Genomic Selection for Resistance to Fusarium Head Blight

Mark E. Sorrells, Jessica Rutkoski, Jared Benson and Gina Guedira-Brown

USDA
United States Department Of Agriculture
Agricultural Research Service

Cornell University
Department of Plant Breeding & Genetics
Celebrating 100 Years of Crop Improvement

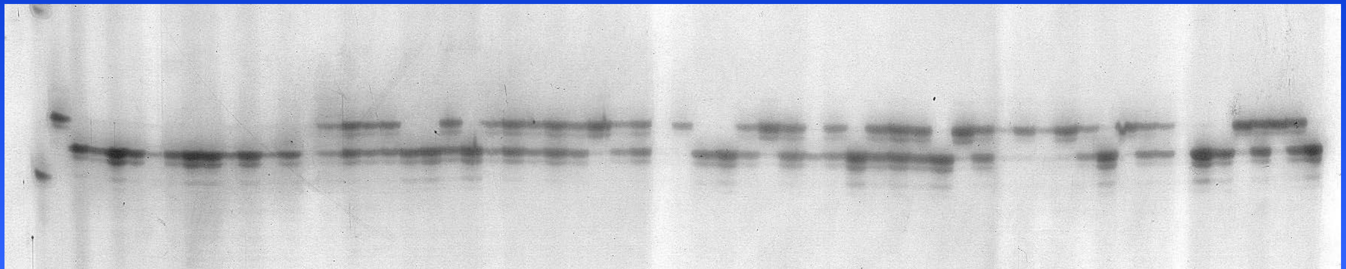
Strategies for Molecular Breeding

- Marker Assisted Selection
  - Only significant markers are used for selection, usually qualitative traits
- Association Breeding (Breseghello & Sorrells 2006)
  - Association mapping methods are superimposed on a conventional hybridization/selection/testing program for allele discovery and validation
- Genomic Selection (Meuwissen, Hayes & Goddard 2001)
  - Genome-wide markers that are used to develop prediction models for estimating genomic breeding values

Marker Assisted Selection

Significant impacts in backcrossing

- Simple, monogenic trait improvement
  - i.e. BC major genes into elite varieties
  - Gene pyramiding
  - Forward selection of multiple major genes



FHB QTL

(Buerstmayer, Ban & Anderson 2009)

- 100 QTL on all wheat chromosomes have been reported
- 22 QTL regions on 16 chromosomes have been detected in more than 1 population
- The most repeatable QTL are those on chromosomes 3BS (Fhb1), 5AS (Qfhs.ifa-5A) and 6BS (Fhb2)
- Only Fhb1 has reliable diagnostic markers

Phenotyping Fusarium Head Blight Resistance

Evaluation for FHB resistance:

- Design: RCB 4-5 replicates, single 1 M row
- Plots are inoculated with a spore suspension 3 times starting at flowering time (inoculum provided by Gary Bergstrom's lab)

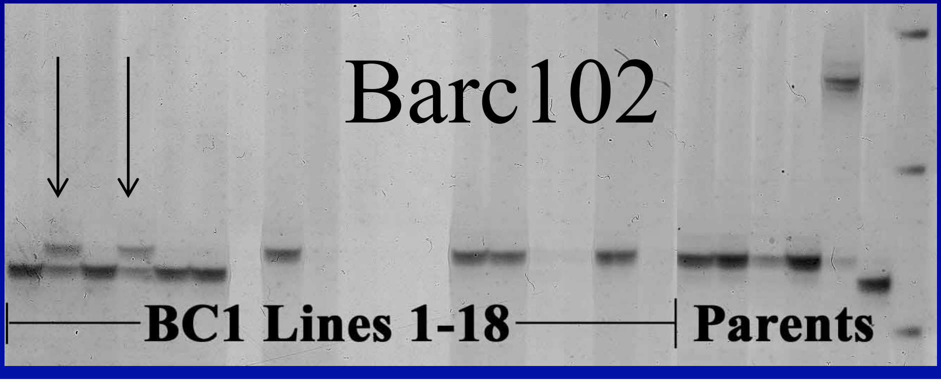
- ~3 weeks post-flowering we count 20 spikes in each plot and score each spike 0-5 for severity

- Incidence: convert #spikes with symptoms to percent
- Severity: calculate average severity for spikes with symptoms and convert to percent

Issues:

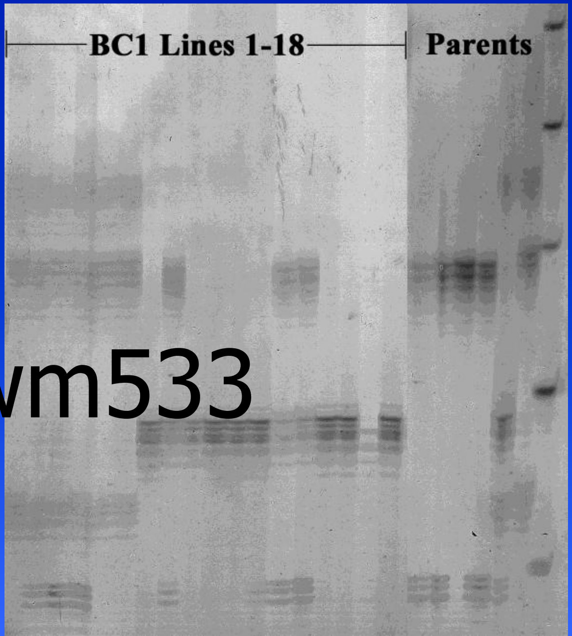
- Timing of scoring
- Confounding with other diseases
- Confounding with heading date
- Secondary tillers

Molecular markers linked to Fusarium Head Blight Resistance




Sorrells Benscher Bergstrom

Cooperative Fusarium Head Blight Trial

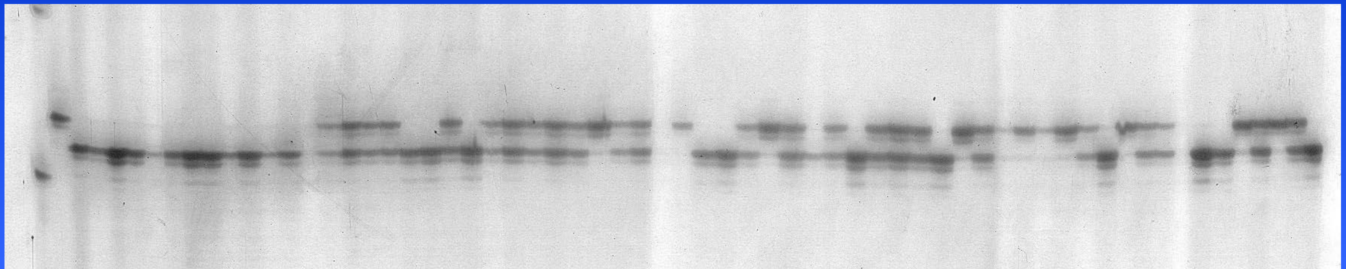


Marker Assisted Selection

Limitations:

- Best suited for major genes
  - BC is the most conservative breeding method
  - Pyramiding limited to a few target genes

Genes with small effects that underlie most of the important traits determine the success of new varieties



Genomic Selection Methodology

Meuwissen et al. 2001 Genetics 157:1819-1829; Goddard & Hayes 2007

A Training Population is genotyped with a large number of markers and phenotyped for important traits

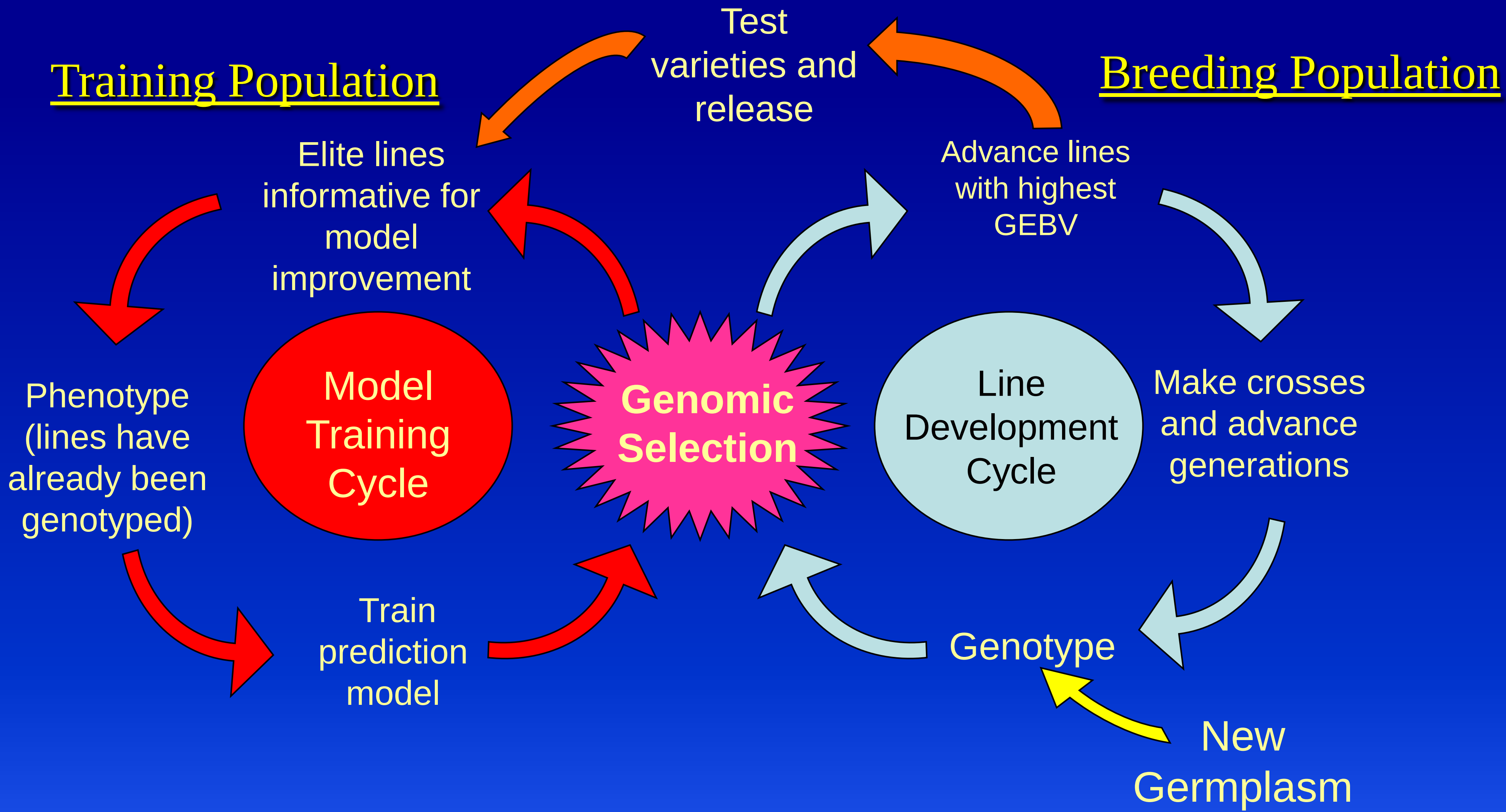
- Genome-wide markers are considered to be random effects
- The magnitude of each marker's effect on the phenotype is estimated simultaneously
- One or more markers are assumed to be in LD with each QTL affecting trait
- Prediction model attempts to captures the total additive genetic variance to estimate breeding value of individuals based on sum of all marker effects

In a Breeding Population individuals are genotyped but not phenotyped

- A genomic estimated breeding value (GEBV) for each individual is obtained by summing the marker effects for that genotype
- Prediction model is used to impose multiple generations of selection

GS in a Plant Breeding Program

Heffner, Sorrells & Jannink. Crop Science 49:1-12



Test varieties and release

Training Population

Breeding Population

Elite lines informative for model improvement

Advance lines with highest GEBV

Phenotype (lines have already been genotyped)

Model Training Cycle

Genomic Selection

Line Development Cycle

Make crosses and advance generations

Train prediction model

Genotype

New Germplasm

Genomic selection reduces cycle time & cost by reducing frequency of phenotyping

Genomic Selection for FHB Resistance

Jessica Rutkoski and Jared Benson

Dataset from 2008, 2009, 2010 Cooperative FHB Nurseries previously used for association mapping

Photos Courtesy of Marcia McMullen

Photo Courtesy of Jim Miller

Materials and Methods

Traits: *Fusarium* head blight Incidence and Severity

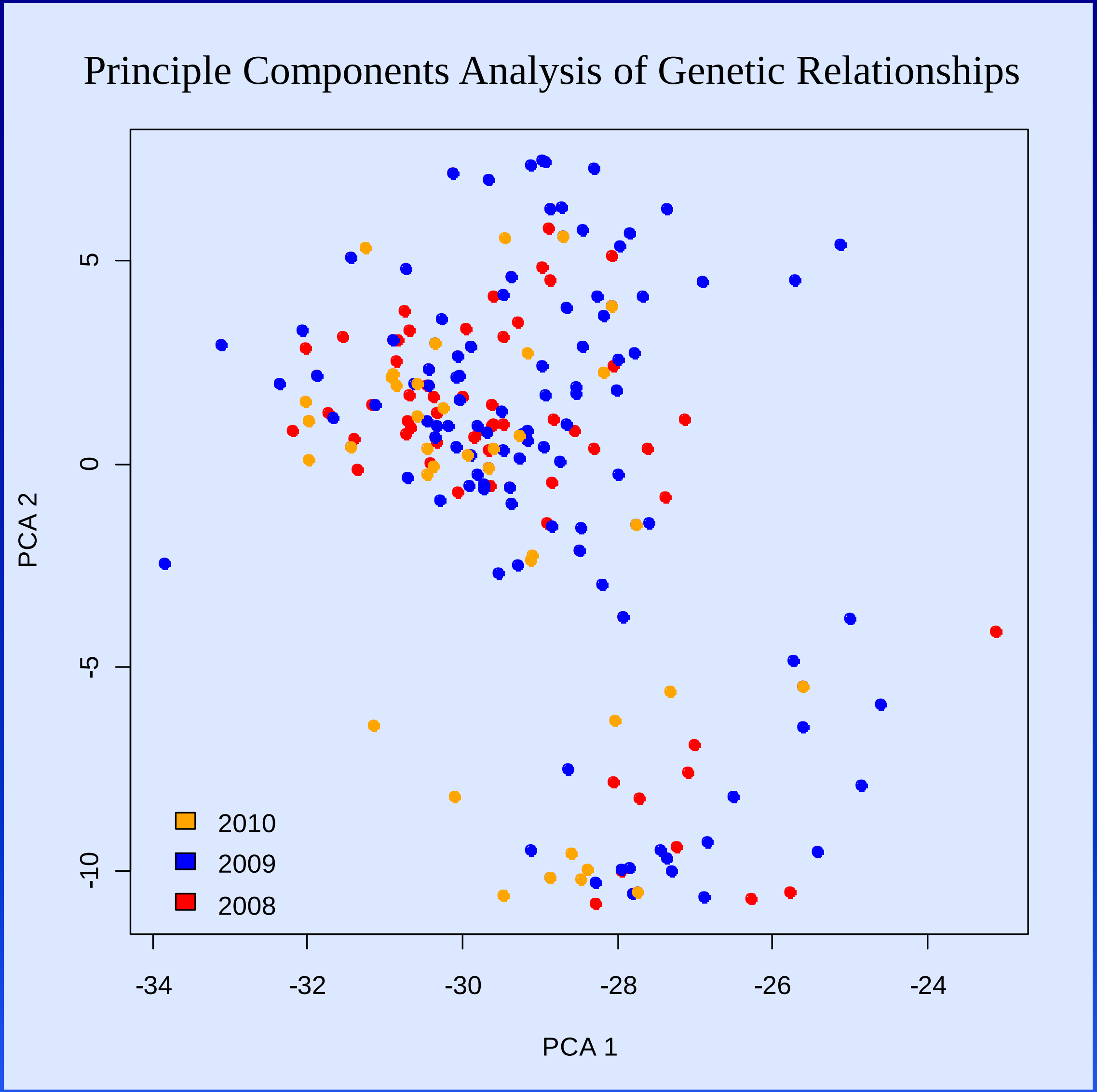
Nurseries, years, and environments:

Genotypes: Unique in each year-nursery combination

Markers: 2,402 DArT markers

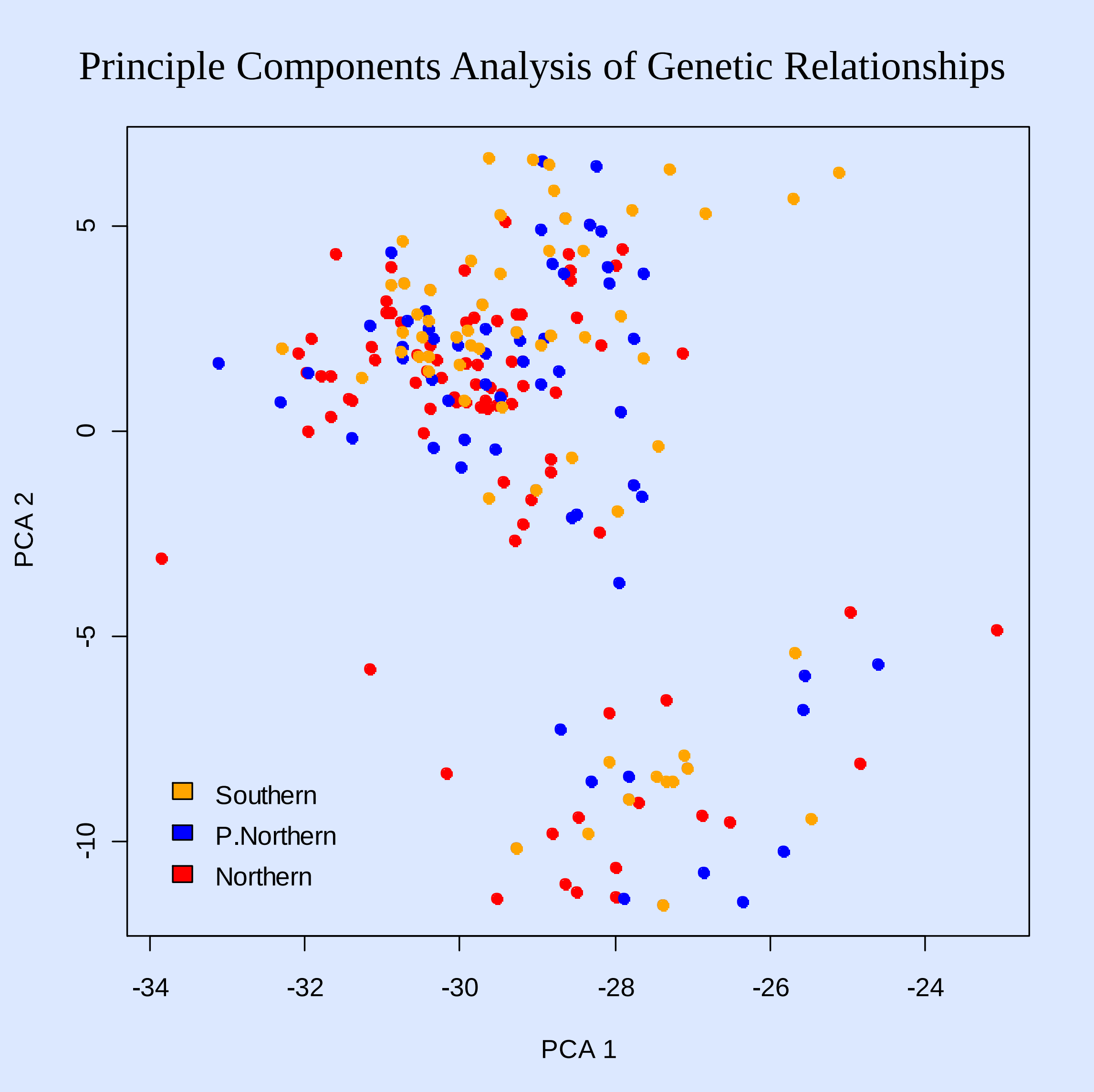
Year	2010	2009			2008		
Nursery	Northern	P. Northern	Southern	Northern	P. Northern	Southern	Northern
Locations	INLAY	INLAY	ILURB	MIELA	ILURB	NCKIN	INLAF
	MOCOL	VABLA	NCKIN	MDSAL	MOCOL	MOCOl	ILURB
	ILURB	MIELA	VABLA	ILURB	MIELA	VABLA	MIELA
	MIELA	ILURB	MOCOL	VABLA	VABLA	MDSAl	NYITH
	VABLA	MOCOL	MDSAL	NYITH	INLAF	ILURB	MOCOL
	INBRO	INBRO	INBRO	MOCOL			OHWOO
	OHWOO			OHWOO			MDCLA
	NYITH			INBRO			VABLA
	X2010N						
	MDSAL						
No. of lines	38	109			85		

Materials and Methods



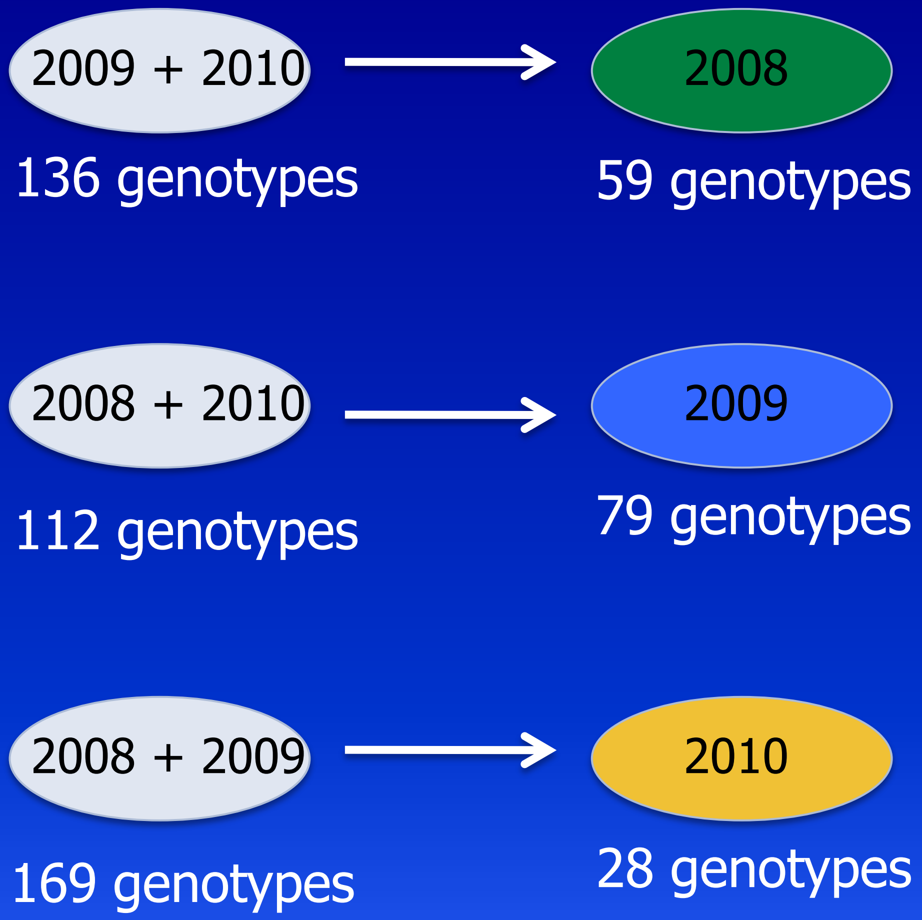
No apparent population structure differentiating the years

Materials and Methods



No apparent population structure differentiating the nurseries

Materials and Methods



2009 + 2010 → 2008

136 genotypes → 59 genotypes

2008 + 2010 → 2009

112 genotypes → 79 genotypes

2008 + 2009 → 2010

169 genotypes → 28 genotypes

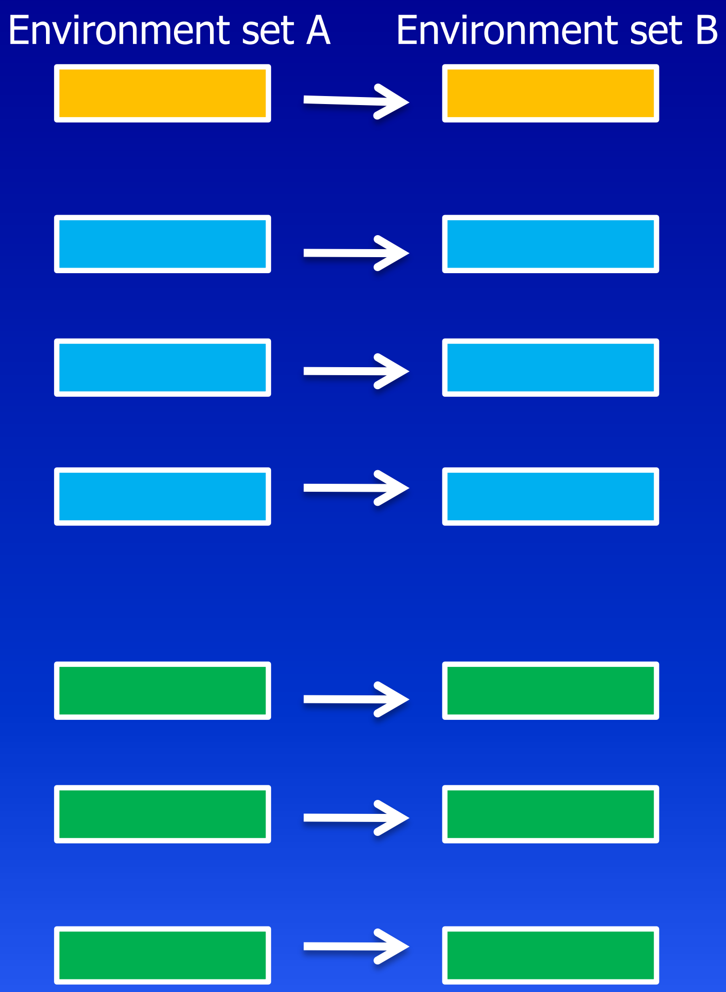
<u>Validation</u>

- **Training set**= 2 years of data
- **Validation set**= remaining year
- **No overlapping genotypes**

Materials and Methods

Accuracy of phenotypic selection (*rP*) calculation:

Calculated within nursery and within year across locations. *rGS/rP* values for each nursery were averaged within years



Year	**Nursery**	**Trait**	***rP***
2010	Northern	Incidence	0.74
		Severity	0.85
2009	Northern	Incidence	0.57
		Severity	0.81
2009	Southern	Incidence	0.66
		Severity	0.81
2009	P. Northern	Incidence	0.56
		Severity	0.64
2008	Northern	Incidence	0.68
		Severity	0.48
2008	Southern	Incidence	0.65
		Severity	0.64
2008	P. Northern	Incidence	0.24
		Severity	0.59

Materials and Methods

Statistical Models		
Model	**Marker variance Assumption**	**Variable selection**
Ridge regression (RR-BLUP)	Equal marker variances	No- all markers included
Bayes Cπ	Equal marker variances	Yes- markers with effects close to 0 are excluded with probability π π is estimated using training population data

Materials and Methods

Genomic estimated breeding value (GEBV) $\xrightarrow{rGS}$ True breeding value (TBV) $\xleftarrow{h}$ Phenotype (TBˆVs)

Accuracy measurements:

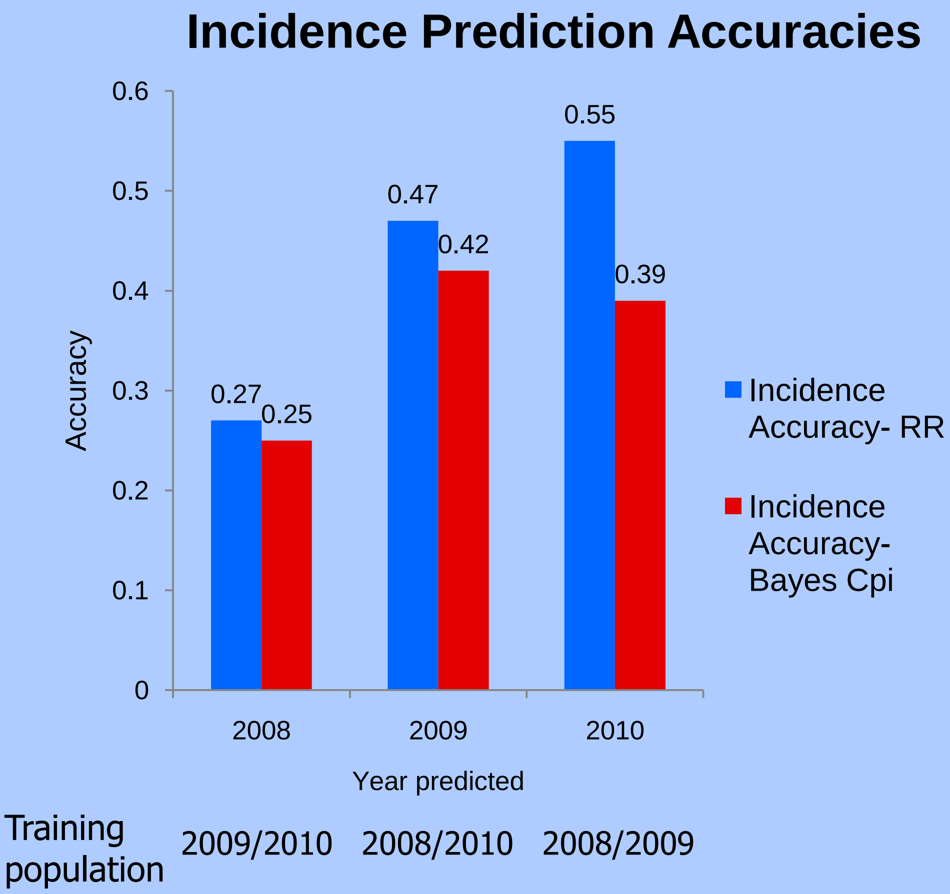
1) **GS Accuracy**= Correlation of the GEBVs and the TBˆVs

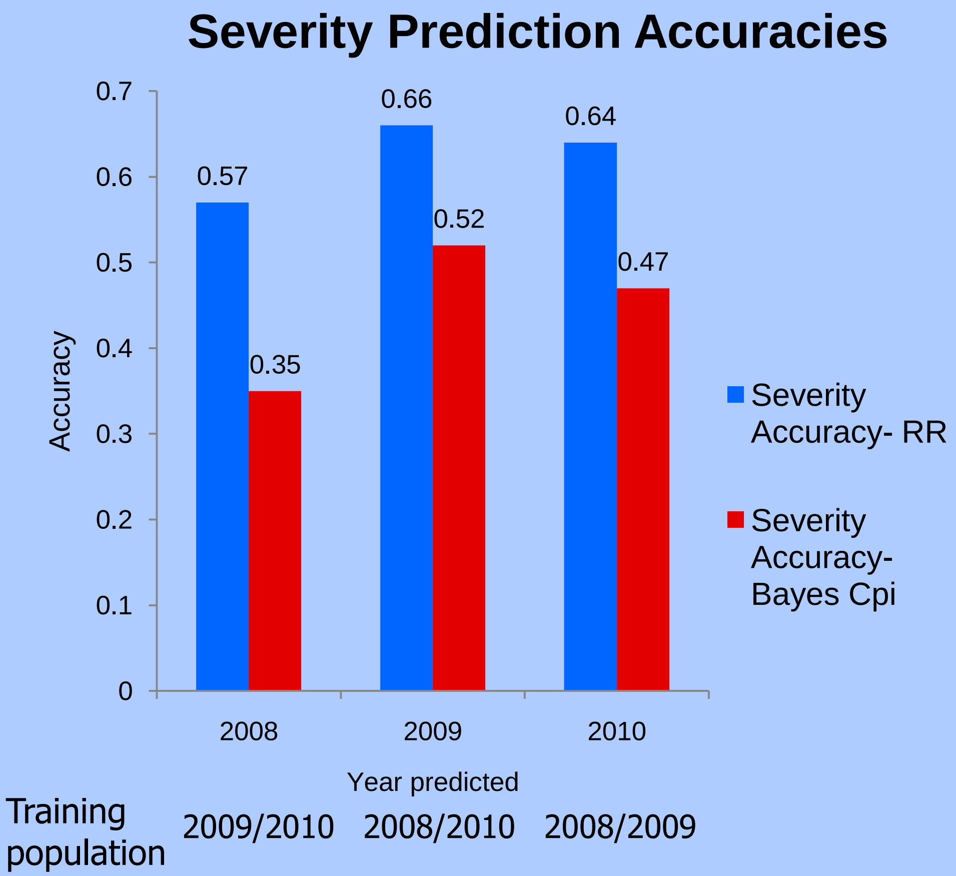
Cor(GEBVs, TBˆVs) = $\hat{r}$

2) **Corrected GS accuracy** = $\hat{r}/H$ of the validation set (Dekkers, 2007)

Correcting $\hat{r}$ by $1/H$ gives us a closer estimate of r, the correlation of the genomic estimated breeding values and the true breeding values.

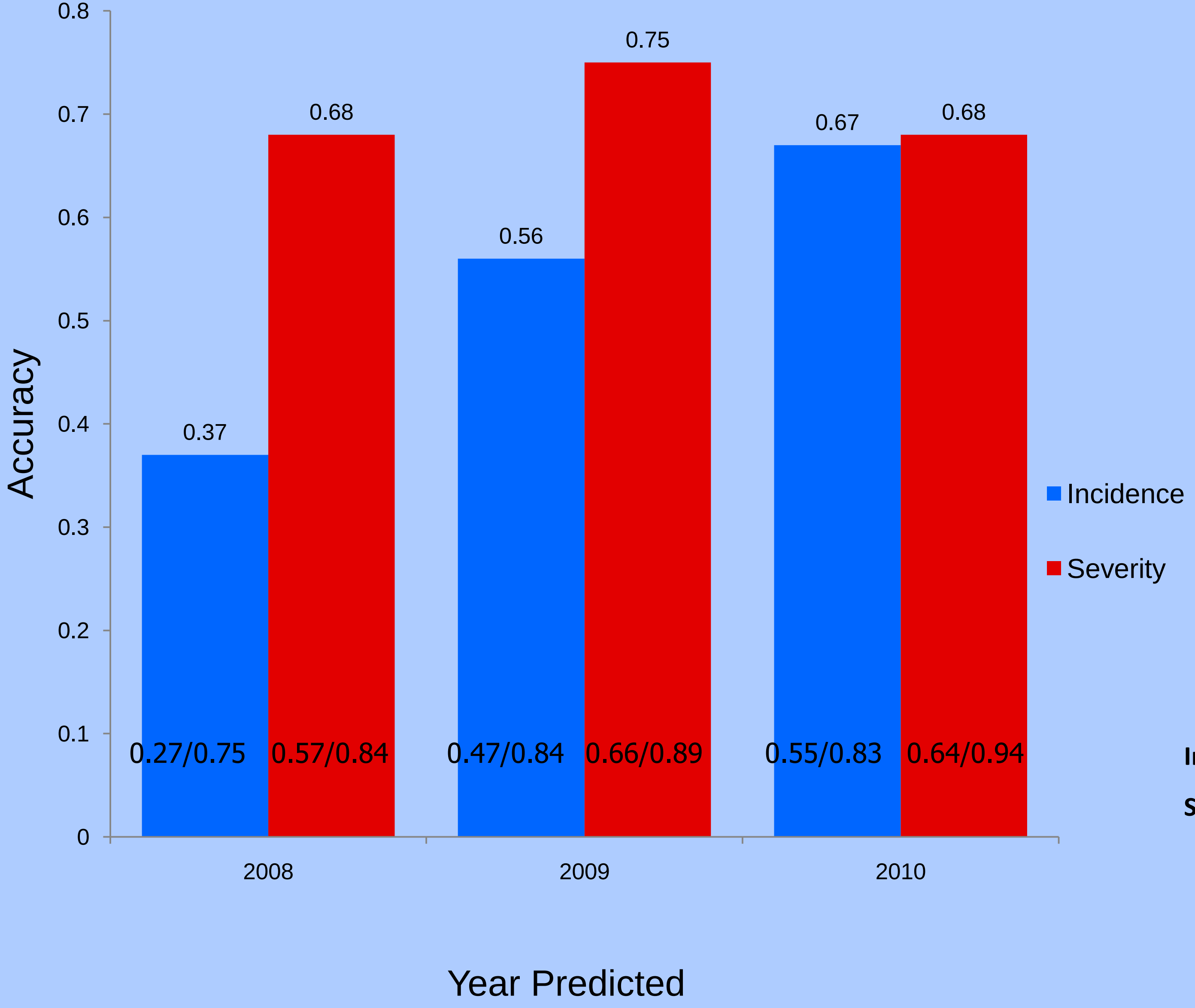
Results





Ridge regression performs better for both incidence and severity

Corrected Ridge Regression Prediction Accuracies



H^2

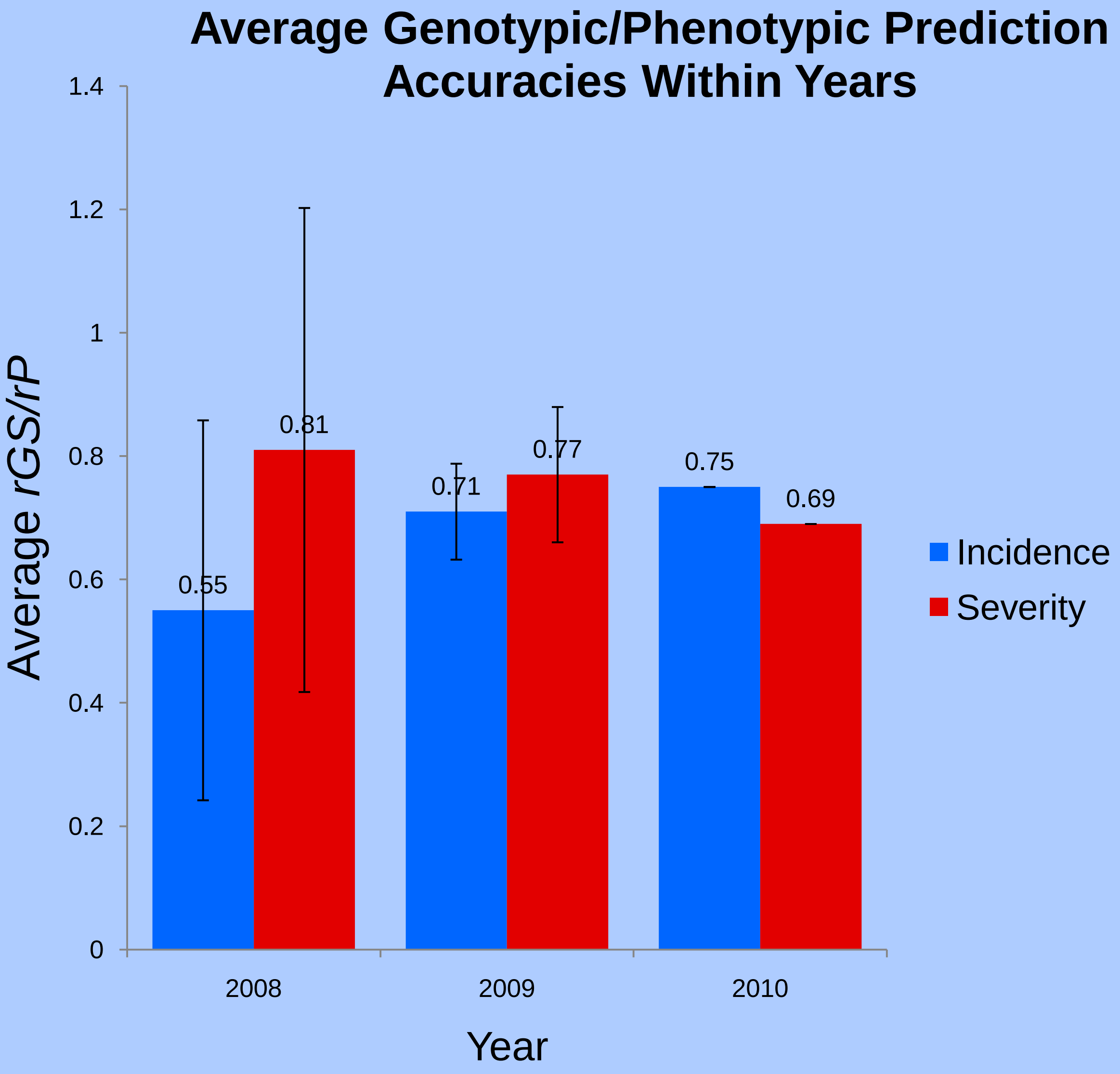
	2008	**2009**	**2010**
Incidence	0.56	0.71	0.7
Severity	0.71	0.79	0.88

H

	2008	**2009**	**2010**
Incidence	0.75	0.84	0.83
Severity	0.84	0.89	0.94

_RR_ accuracy _r̂_

	2008	**2009**	**2010**
Incidence	0.27	0.47	0.55
Severity	0.57	0.66	0.64

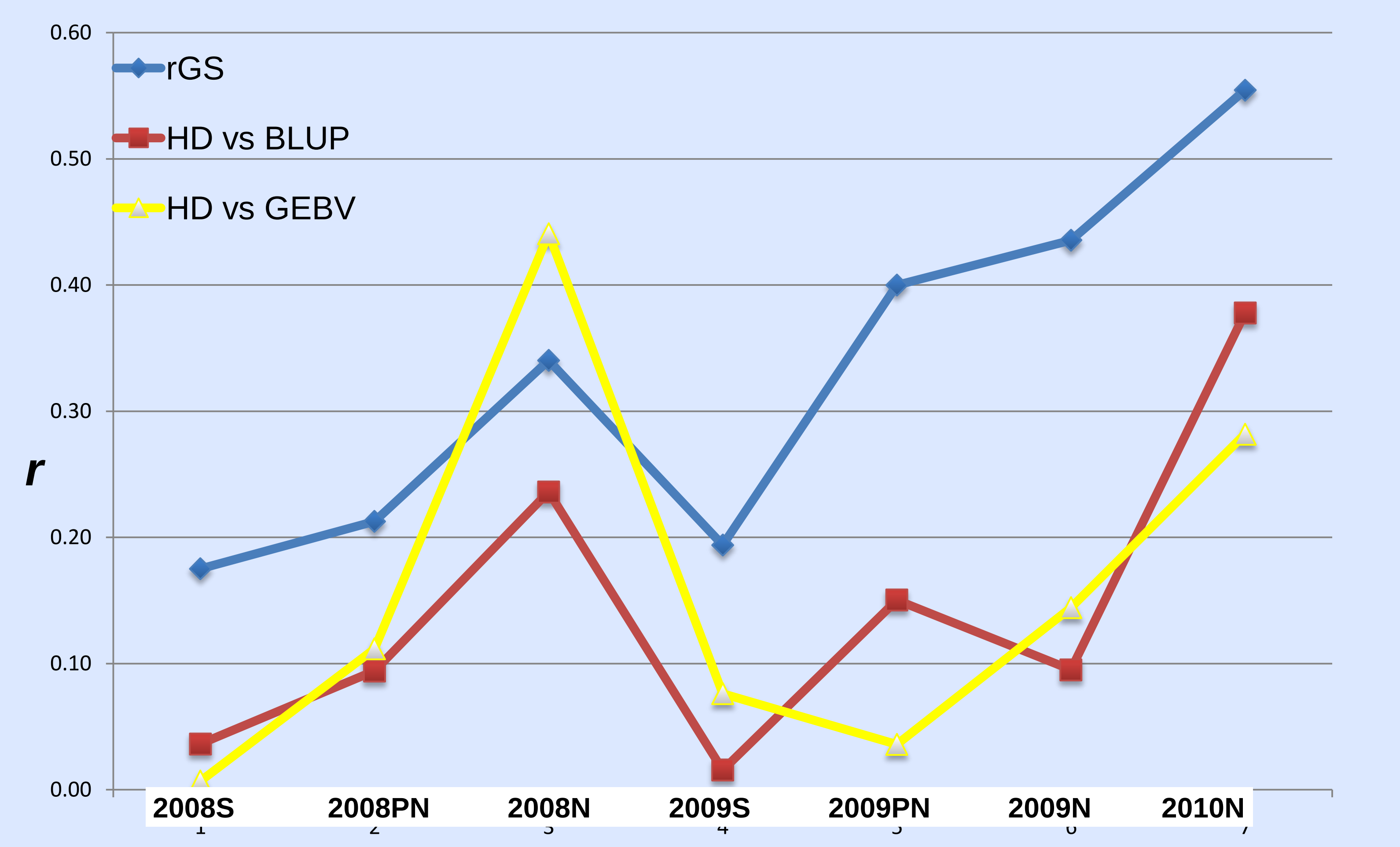


Average Genotypic/Phenotypic Prediction Accuracies Within Years

Year	Nursery	rP	Mean rGS	rGS/rP
Incidence				
2010	Northern	0.74	0.55	0.75
2009	Northern	0.57	0.44	0.76
2009	Southern	0.66	0.43	0.65
2009	P. Northern	0.56	0.4	0.71
2008	Northern	0.68	0.34	0.5
2008	Southern	0.65	0.18	0.27
2008	P. Northern	0.24	0.21	0.88
Severity				
2010	Northern	0.85	0.59	0.69
2009	Northern	0.81	0.57	0.71
2009	Southern	0.81	0.57	0.71
2009	P. Northern	0.64	0.57	0.9
2008	Northern	0.48	0.54	1.13
2008	Southern	0.64	0.24	0.37
2008	P. Northern	0.59	0.55	0.92

Estimates of genotypic/phenotypic prediction accuracies are conservative because *rP* is calculated within each year and nursery

GS Accuracy Compared to Correlations between Days to Heading BLUP & GEBV



Wide range in HD in Southern nurseries reduces prediction accuracy

	2008S	2008PN	2008N	2009S	2009PN	2009N	2010N
HD vs BLUP	0.04	-0.09	-0.24	0.02	0.15	0.10	-0.38
HD vs GEBV	-0.01	-0.11	-0.44	-0.08	-0.04	0.14	-0.28

Eastern Region Coordinated Projects

Two multi-PI projects (Northern CP) set the stage for using MAS and GS for accelerating improvement of FHB resistance in soft winter wheat:

Multi-PI #4 Improved Breeding Methods for FHB (Clay Sneller)

1. Elucidate the genetic structure of FHB resistance
2. Develop models to implement genomic selection for multiple FHB traits.
3. Document resistance to kernel infection and toxin accumulation

Multi-PI #5 Male-Sterile Facilitated Recurrent Selection for FHB (Ed Souza)

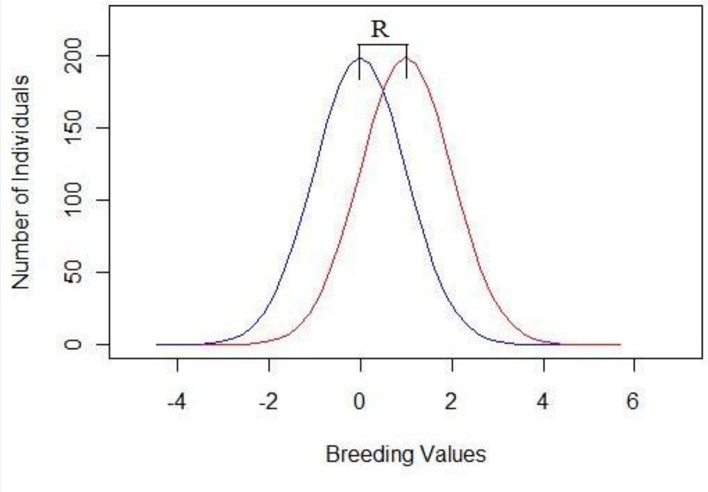
1. Early maturity MS selections: Southern-mid-Atlantic population.
2. Mid-maturity MS selections: Early Midwest population.
3. Mid-maturity and late MS selections: Late Midwest population.
4. Late maturity MS selections: Late population.

Both depend on being able to predict FHB resistance across diverse germplasm rather than within biparental populations

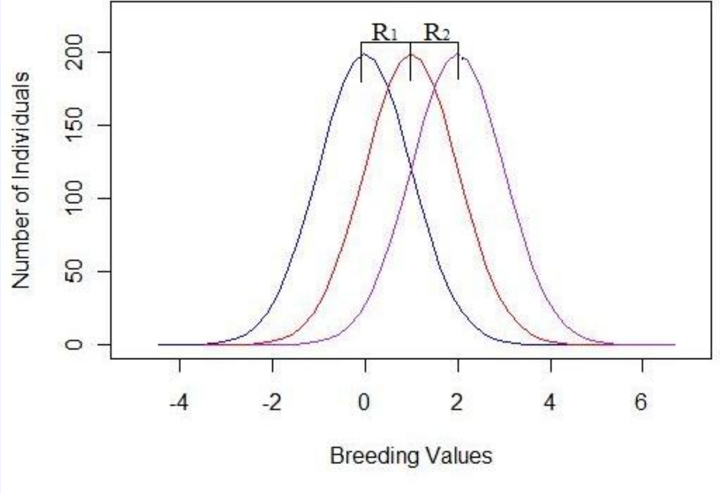
Conclusions: Genomic Selection for FHB Resistance

- Small training populations limited the accuracy of GS models
- Ridge Regression performed better than Bayes Cpi for this dataset
- Prediction accuracies for the 2008 Southern FHB Nursery were low because there were fewer locations and the wide range in heading date affected FHB scores
- Ridge regression models provided sufficient accuracy to increase genetic gain with two cycles of selection per year without phenotyping

PS: OneCycle per year



GS: Two Cycles per year



Summary: Genomic Selection

- GS differs from MAS and Association Breeding in that the underlying genetic control and biological function is not known.
- Breeders can implement GS without the upfront cost of obtaining that knowledge.
- GS preserves the creative nature of phenotypic selection to sometimes arrive at solutions outside the engineer’s scope.
- <u>Most important advantages are reductions in the length of the selection cycle and associated phenotyping cost resulting in greater genetic gain per year.</u>

Acknowledgements

- Supported in part by a grant from the Bill & Melinda Gates Foundation to Cornell University for the Borlaug Global Rust Initiative Durable Rust Resistance in Wheat Project.
- USDA Cooperative State Research, Education and Extension Service, Coordinated Agricultural Project 2005-05130-Wheat Applied Genomics
- USDA National Needs Fellowship Grant 2005-38420-15785: Provided Fellowship for Elliot Heffner
- USDA National Needs Fellowship Grant 2008-38420-04755: Provided Fellowship for Jessica Rutkoski
- USDA National Needs Fellowship Grant 2008-38420-04755: Provided Fellowship for Anna Bishop-Tran
- Pioneer HiBred International: Provided partial Fellowships for Elliot Heffner and Anna Bishop-Tran

USDA United States Department of Agriculture

CSREES EDUCATION USDA RESEARCH · EXTENSION

Questions?