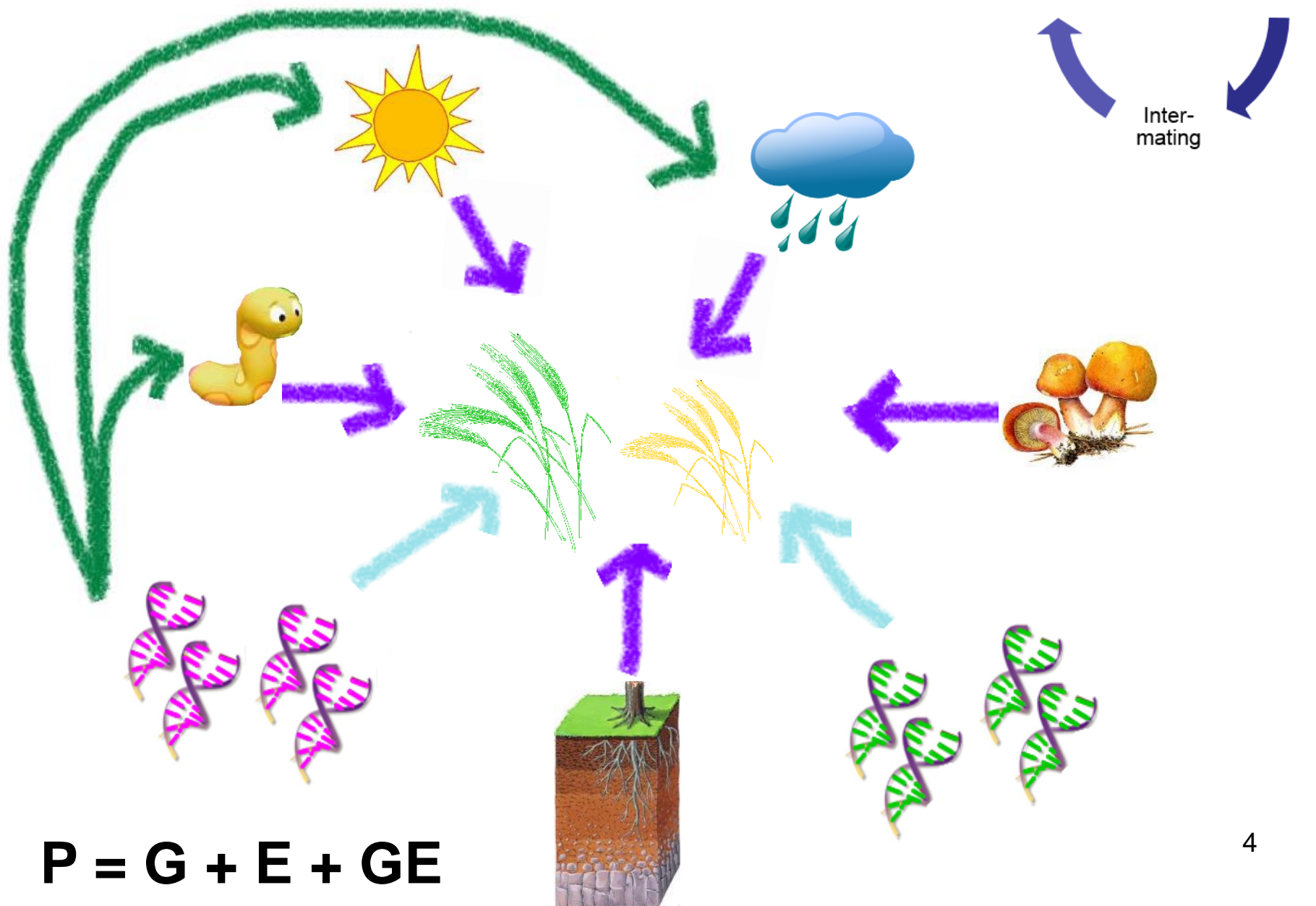


GENOMIC SELECTION

Uses a lot of
genotypic information
to improve selection

Phenotypes!

What is the Phenotype?



$$P = G + E + GE$$

GxE Interaction

$$P_{ijk} = \mu + G_i + E_j + \mathbf{GE}_{ij} + \varepsilon_{ijk}$$



Genotype 1



Genotype 2

ENV 1

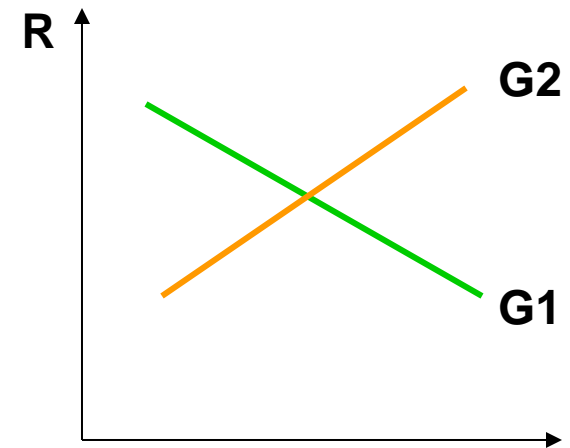


Genotype 1



Genotype 2

ENV 2



E1

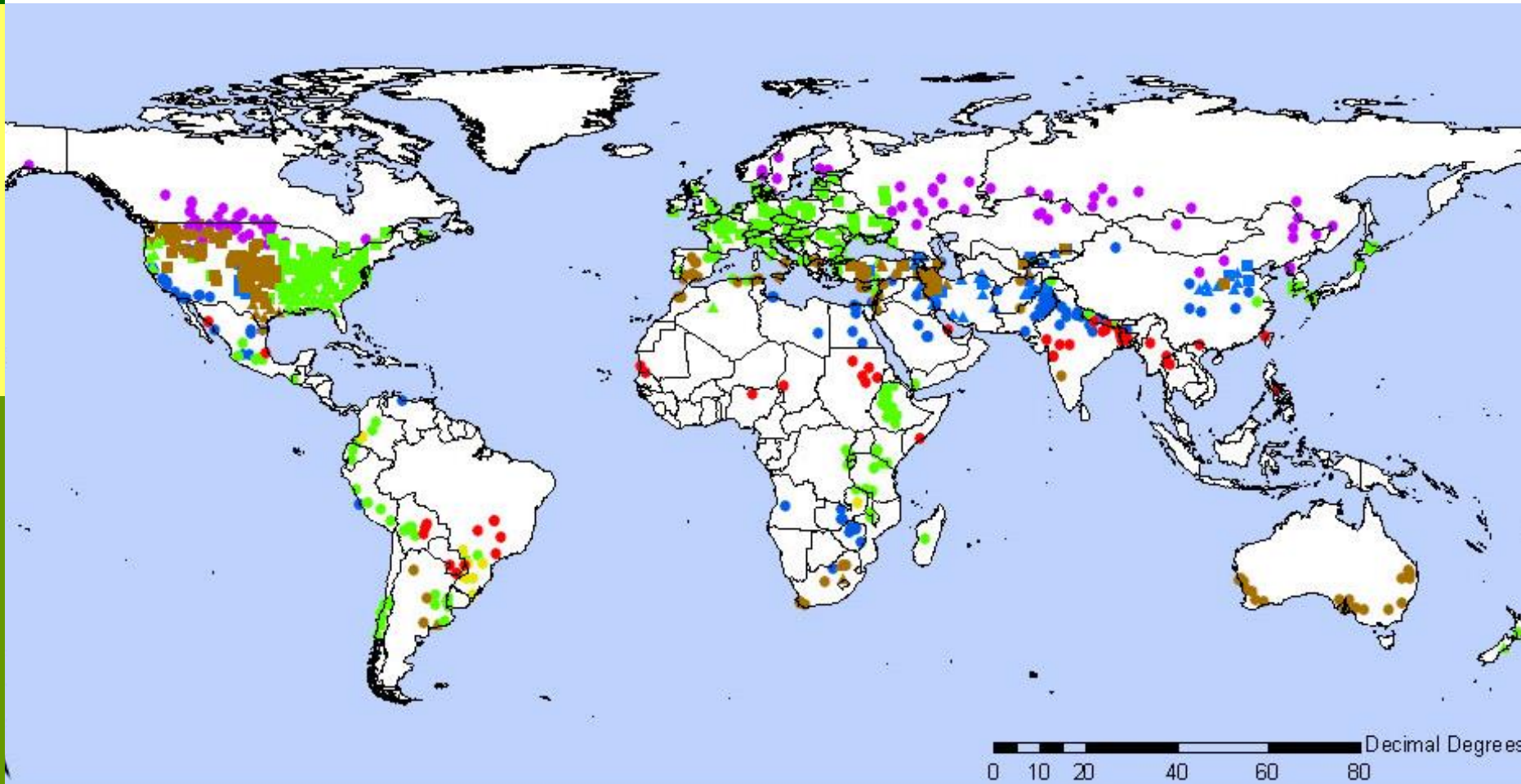


E2

Dealing with GxE

1. Ignore
2. Avoid
3. Exploit

Bernardo (2010)



QTLxE

Genomewide scan with QTLxE

$$\underline{P}_{ij} = \mu + E_j + \underline{G}_i + \underline{GE}_{ij}$$

$$\underline{P}_{ij} = \mu + E_j + x_i \alpha + \underline{G}_i^* + x_i \alpha_j^* + \underline{GE}_{ij}^*$$

$$\underline{P}_{ij} = \mu + x_i \alpha_j + \underline{\varepsilon}_{ij}$$

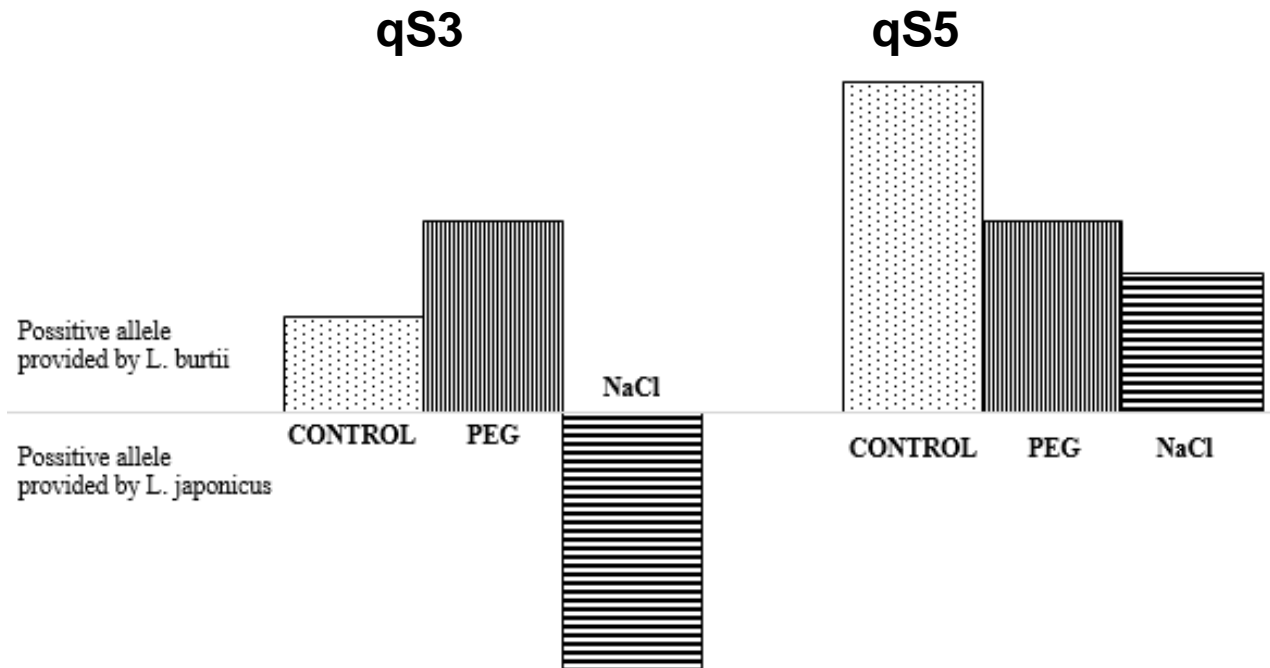
α = QTL main effect

α_j^* = QTLxE = deviation of QTL main effect

α_j = Environment - specific QTL effect

Handling GxE for QTL Mapping

QTLxE in a RIL population for saline stress in *Lotus*



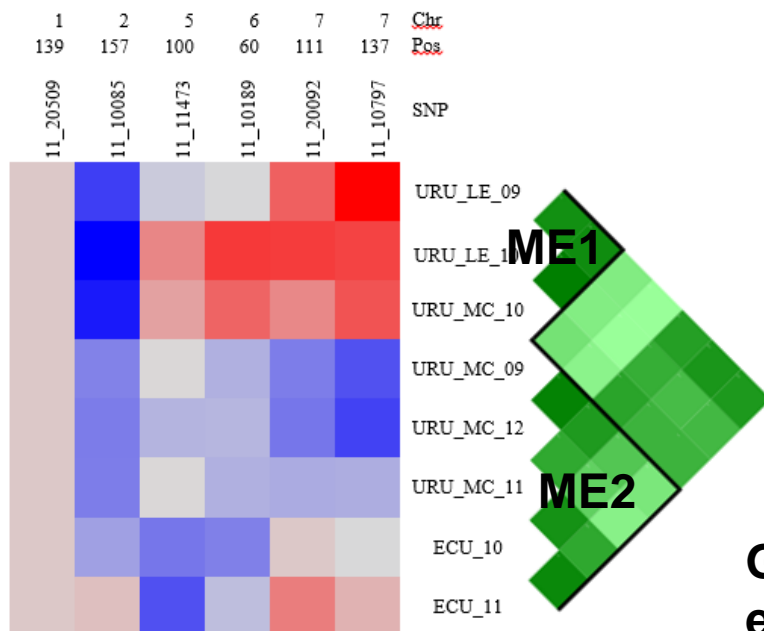
Handling GxE for QTL Mapping



www.fagro.edu.uy

QTLxE on a GWAS population for disease resistance traits in Barley

Leaf Rust



Global QTL as well as Mega-environment specific QTL were detected. Different strategies should be used for each type.

QTL for Marker Assisted Selection



www.fagro.edu.uy



Instituto Nacional de Investigación Agropecuaria
URUGUAY



Spot Blotch



Leaf Rust

Gutierrez et al., 2013 (Adv. Barley Sci.: 209-216)



Winterhardness



Malting Quality



Zitzewitz et al., 2011

(The Plant Genome 4(1): 76-91)

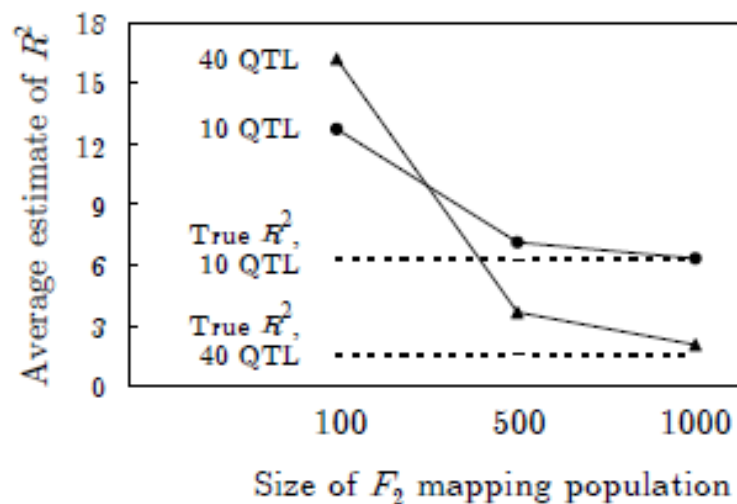
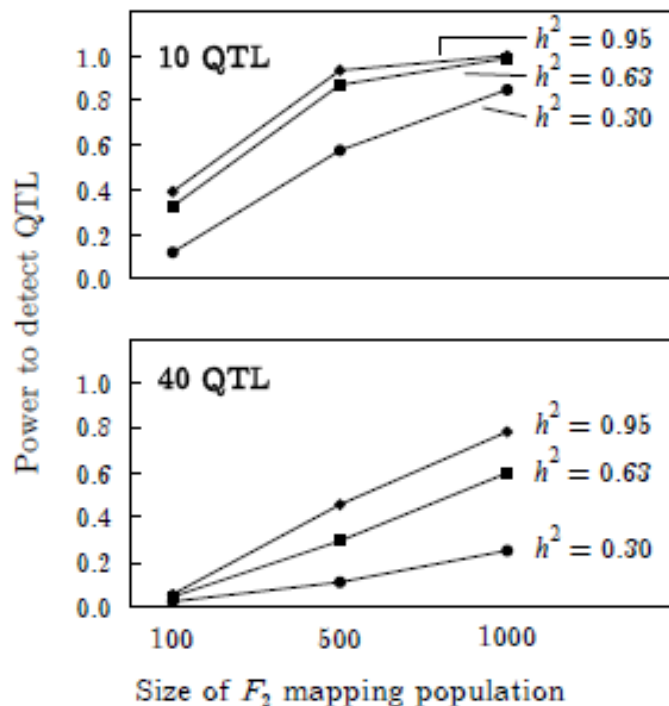
Gutierrez et al., 2011

(The Plant Genome 4:256 – 272)

Some Limitations of MAS

LIMITATIONS OF MAS (BEAVIS EFFECT)

1. Underestimation of the number of QTL
2. Over-estimation of effects



Molecular Breeding

SELECTION RESPONSE VS. UNDERLING CAUSES

Trait	Genetic control	Population used for selection	Accuracy of gene/QTL position	Molecular Breeding (MB) method
Qualitative	Few major genes	Elite X Elite	High accuracy - Ideally marker = gene - Often, marker close to gene (stable association)	MAS
		Elite x Non-elite		
Quantitative	Few QTL of large effects	Elite x Non-elite	Medium to low accuracy (CI >10cM) - Markers linked to QTL - Associations marker-QTL depend on the genetic background	MABC
		Elite x Elite		
	Numerous QTL of small effects	Elite x Elite		

Building blocks



Black box



Genomic Selection

STEPS IN GENOMIC PREDICTION

- 1. Create/choose a Training Population:** this is a group of individuals that will be used to train the model (i.e. obtain marker predictors).
Things to take into account:
 - a) Population size
 - b) Trait: h^2 , number of QTL, and trait *per se*
 - c) Population structure
 - d) Relationship to Testing Population
 - e) Markers: Number and platform
 - f) **Genotype by Environment Interaction**
- 2. Genotype and Phenotype individuals in the Training Population.**
- 3. Train the model:** use one of the methods described previously to obtain marker predictors.
- 4. Create/choose a Testing Population:** this is a set of individuals whose merit we wish to determine.
- 5. Genotype individuals in the Testing Population.**
- 6. Predict Breeding Values for the Testing Population:** use the model created with the training population and the genotypic information of the testing population to determine their merit (i.e. to get the GEBV).^{1,3}

Handling GxE in GS



Genomic Selection with GxE

$$y = 1\mu + Zg + \varepsilon$$

GBLUP_(M):

$y_{(N \times 1)}$: vector of mean performance

$g \sim N(0, A_{(N \times N)} \sigma^2_g)$

N: number of genotypes

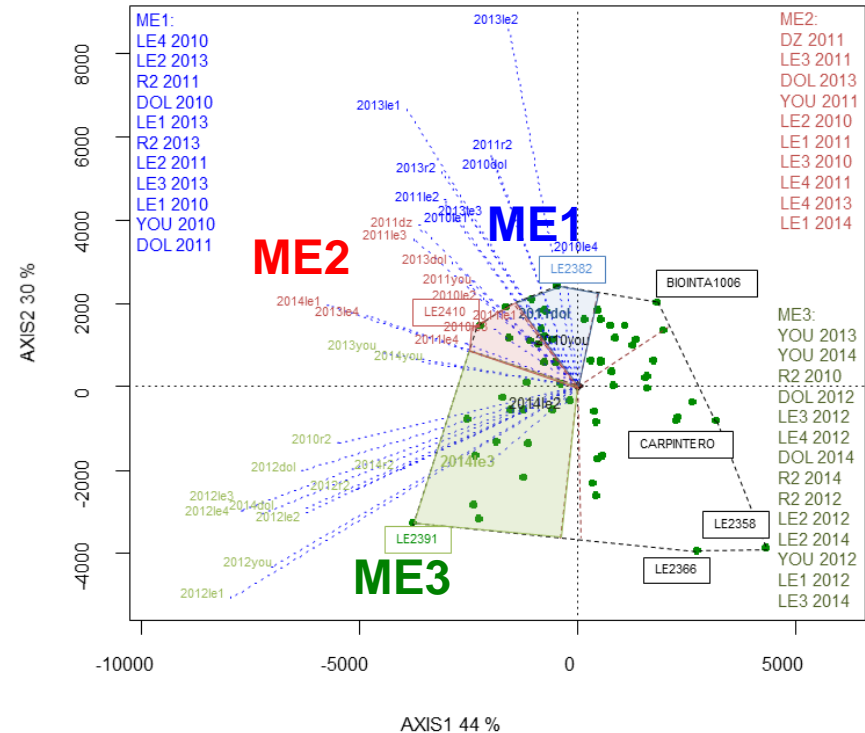
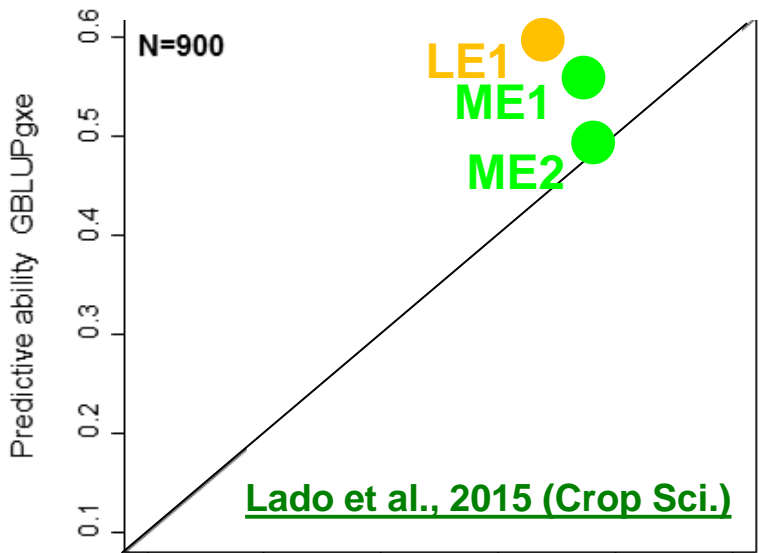
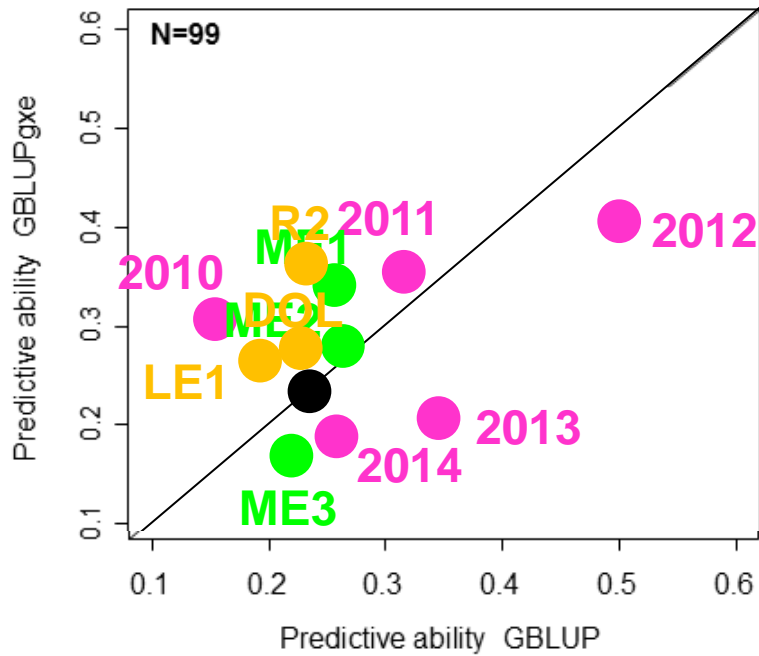
GBLUP_(gxe):

$y_{(n \times 1)}$: vector of mean performance

$g \sim N(0, A_{(N \times N)} \otimes \rho_{(k \times k)} \sigma'^2_g)$

n: number of genotypes (N) by number of environments (k)

Handling GxE for GS



Handling GxE for GS



Set	Training set	Testing Year	Testing location							
			LE		YOU		DOL		R2	
			1	2	1	2	1	2	1	2
By year [†]	All locations in testing year but one	2010	0.31	0.31	0.12	0.17	0.71	0.71	0.57	0.44
		2011	0.42	0.39	0.39	0.31	0.52	0.37	0.44	0.57
		2012	0.74	0.79	0.44	0.54	0.78	0.81	0.57	0.65
		2013	0.03	0.05	0.40	0.74	0.36	0.59	0.28	0.36
		2014	0.23	-0.14	0.30	0.19	0.58	0.71	0.30	0.19
By location [‡]	All years for a location but one	2010	0.16	0.17	0.01	-0.02	-0.06	0.04	0.01	0.11
		2011	0.23	0.18	0.27	0.34	0.08	0.14	0.19	0.29
		2012	0.50	0.31	0.10	0.04	0.33	0.29	0.05	0.15
		2013	0.23	0.10	-0.05	0.08	0.31	0.56	0	0.19
		2014	0.21	0.20	0.06	-0.05	0.08	0.36	-0.05	0.33
By mega-environment [§]	All environments for a specific mega-environment but one	2010	0.10 (1) [¶]	0.20 (1)	0.11 (1)	0.18 (1)	0.72 (1)	0.72 (1)	0.32 (3)	0.35 (3)
		2011	0.36 (2)	0.35 (2)	0.36 (2)	0.44 (2)	0.39 (1)	0.28 (1)	0.38 (1)	0.39 (1)
		2012	0.63 (3)	0.45 (3)	0.46 (2)	0.54 (2)	0.75 (3)	0.80 (3)	0.63 (3)	0.30 (3)
		2013	0.28 (1)	0.18 (1)	0.46 (3)	0.43 (3)	0.43 (2)	0.50 (2)	0.44 (1)	0.32 (1)
		2014	0.24 (2)	0.25 (2)	-0.01 (2)	0.11 (2)	0.63 (3)	0.78 (3)	-0.03 (3)	0.20 (3)

Handling GxE for GS



USING ENVIRONMENTAL COVARIATES

Year	GBLUP(M)	GBLUP(gxe)	GBLUP(gxe) + Cov	GBLUP(cov)	GBLUP(M) + Cov
2010	0.204	0.669	0.742	0.243	0.324
2011	0.469	0.652	0.688	0.236	0.237
2012	0.475	0.372	0.484	0.176	0.185
2013	0.551	0.622	0.496	0.292	0.286
2014	0.234	0.646	0.679	0.243	0.269

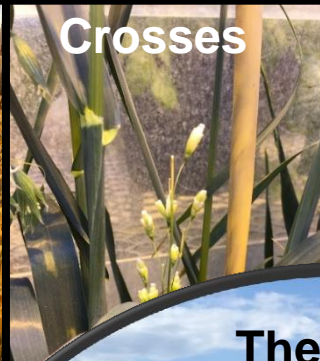
Oat Breeding at UW-Madison



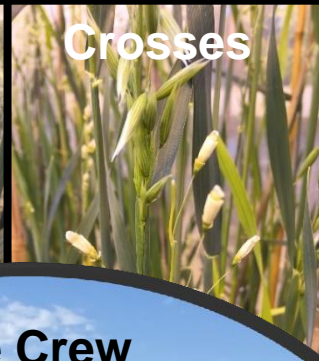
Right off my plane



Crosses



Crosses



Crosses



Increases



Increases



Planting



Planting



The Crew

Quite a sight!



2-leaf



Advanced



Elite



Uniform



Oat Breeding at UW-Madison



Madison



Marshfield



Spooner



Marshfield



Arlington



Sturgeon Bay

Oat ideas

Use the Information provided on T3:

- CORE
- POGI
- BREEDING PROGRAMS

GBS Data (Nick Tinker)

Historical Phenotypic data (highly unbalanced)

Good GS predictions (Jean-Luc Jannink)

GxE characterization

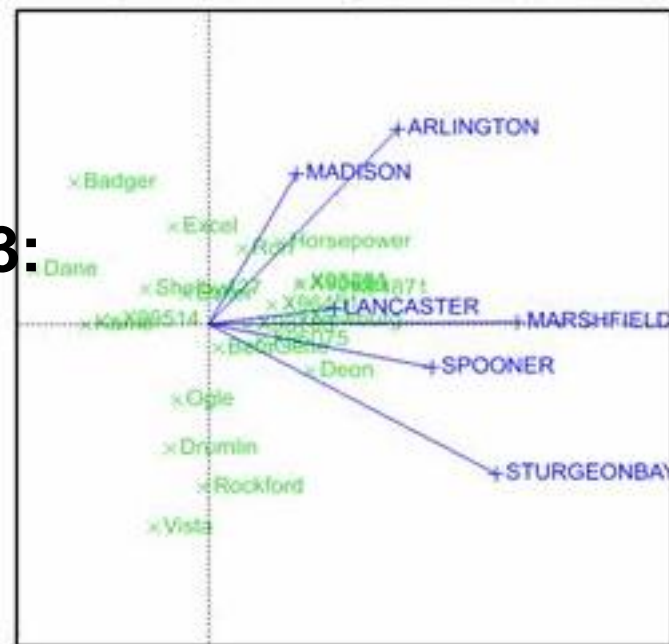
Strategic Phenotyping

Strategic Phenotyping

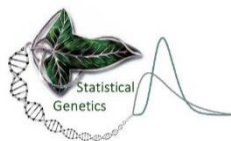
Devise a training GLOBAL Oats data-set

Use GxE to enhance prediction ability

PC2 - 18.96%



Funding and Acknowledgement



Statistical Genetics:

Gaston Quero
Sofia Brandariz
Bettina Lado
Ines Berro
Pablo Gonzalez
Agustin Gonzalez
Alejandra Borges
Pablo Sandro
Natalia Berberian
Juan Rosas



Barley Breeding:

Ariel Castro
Silvia German



Instituto Nacional de Investigación Agropecuaria
URUGUAY

Rice Breeding:

Alvaro Roel
Gonzalo Zorrilla
Marco Dalla Rizza
Victoria Bonnacarrere
Pedro Blanco
Fernando Pérez de Vida
Schubert Fernández
Silvia Garaycochea

Wheat Breeding:

Martin Quincke
Paula Silva
Silvia German

Animal Breeding:

Ignacio Aguiar



Fred van Eeuwijk
Marcos Malosetti



Pat Hayes
Alfonso Cuesta-Marco



Cornell University

Jean-Luc Jannink
Susan McCouch
Eliana Monteverde



WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON



Funding:



Monica Balzarini
Cecilia Bruno
Andrea Pena



Jesse Poland

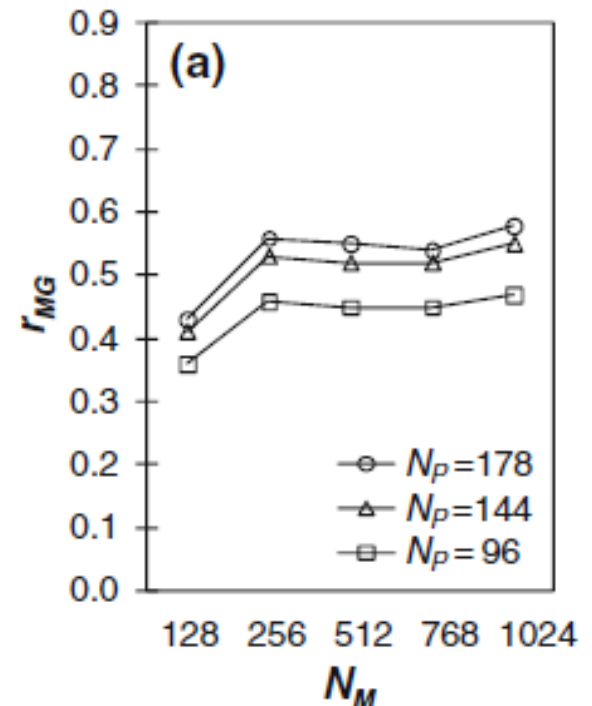
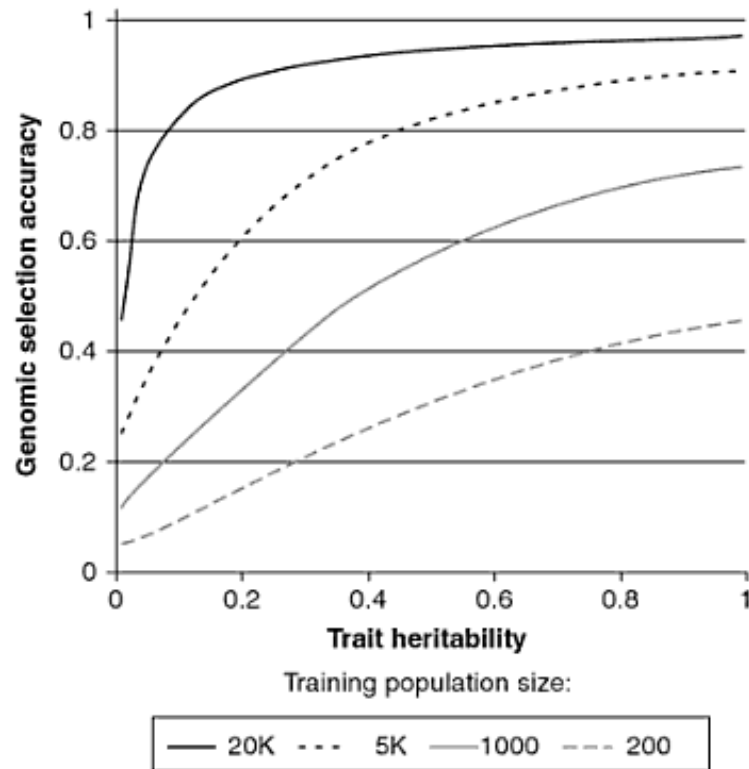


Gustavo de los Campos
Ana Ines Vazquez

Genomic Selection

A) POPULATION SIZE AND B) TRAIT h^2 :

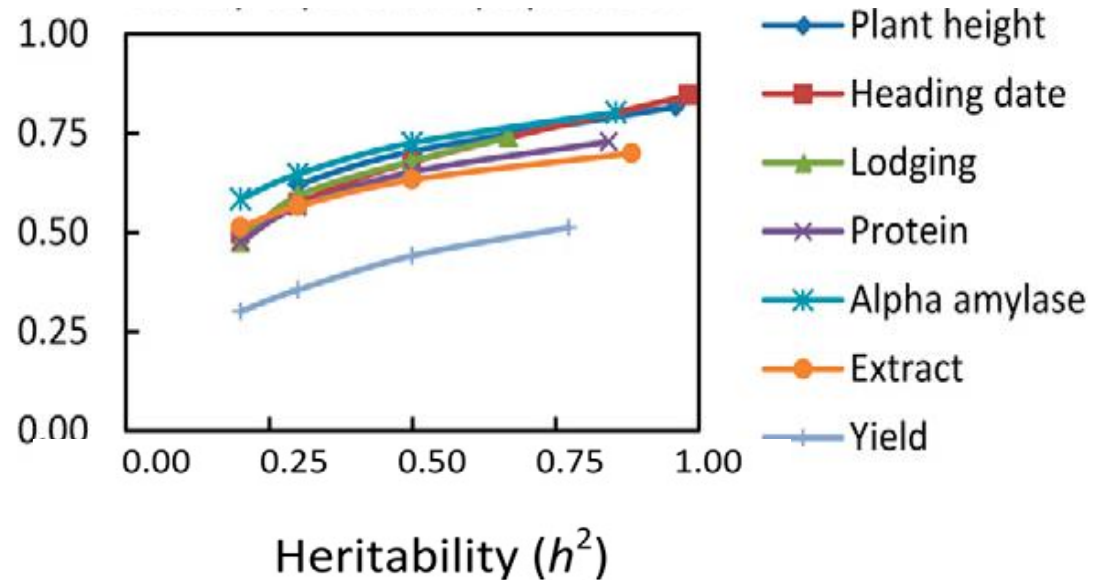
Larger population sizes and higher heritabilities increase genomic selection accuracy



Genomic Selection

B) TRAIT: *per se*

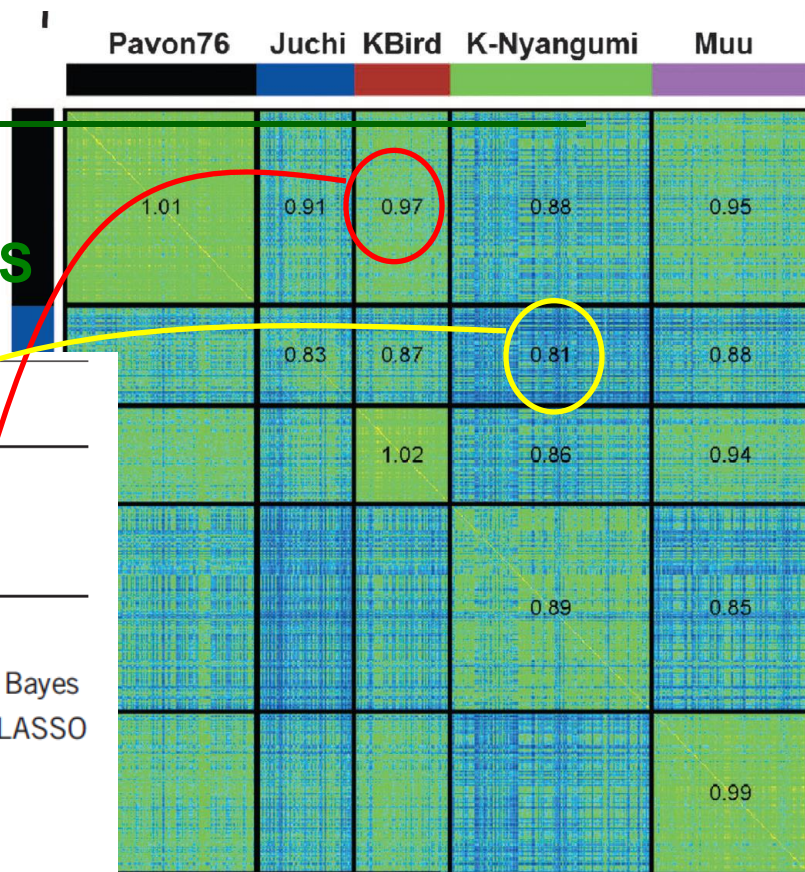
The trait is relevant not only for its heritability, but for the trait *per se*.



150 DH from Steptoe x Morex with 233 Markers
(Hayes et al., 1993)

Genomic Selection

C) POPULATION STRUCTURE AND D) RELATIONSHIP TO INDIVIDUALS



Training ^a						
	PBW343/ Juchi	PBW343/ Kingbird	PBW343/K- Nyangumi	PBW343/ Muu	PBW343/ Pavon76	
Testing						
PBW343/ Juchi	—	0.48	0.14	0.28	0.31	Bayes LASSO
PBW343/ Kingbird	0.53	—	0.29	0.25	0.54	
PBW343/ K-Nyangumi	0.14	0.30	—	0.28	0.28	
PBW343/ Muu	0.18	0.30	0.33	—	0.29	
PBW343/ Pavon76	0.37	0.51	0.22	0.33	—	

GBLUP

5 wheat populations with N= 92, 176, 148, 90, 176 for stem rust with DARs

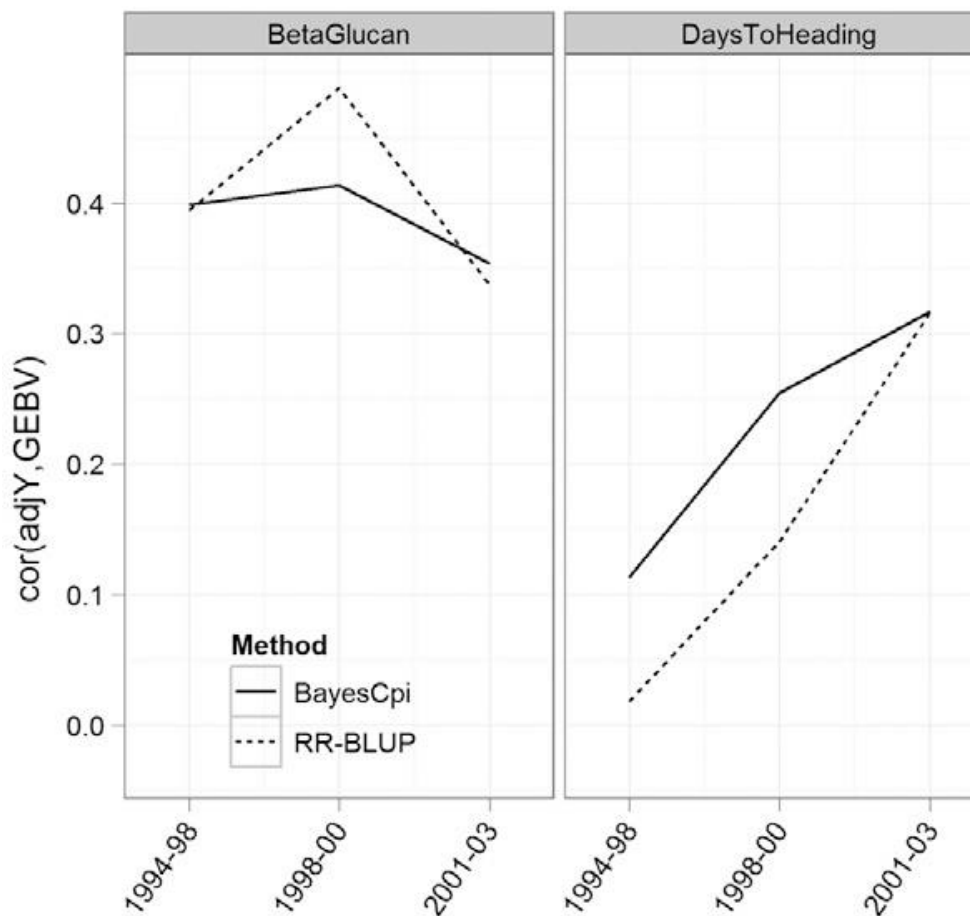
Genomic Selection

C) POPULATION STRUCTURE AND D) RELATIONSHIP TO INDIVIDUALS

Older training populations have smaller or no prediction accuracy difference.

Related training and testing populations have larger prediction accuracy.

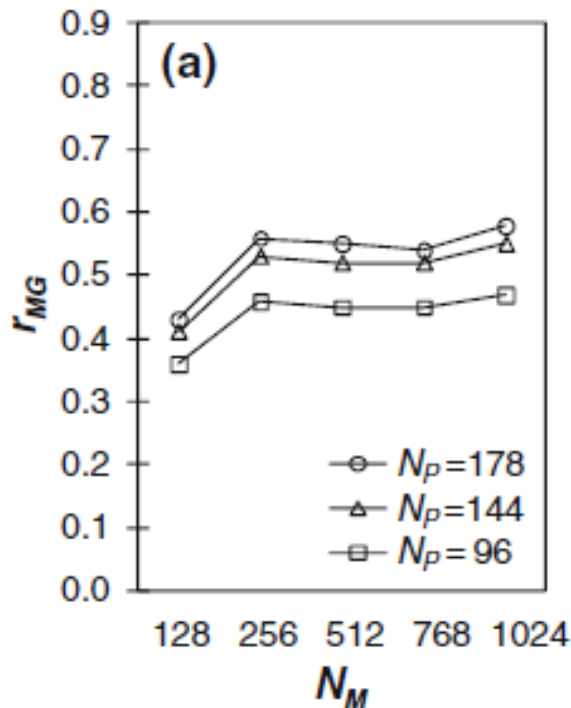
Mixed training populations have lower accuracy unless mixed populations are larger.



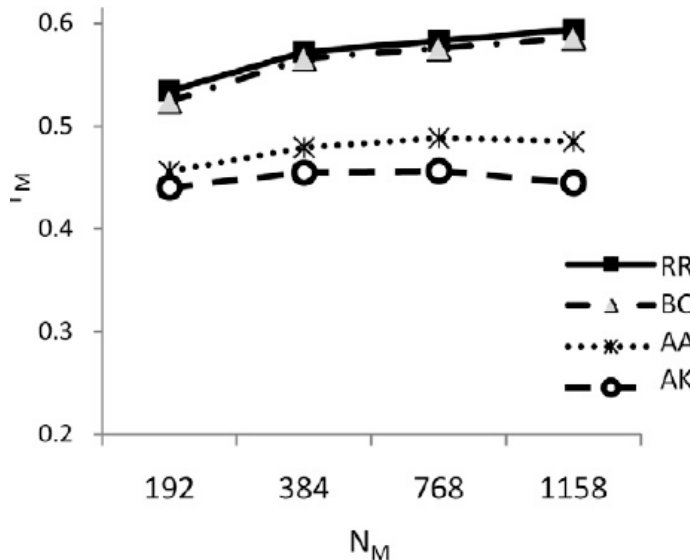
421-446 Oat lines, 1005 DArT markers

Genomic Selection

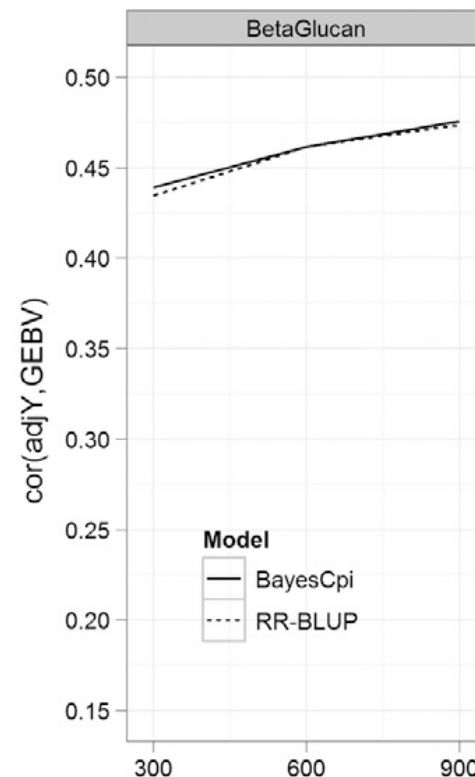
E) NUMBER OF MARKERS



223 RIL maize BM-TC1 test-crossed 1339 SSR or RFLP



374 winter wheat F5 lines with 5000 DArTs



421-446 Oat lines, 1005 DArT markers

Genomic Selection

PREDICTION METHOD

Prediction method (BayesC, RR-BLUP) did not affect accuracy as much as trait, marker density, and training population size, depth (increasing population size by including older lines increased selection accuracy) and age (older populations have lower accuracy for some traits).

Source of variation	df	Marker density	Training population size	Training population depth	Training population age
Trait [†]	4	<0.0001	<0.0001	<0.0001	<0.0001
Method [‡]	1	0.22	0.02	0.21	0.06
Design [§]	2	<0.0001	<0.0001	<0.0001	<0.01
Trait × method	4	0.03	0.14	0.56	0.31
Trait × design	8	<0.01	<0.01	<0.0001	<0.01
Method × design	2	0.26	0.11	0.64	0.52
Error	8				
Total	29				

421-446 Oat lines, 1005 DArT markers