



Implementing Association Mapping and Genomic Selection in Soybean Breeding Program

Yong Bao

Advisors: Dr. Nevin Young & Dr. Jim Orf

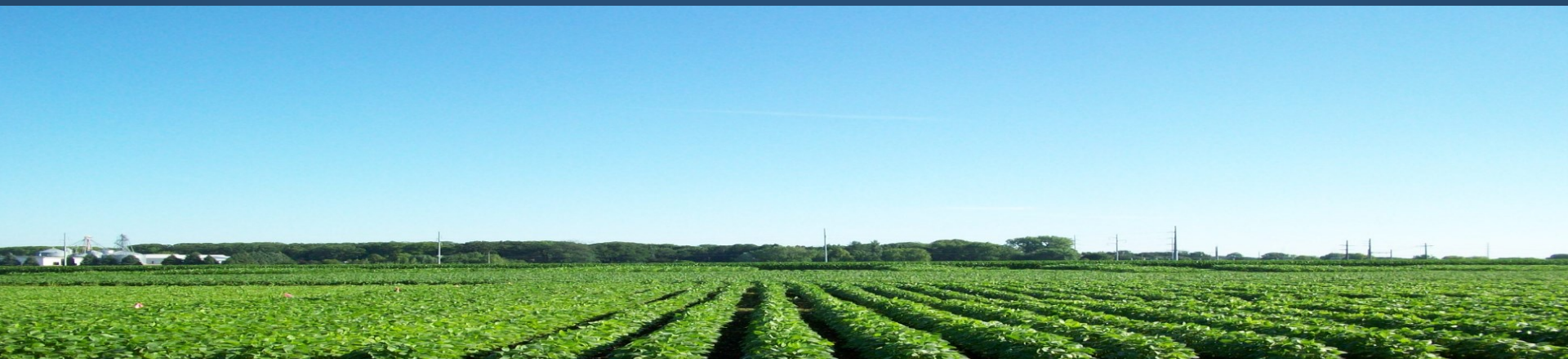
Funding: Minnesota Soybean Research & Promotion Council



UNIVERSITY OF MINNESOTA
Driven to DiscoverSM

Outline

- **Overview of all on-going projects**
- Introduction of association mapping (AM) and genomic selection (GS)
- Association mapping for PI88788 derived soybean cyst nematode (SCN) resistance
- Prediction accuracy of GS for quantitative resistance to SCN
- Conclusion and perspectives



Overview



Yield

Oil



Protein



SCN



SDS



BSR

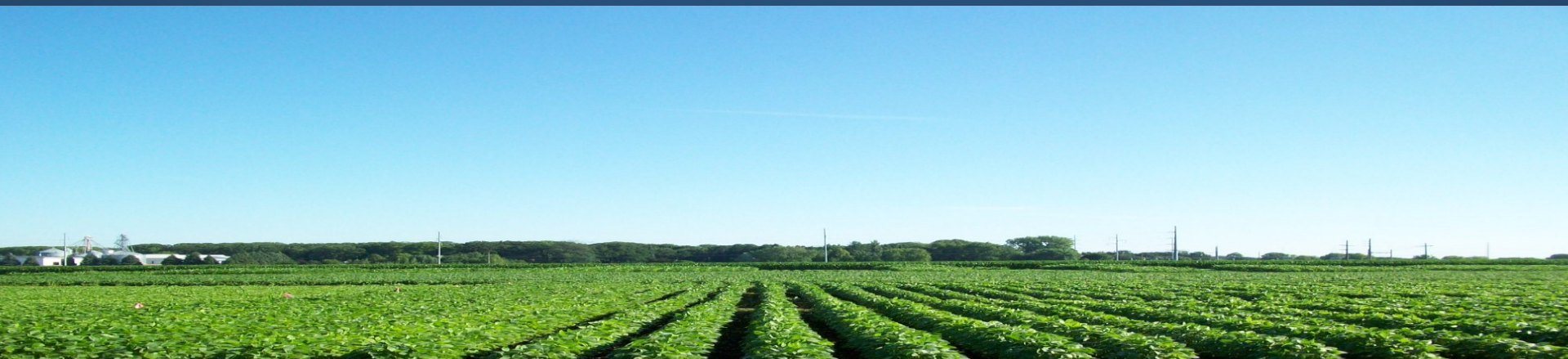


Dr. Jim Kurle, UMN

Bao, Y., **T. Vuong**, **C. Meinhardt**, R. Denny, P. Tiffin, S. Chen, **H. Nguyen**, J. H. Orf, and N. D. Young (2014): "Potential of *Association Mapping and Genomic Selection to Explore PI88788 Derived Soybean Cyst Nematode Resistance.*" **The Plant Genome** DOI:10.3835/plantgenome2013.11.0039.

Outline

- Overview
- **Introduction of association mapping (AM) and genomic selection (GS)**
- Association mapping for PI88788 derived soybean cyst nematode resistance
- Prediction accuracy of GS for quantitative resistance to SCN
- Conclusion and perspectives

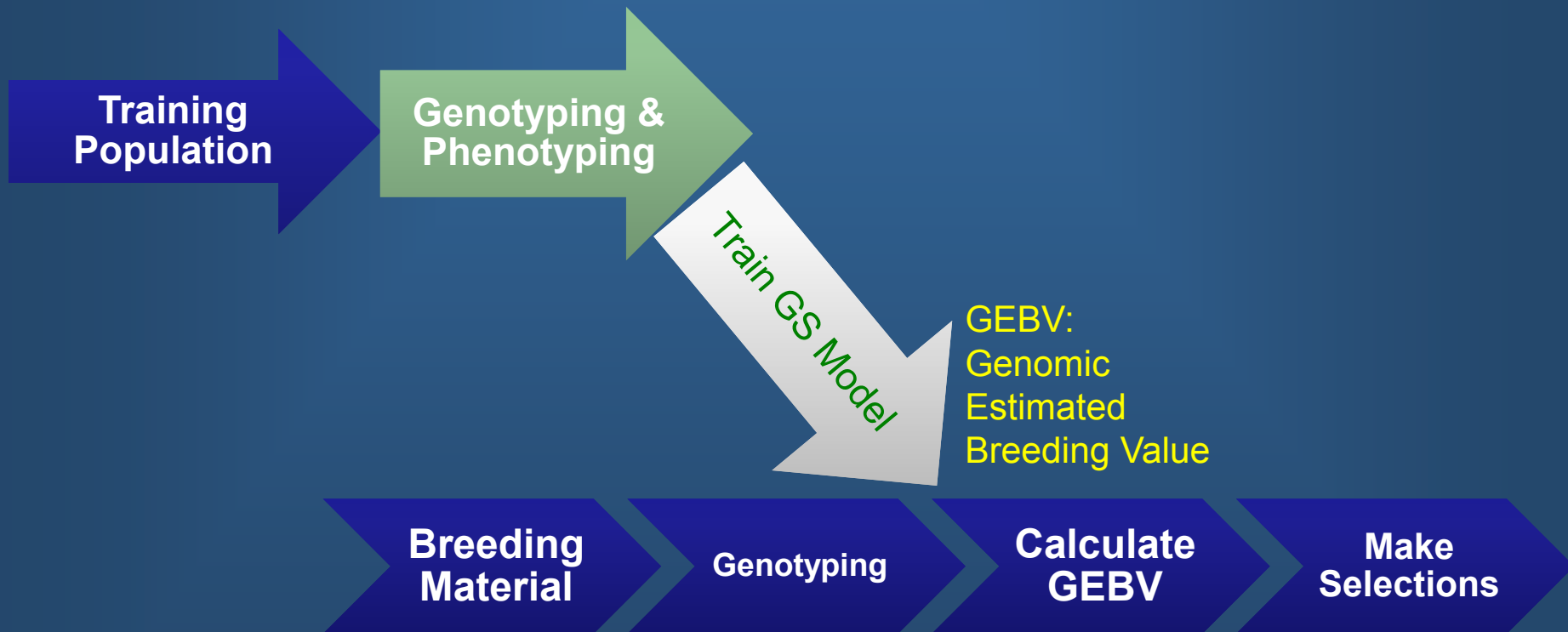


Genome-wide Association Mapping

- Allele mining for quantitative traits in **diverse germplasm** collection including advanced breeding lines, cultivars, landraces, etc.
- Potentially **high mapping resolution** with high-density markers
- Constraint: the existence of subpopulations may cause **false positives**

Genomic Selection:

An improved marker-base selection without QTL mapping

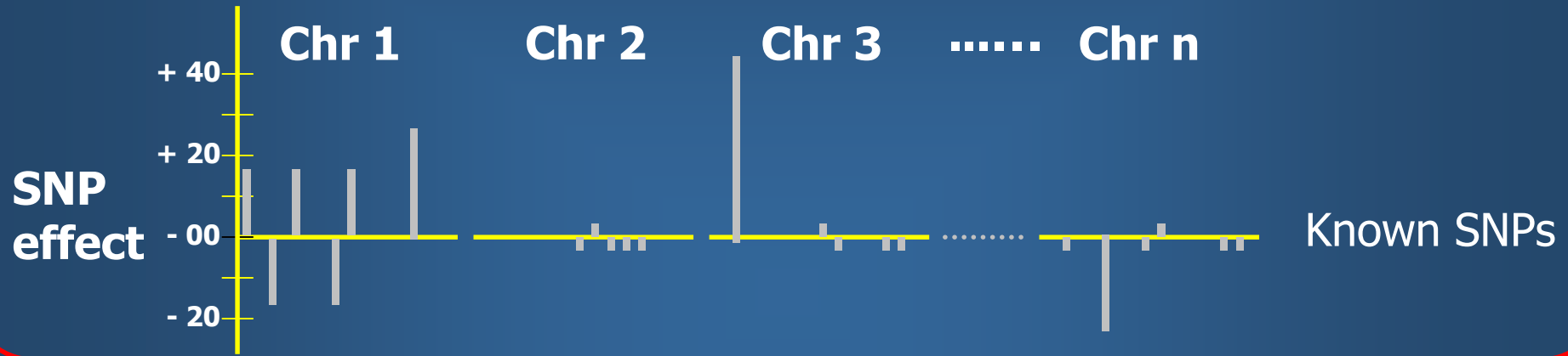


Training Population

$$y = \mu 1_n + X_i g_i + \varepsilon$$

Marker matrix

SNP marker effect



Breeding Material

$$GEBV = \sum_{i=1}^p X_i g_i$$

Sum of all marker effects

Marker matrix

SNP marker effect

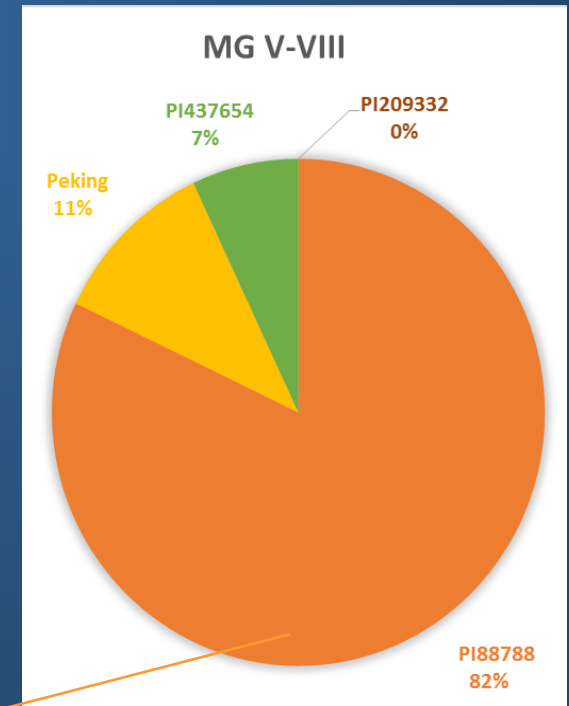
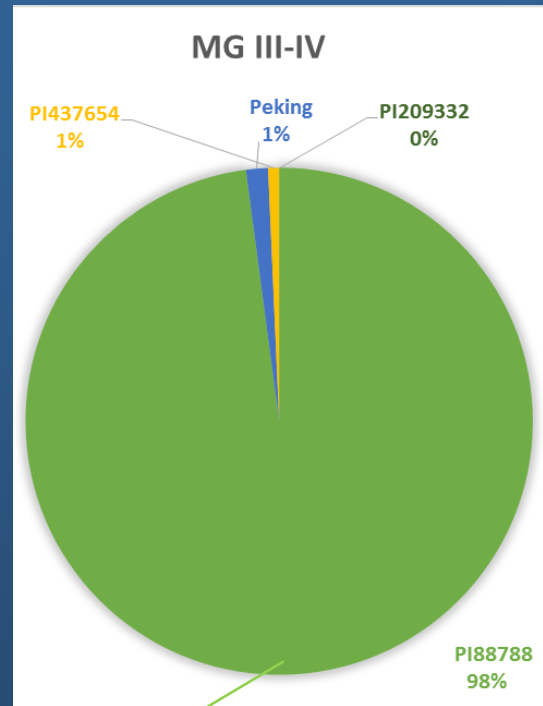
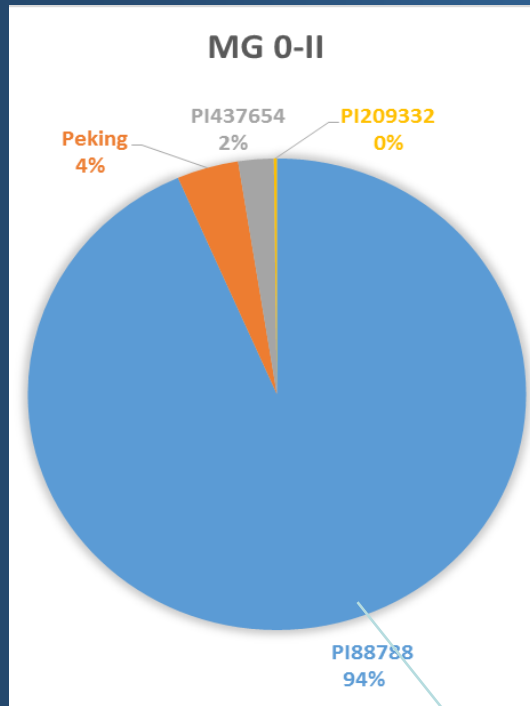
$$1 + 1 - 1 - 1 + 1 + 25 - 1 + 1 - 1 - 1 - 1 + 42 + 1 - 1 - 1 - 1 - 1 - 22 - 1 + 1 - 1 - 1 = +38$$

Outline

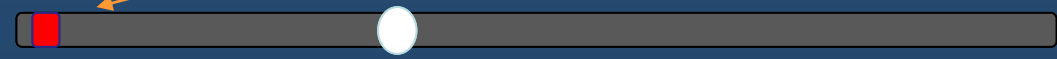
- Overview
- Introduction of association mapping (AM) and genomic selection (GS)
- **Association mapping for PI88788 derived soybean cyst nematode resistance**
- Prediction accuracy of GS for quantitative resistance to SCN
- Conclusion and perspectives



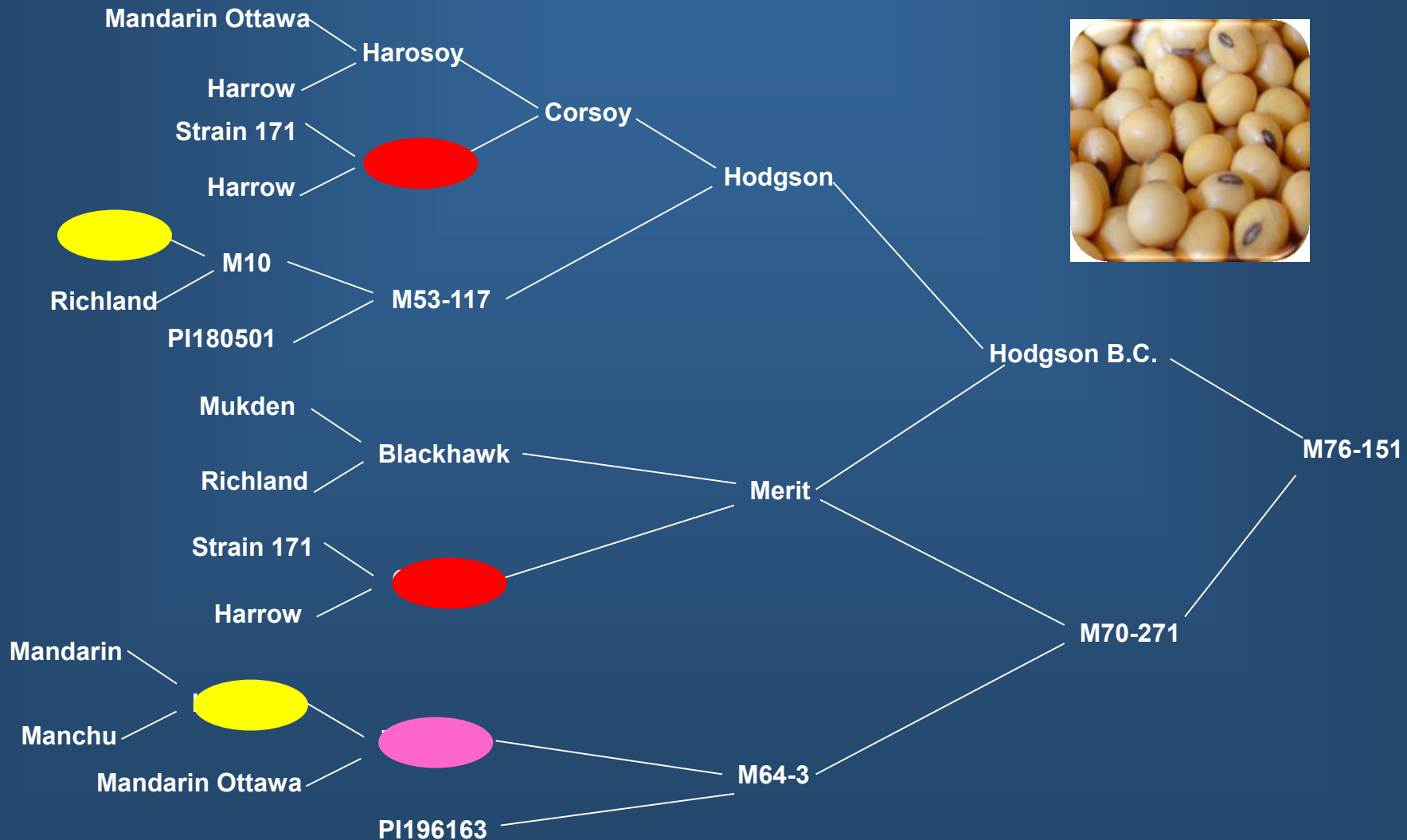
Sources of resistance of 1,500 commercial SCN-resistant soybean cultivars



rhg1



Select 282 Germplasm Based on Footprints



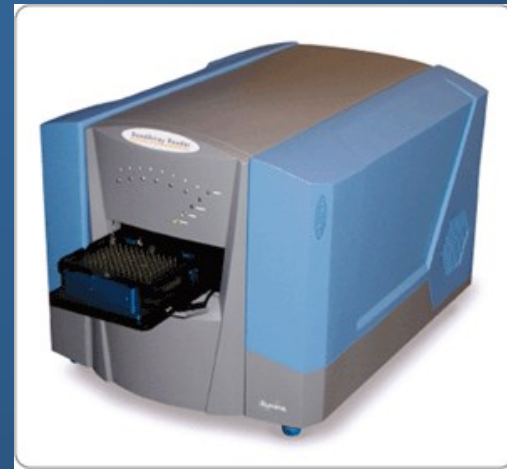
Footprint value: Lincoln > Capital > Traverse

Phenotyping and Genotyping



SCN race 3 (HG type 0)
Two replications of five plants for each genotype
(*Guo et al., 2005, U of Missouri*)

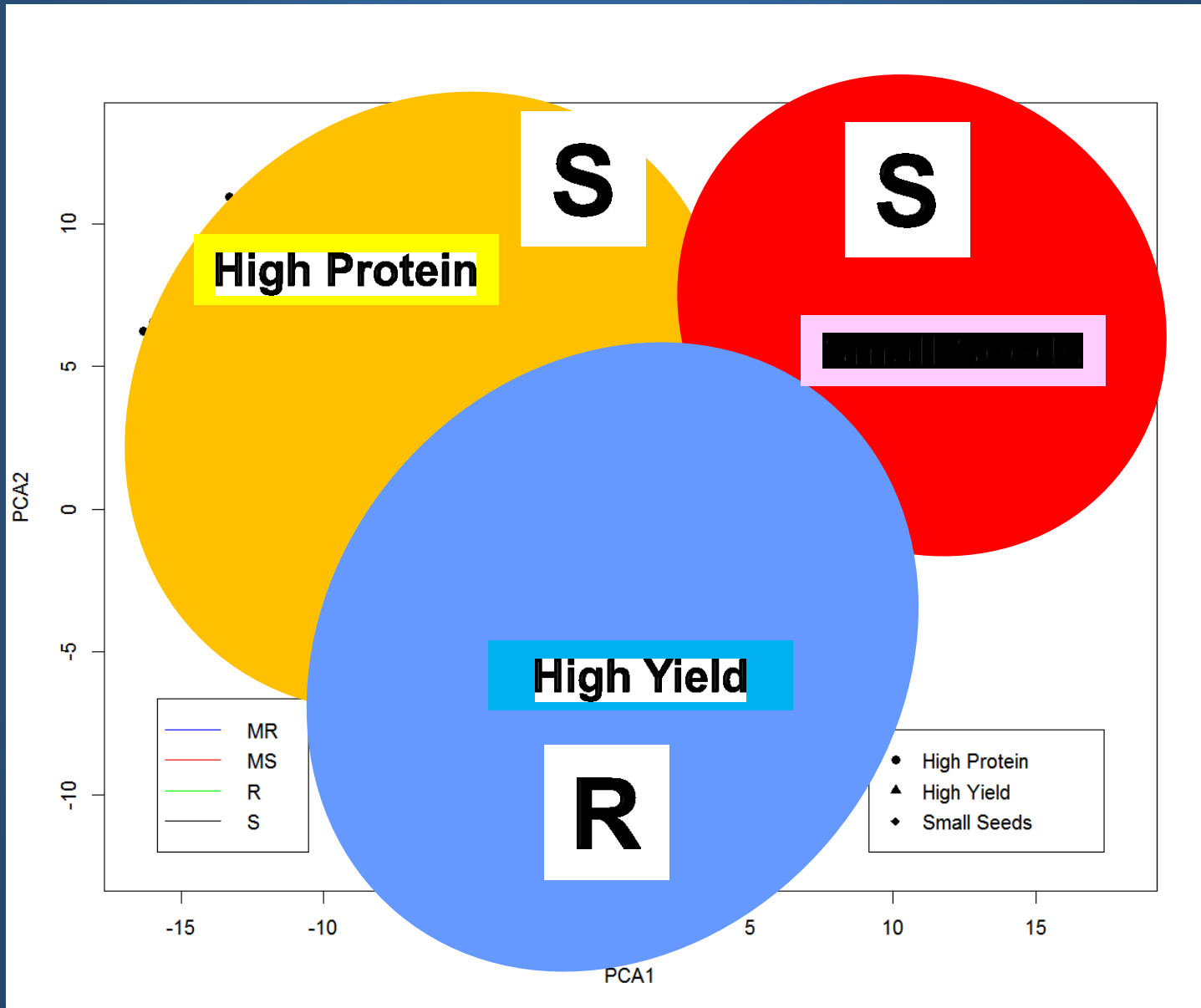
$$FI = \frac{\text{Number of female nematodes on a given individual}}{\text{Mean number of female nematodes on Hutcheson}} \times 100$$



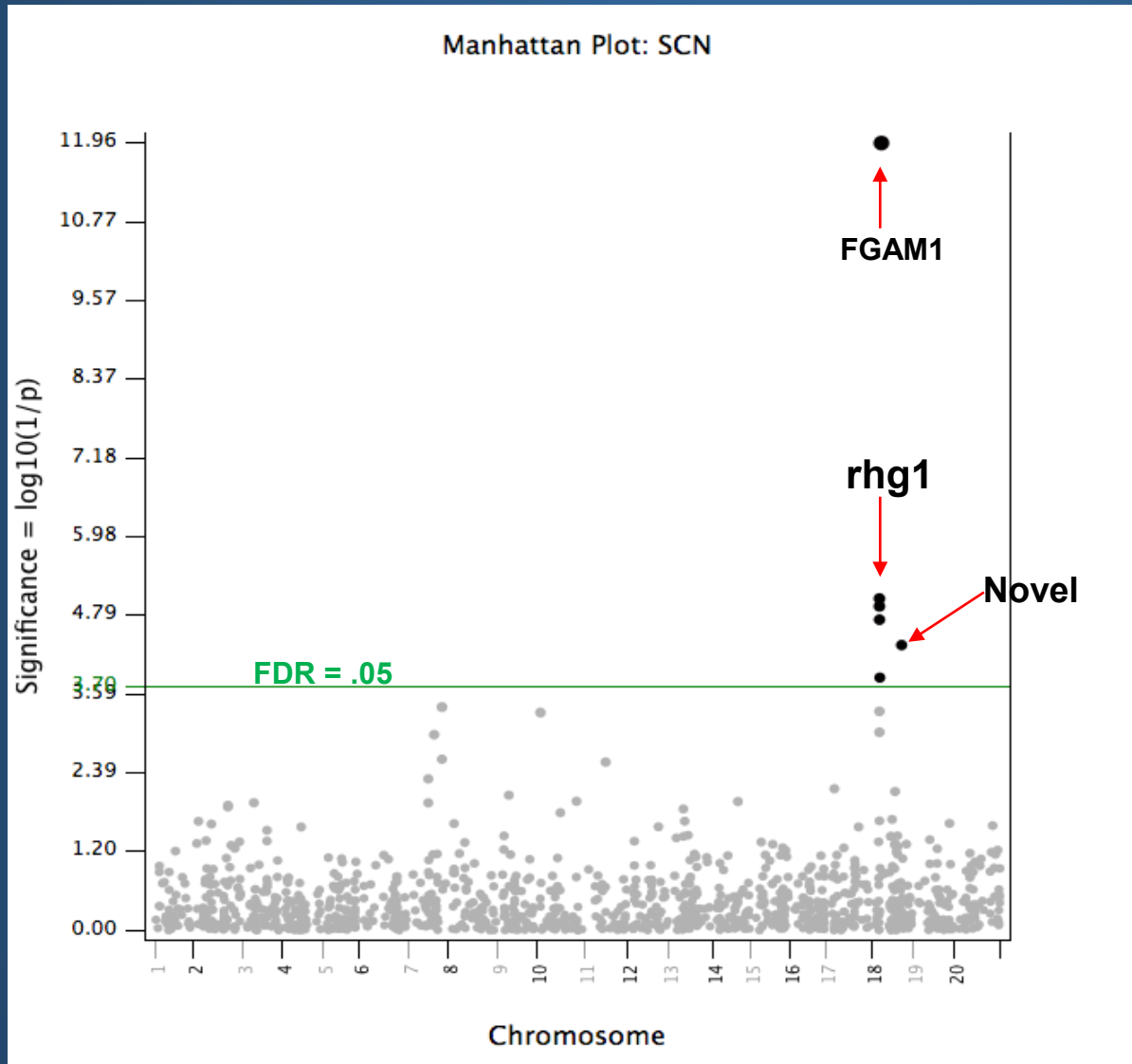
1,536 genome-wide SNPs Goldengate Assay (*Hyten et al., 2010*)

Minor allele frequency > 5%
Missing rate < 50%

Three Groups Exist in Panel



Three Loci Identified in AM

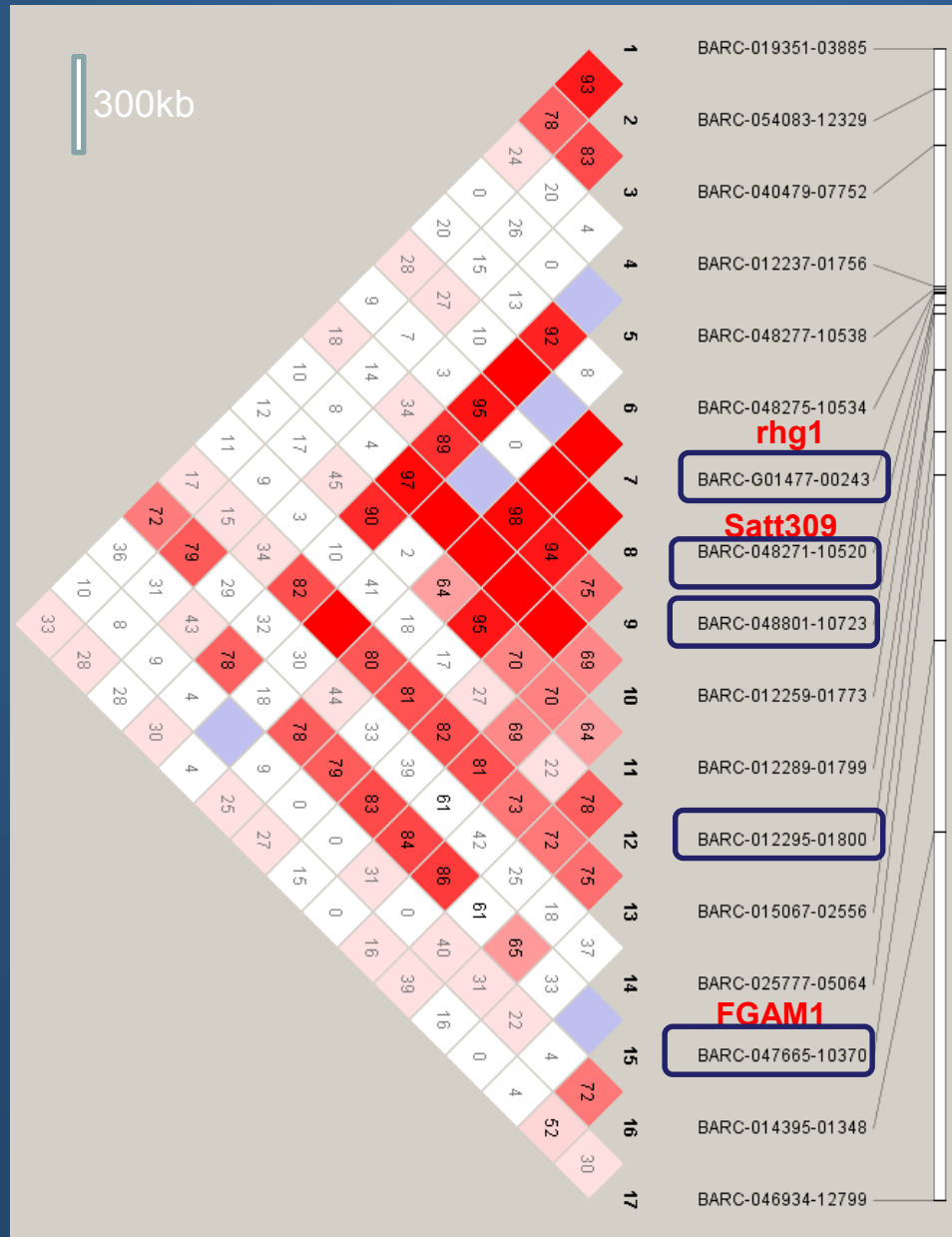


rhg1

FGAM1

Novel

LD Plot for rgh1 and FGAM1 Regions



~ 1.1 Mb

LD: $r^2 = \sim 0.8$

Outline

- Overview
- Introduction of association mapping (AM) and genomic selection (GS)
- Association mapping for PI88788 derived soybean cyst nematode resistance
- **Prediction accuracy of GS for quantitative resistance to SCN**
- Conclusion and perspectives

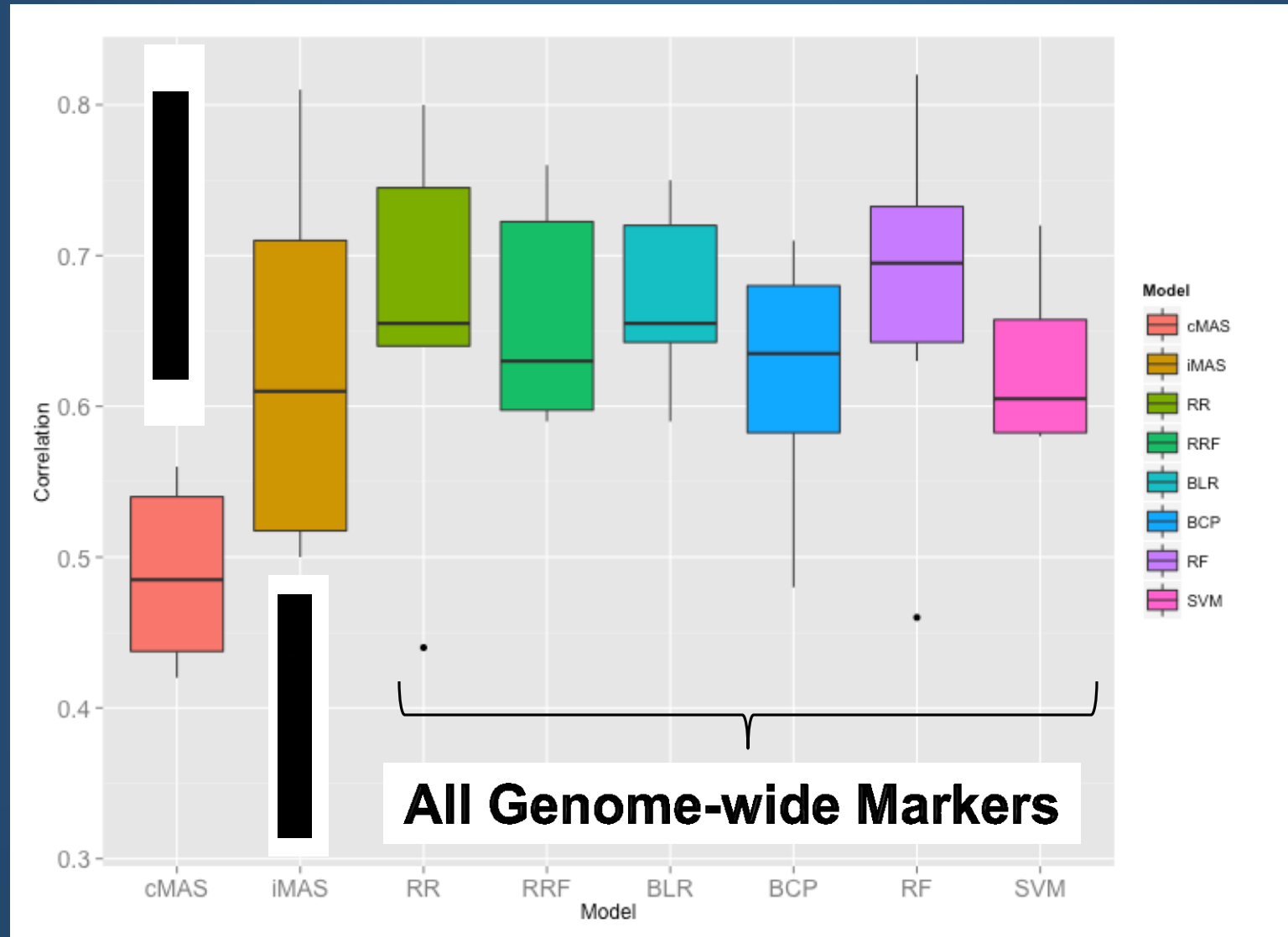


Six-fold Cross-validation

- Each subset: $282/6 = 47$ lines
- Training set: $47*5 = 235$ lines
- Test set: 47 lines
- Prediction accuracy = correlation of GEBV and phenotypic value in test set

	Subset1	Subset2	Subset3	Subset4	Subset5	Subset6
Fold 1	Test	Training	Training	Training	Training	Training
Fold 2	Training	Test	Training	Training	Training	Training
Fold 3	Training	Training	Test	Training	Training	Training
Fold 4	Training	Training	Training	Test	Training	Training
Fold 5	Training	Training	Training	Training	Test	Training
Fold 6	Training	Training	Training	Training	Training	Test

GS is significantly more accurate than MAS & All GS prediction algorithms are equivalent

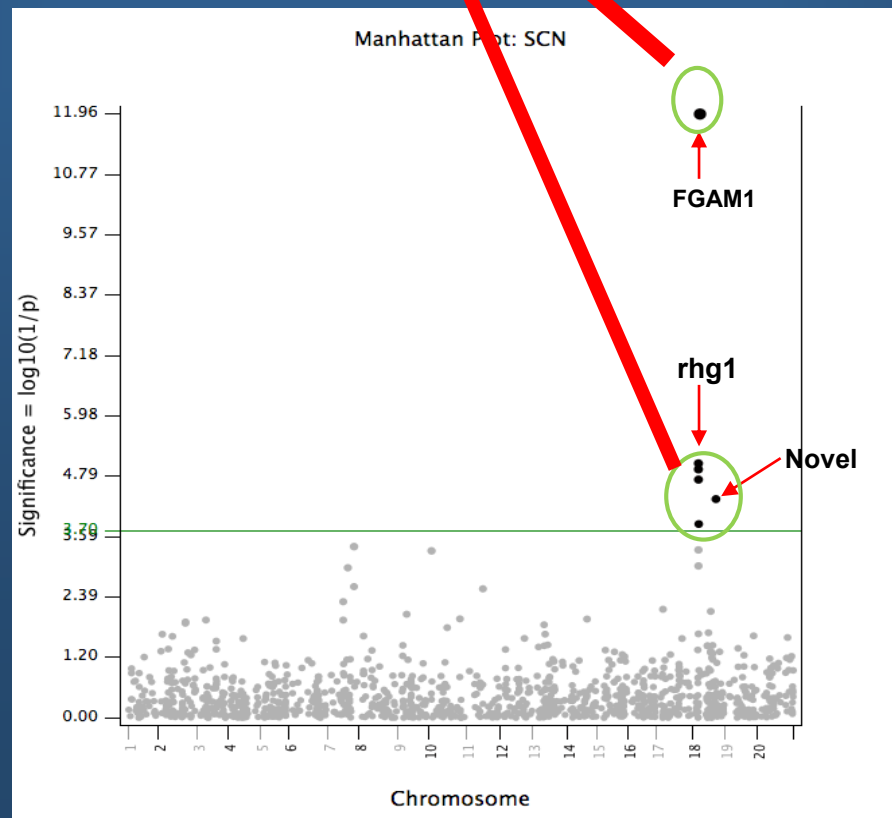


Account for Major QTL in GS

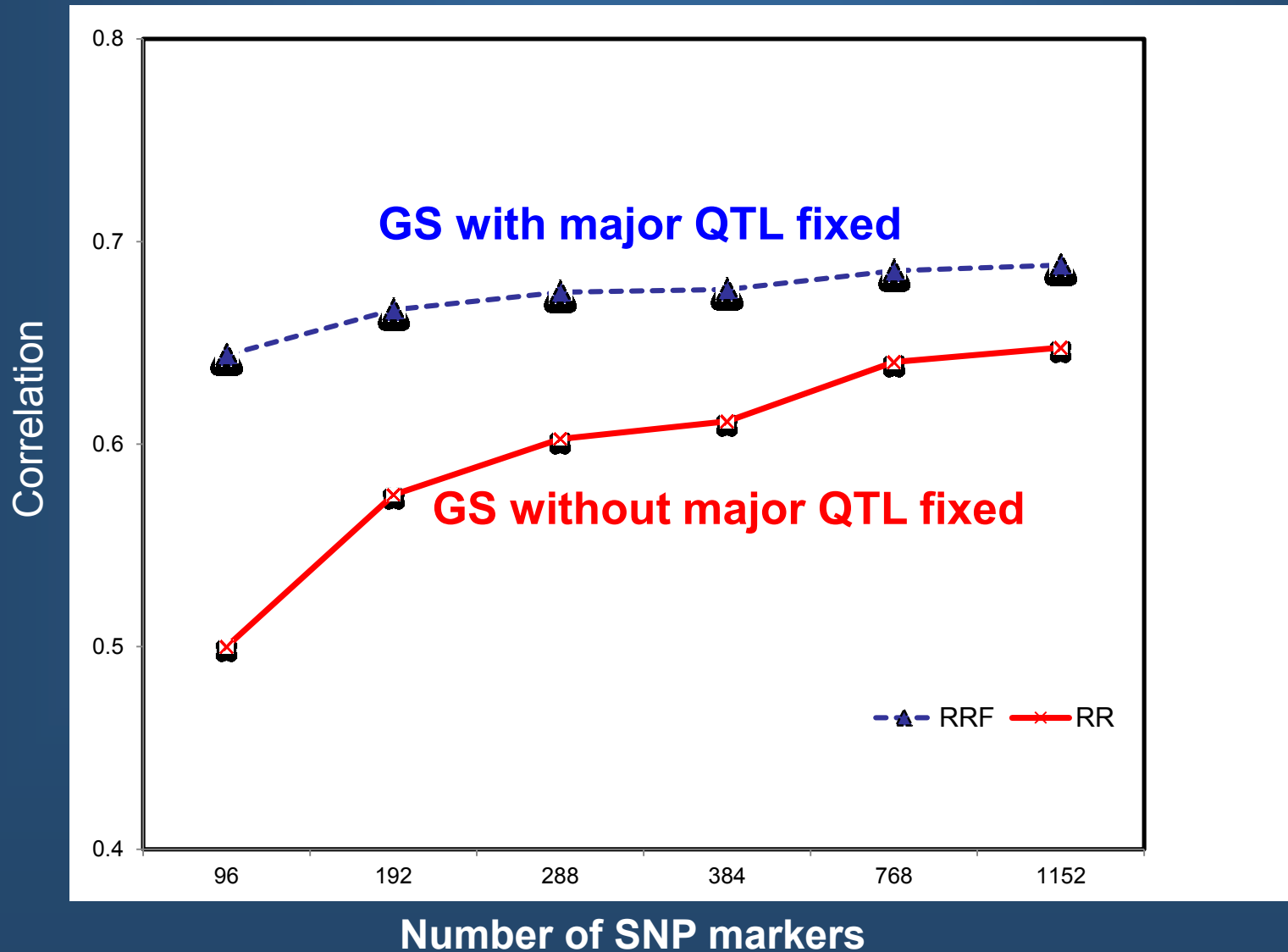
RRF-BLUP Model

$$y = \mu 1_n + \boxed{Wm} + X_i g_i + e$$

SNPs to be fixed



Compare GS w/o Major QTL Fixed



Conclusion and Perspective

- AM detected significant signals at *rhg1* and *FGAM1*, plus the third locus located on chromosome 18.
- GS was more accurate than marker-assisted selection (MAS) strategies using two significant markers alone.
- AM was extended to SDS, and GS was extended to yield, protein, and oil for MN germplasm.



Acknowledgements

Committee

- **Dr. Nevin Young (co-advisor)**
- **Dr. James Orf (co-advisor)**
- **Dr. Rex Bernardo**
- **Dr. Senyu Chen**
- **Dr. James Kurle**
- **Dr. Peter Tiffin**

Collaborators

- **Dr. Henry Nguyen, Dr. Tri Voung, and Clinton Meinhardt** University of Missouri

Genomics and Bioinformatics Lab

- **Roxanne Denny**
- **Kevin Silverstein**
- **Surabhi Mithal**
- **Peng Zhou**
- **Joseph Guhlin**
- **Diana Trujillo**
- **Shaun Curtin**

Soybean Breeding Group

- **Gerald Decker**
- **Darcy Weston**
- **Phil Schaus**
- **Rafael Echenique**
- **Leo Moros**
- **MAST Students**

Soybean Pathology Lab

- **Grace Anderson**
- **Colin Zumwalde**
- **John Lencowski**
- **Erin Walch**
- **Adam Barbeau**
- **Dante Leyva**
- **Marissa Scherven**

Graduate Students

- **Ahmad Sallam**
- **Landon Ries**
- **JoAnn Kirsch**
- **Ben Campbell**
- **Lian Lian**
- **Amy Jacobson**

Funding

