



# **SoyNAM Project Update Preliminary Results**

**Brian Diers, Carol Bonin, Neil Yu, Randy Nelson,  
Troy Cary, John Meharry, UIUC**

**Bill Beavis, Dawn Gibson, Reka Howard Iowa State University  
Katy Martin Rainy, Bill Muir, Alencar Xavier, Purdue University**

**Jim Specht and George Graef, University of Nebraska**

**Perry Cregan, Qijian Song, USDA-ARS**

**William Schapaugh, Kansas State University**

**Stella Kantartzi, Southern Illinois University**

**Dechun Wang, Michigan State University**

**Grover Shannon, University of Missouri**

**Leah McHale, The Ohio State University**

**Rouf Mian, USDA-ARS**





# **SoyNAM Project Update Preliminary Results**

**Bill Beavis, Dawn Gibson, Reka Howard Iowa  
State University**

**Carol Bonin, Neil Yu, UIUC**

**Katy Martin Rainy, Bill Muir, Alencar Xavier,  
Purdue University**

# Outline

---

- Objectives
- Background
- Materials and methods
- Preliminary results





# Objectives

---

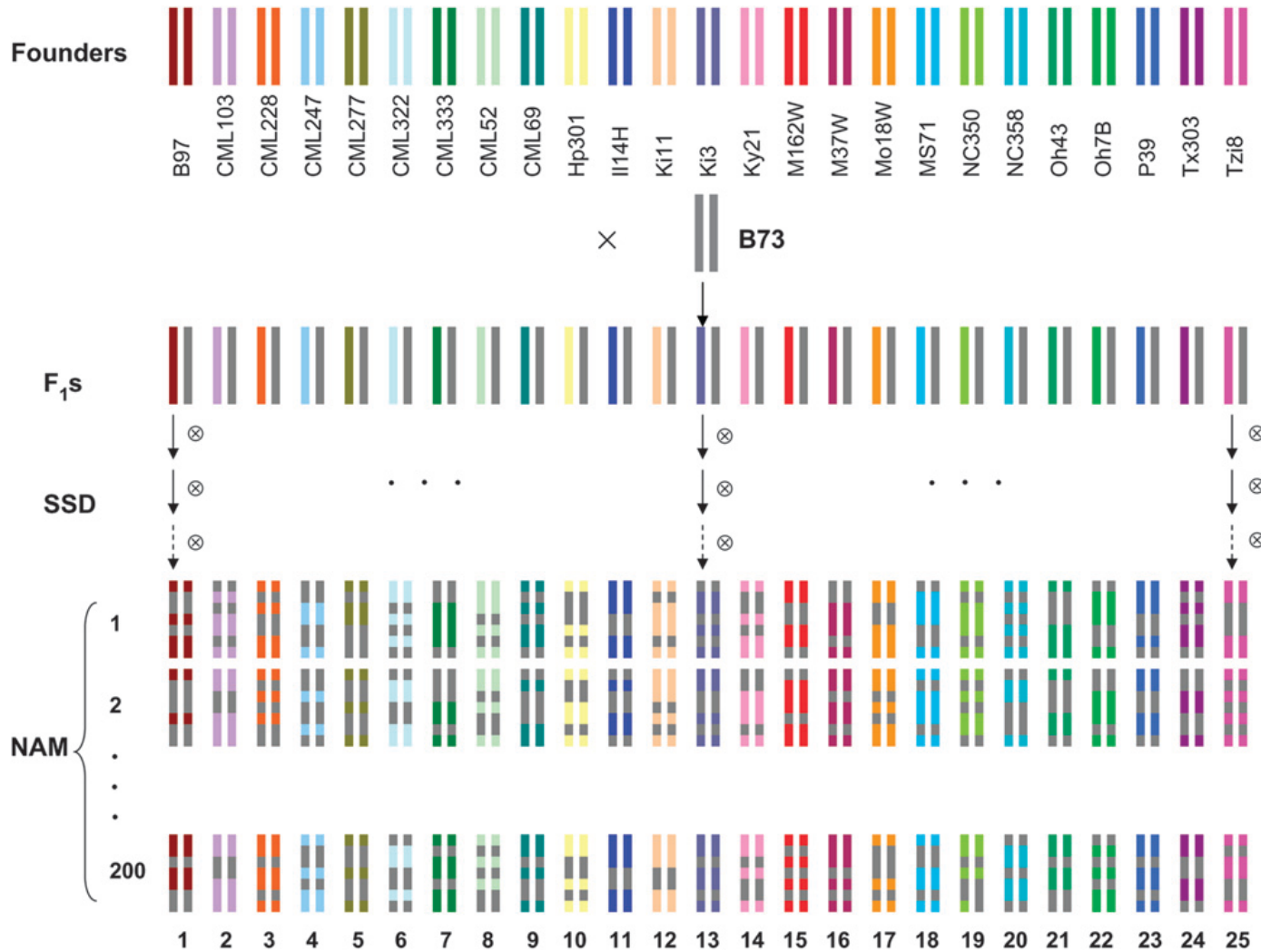
- Map QTL controlling agronomic, composition, physiological and resistance traits across a wide range of soybean germplasm.
- Identify beneficial QTL alleles from elite and exotic germplasm
- Use information to develop selection models for forward breeding.

# Nested Association Mapping (NAM)

---

- Nested association mapping (NAM) combines advantages of linkage and association mapping.
  - Linkage mapping - Advantage of power in identifying QTL but disadvantage of poor map resolution.
  - Association mapping – Advantage of high map resolution but disadvantage of poor power in identifying QTL.

# Structure of Maize NAM



From Yu et al. Genetics 178:539-551.



# Parents

- 40 Parents include:
  - 17 high yielding parents from 8 states.
  - 15 lines with diverse ancestry from R. Nelson's program.
  - 8 plant introductions.

Parent	Origin	Parent	Origin
<b>4J105-3-4</b>	Purdue Univ.	<b>LG94-1128</b>	USDA-ARS
<b>5M20-2-5-2</b>	Purdue Univ.	<b>LG94-1906</b>	USDA-ARS
<b>CL0J095-4-6</b>	Purdue Univ.	<b>LG97-7012</b>	USDA-ARS
<b>CL0J173-6-8</b>	Purdue Univ.	<b>LG98-1605</b>	USDA-ARS
<b>HS6-3976</b>	Ohio State	<b>Magellan</b>	Univ. of Missouri
<b>LD00-3309</b>	Univ. of Illinois	<b>Maverick</b>	Univ. of Missouri
<b>LD01-5907</b>	Univ. of Illinois	<b>NE3001</b>	Univ. of Nebraska
<b>LD02-4485</b>	Univ. of Illinois	<b>Prohio</b>	Ohio State Univ.
<b>LD02-9050</b>	Univ. of Illinois	<b>S06-13640</b>	Univ. of Missouri
<b>LG03-2979</b>	USDA-ARS	<b>Skylla</b>	Mich. State Univ.
<b>LG03-3191</b>	USDA-ARS	<b>TN05-3027</b>	Univ. of Tenn.
<b>LG00-3372</b>	USDA-ARS	<b>U03-100612</b>	Univ. of Nebraska
<b>LG04-4717</b>	USDA-ARS	<b>PI 398.881</b>	South Korea
<b>LG04-6000</b>	USDA-ARS	<b>PI 427.136</b>	South Korea
<b>LG05-4292</b>	USDA-ARS	<b>PI 437.169B</b>	Russia
<b>LG05-4317</b>	USDA-ARS	<b>PI 438.164B</b>	China
<b>LG05-4464</b>	USDA-ARS	<b>PI 518.751</b>	Serbia
<b>LG05-4832</b>	USDA-ARS	<b>PI 561.370</b>	China
<b>LG90-2550</b>	USDA-ARS	<b>PI 404.188A</b>	China
<b>LG92-1255</b>	USDA-ARS	<b>PI 574.486</b>	China



# Parents

- 40 Parents include:

- 17 high yielding parents from 8 states.
- 15 lines with diverse ancestry from R. Nelson's program.
- 8 plant introductions.

Parent	Origin	Parent	Origin
4J105-3-4	Purdue Univ.	<b>LG94-1128</b>	USDA-ARS
5M20-2-5-2	Purdue Univ.	<b>LG94-1906</b>	USDA-ARS
CL0J095-4-6	Purdue Univ.	<b>LG97-7012</b>	USDA-ARS
CL0J173-6-8	Purdue Univ.	<b>LG98-1605</b>	USDA-ARS
HS6-3976	Ohio State	<b>Magellan</b>	Univ. of Missouri
LD00-3309	Univ. of Illinois	<b>Maverick</b>	Univ. of Missouri
LD01-5907	Univ. of Illinois	<b>NE3001</b>	Univ. of Nebraska
LD02-4485	Univ. of Illinois	<b>Prohio</b>	Ohio State Univ.
LD02-9050	Univ. of Illinois	<b>S06-13640</b>	Univ. of Missouri
<b>LG03-2979</b>	USDA-ARS	<b>Skylla</b>	Mich. State Univ.
<b>LG03-3191</b>	USDA-ARS	<b>TN05-3027</b>	Univ. of Tenn.
<b>LG00-3372</b>	USDA-ARS	<b>U03-100612</b>	Univ. of Nebraska
<b>LG04-4717</b>	USDA-ARS	<b>PI 398.881</b>	South Korea
<b>LG04-6000</b>	USDA-ARS	<b>PI 427.136</b>	South Korea
<b>LG05-4292</b>	USDA-ARS	<b>PI 437.169B</b>	Russia
<b>LG05-4317</b>	USDA-ARS	<b>PI 438.164B</b>	China
<b>LG05-4464</b>	USDA-ARS	<b>PI 518.751</b>	Serbia
<b>LG05-4832</b>	USDA-ARS	<b>PI 561.370</b>	China
<b>LG90-2550</b>	USDA-ARS	<b>PI 404.188A</b>	China
<b>LG92-1255</b>	USDA-ARS	<b>PI 574.486</b>	China

# Parents

- 40 Parents include:
  - 17 high yielding parents from 8 states.
  - 15 lines with diverse ancestry from R. Nelson's program.
  - 8 plant introductions.

Parent	Origin	Parent	Origin
4J105-3-4	Purdue Univ.	LG94-1128	USDA-ARS
5M20-2-5-2	Purdue Univ.	LG94-1906	USDA-ARS
CL0J095-4-6	Purdue Univ.	LG97-7012	USDA-ARS
CL0J173-6-8	Purdue Univ.	LG98-1605	USDA-ARS
HS6-3976	Ohio State	Magellan	Univ. of Missouri
LD00-3309	Univ. of Illinois	Maverick	Univ. of Missouri
LD01-5907	Univ. of Illinois	NE3001	Univ. of Nebraska
LD02-4485	Univ. of Illinois	Prohio	Ohio State Univ.
LD02-9050	Univ. of Illinois	S06-13640	Univ. of Missouri
LG03-2979	USDA-ARS	Skylla	Mich. State Univ.
LG03-3191	USDA-ARS	TN05-3027	Univ. of Tenn.
LG00-3372	USDA-ARS	U03-100612	Univ. of Nebraska
LG04-4717	USDA-ARS	PI 398.881	South Korea
LG04-6000	USDA-ARS	PI 427.136	South Korea
LG05-4292	USDA-ARS	PI 437.169B	Russia
LG05-4317	USDA-ARS	PI 438.164B	China
LG05-4464	USDA-ARS	PI 518.751	Serbia
LG05-4832	USDA-ARS	PI 561.370	China
LG90-2550	USDA-ARS	PI 404.188A	China
LG92-1255	USDA-ARS	PI 574.486	China

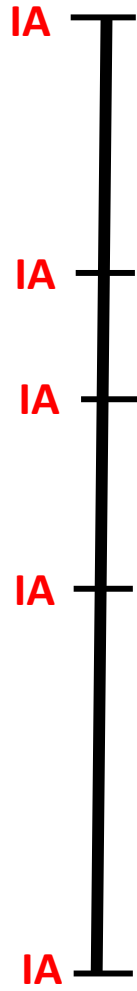
# Methods

---

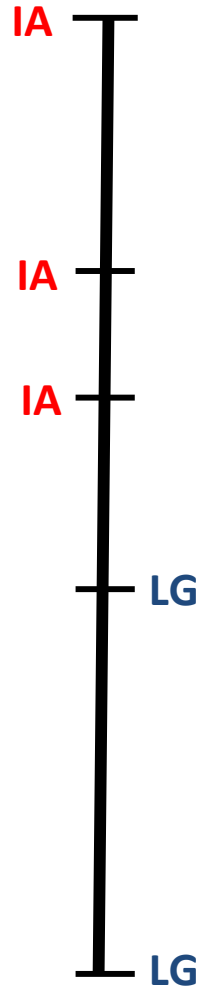
- 40 populations with 140 RILs in each population developed in Illinois and Nebraska (5,600 RILs).
- Genetic marker work done in Beltsville by Cregan and Song.
  - 41 parents of the populations tested with 50,000 SNP markers.
  - Parents resequenced to identify SNPs with rare alleles from IA3023.
  - Lines in populations tested with 5,300 SNP markers.
  - SNPs from 50 K array and resequencing are being projected to lines in the populations (Bill Muir, Purdue, University).



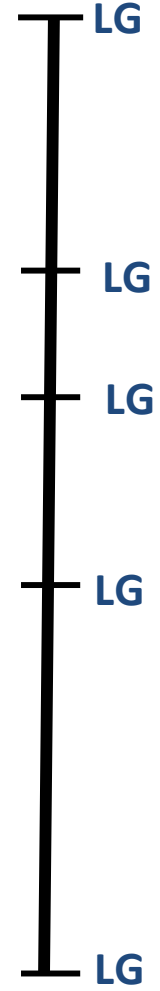
**IA**



**RIL**



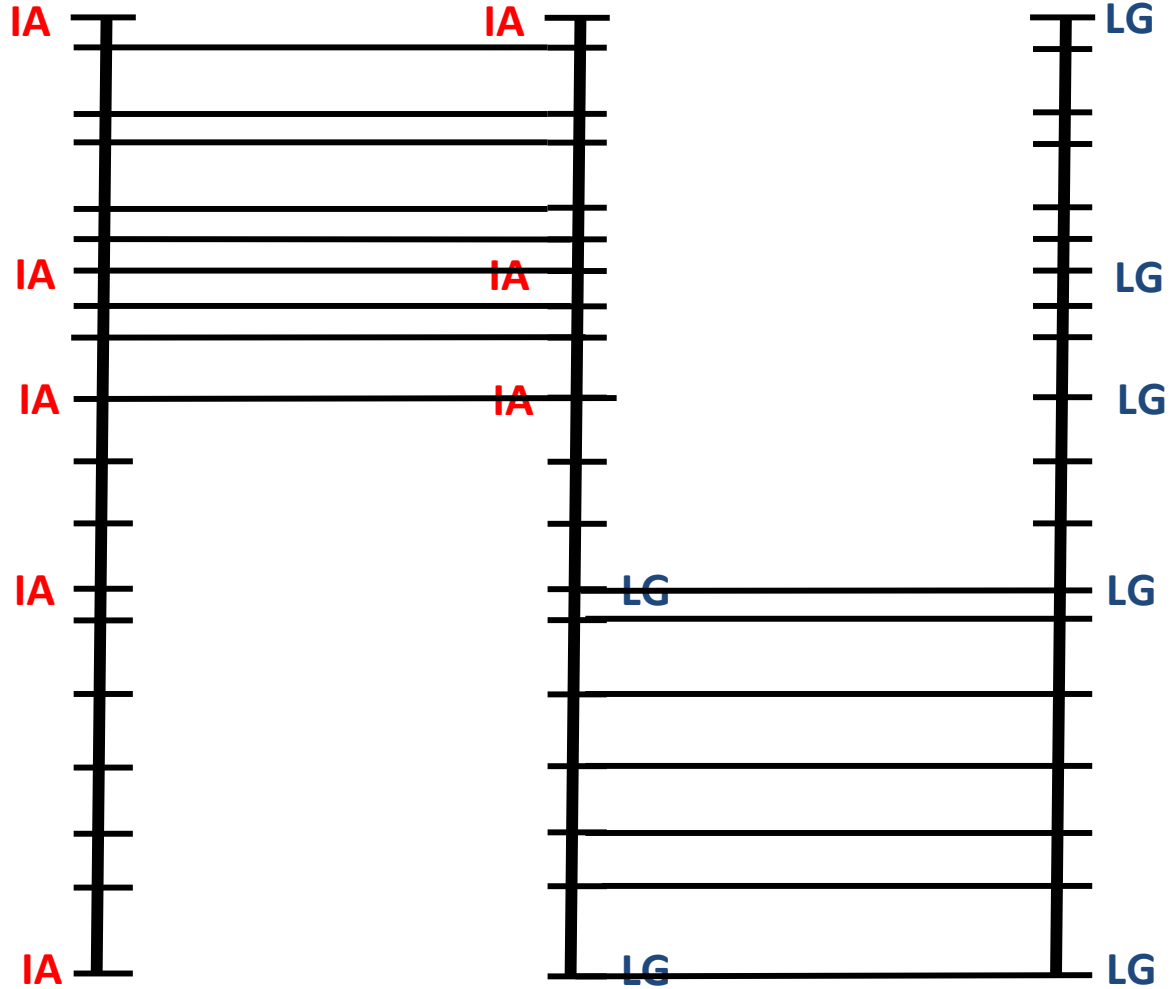
**LG**



**IA**

**RIL**

**LG**



# Phenotyping

- 80,000+ yield plots grown during 2011, 2012, and 2013.

---

Cooperator	State	Institution	Annual no. of plots
Brian Diers and Randall Nelson	Illinois	Univ. of Ill and USDA	6,400
Bill Beavis	Iowa	Iowa State Univ.	6,400
Katy Martin-Rainey	Indiana	Purdue Univ.	6,400
George Graef and James Specht	Nebraska	Univ. of Nebraska	6,400
Stella Kantartzi	Illinois	Southern Illinois Univ.	2,000
William Schapaugh	Kansas	Kansas State Univ.	4,000
Dechun Wang	Michigan	Michigan State Univ.	1,000
Grover Shannon	Missouri	Univ. of Missouri	1,000
Leah McHale	Ohio	Ohio State Univ.	2,000
Rouf Mian	Ohio	USDA-ARS	2,000
Total			37,600

---



# Phenotyping

---

- Challenging environments for phenotyping.
  - 2012 SIU lost from poor emergence.
  - 2013 Nebraska lost from hail.
  - Many locations suffered from significant drought.



# Phenotyping

---

- **Agronomic traits all environments:**
  - Lodging, Height, Maturity, Seed yield.
- **Agronomic traits at some environments:**
  - R1, Seed wt.
- **Composition traits:**
  - Protein, oil, elemental content (ionomics).
- **Physiological traits:**
  - Photosynthetic traits derived from leaf reflectance, canopy closure during flowering (light harvesting ability), canopy structure.
- **Serendipitous traits:**
  - Leaf wilting, SDS, vein streak necrosis virus.

# Preliminary Analysis

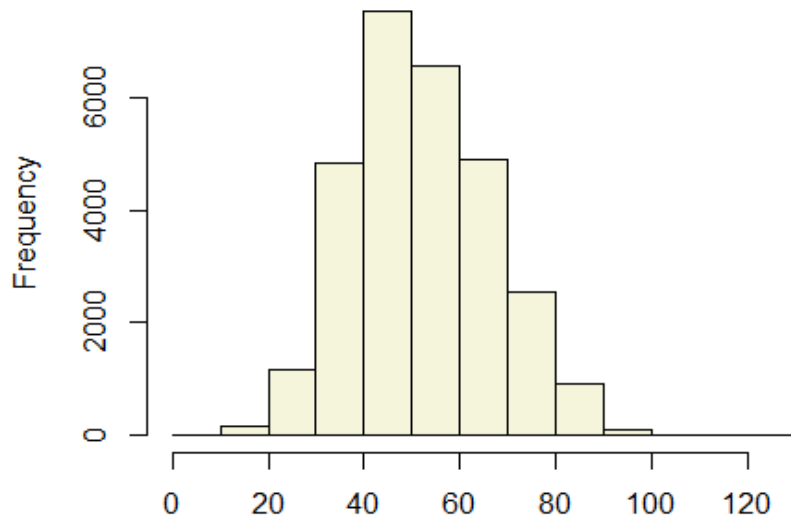
**Bill Beavis, Dawn Gibson, Reka Howard,  
Carol Fox, and Neil Yu**

---

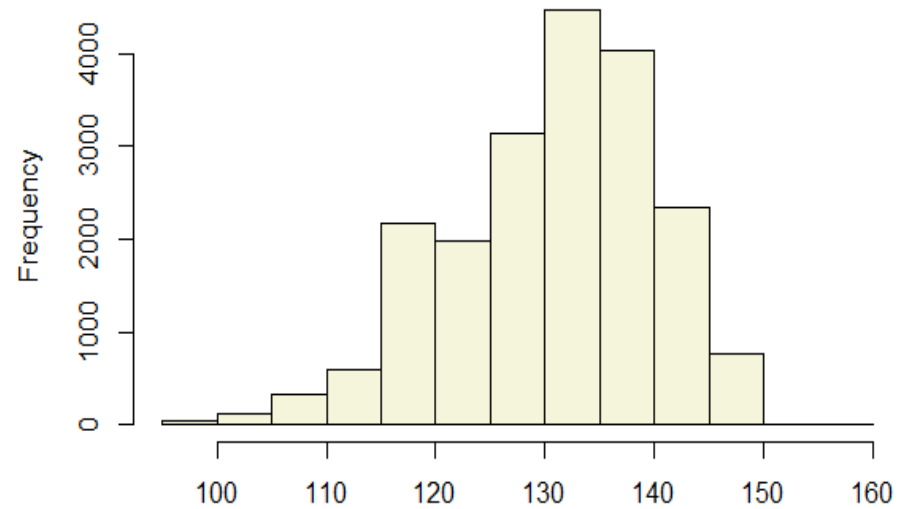
- 2011 and 2012 field data.
- Reviewed raw data.
- BLUPs (best linear unbiased predictions) calculated for each RIL.
  - Unbalanced data.
- Preliminary QTL scans completed.



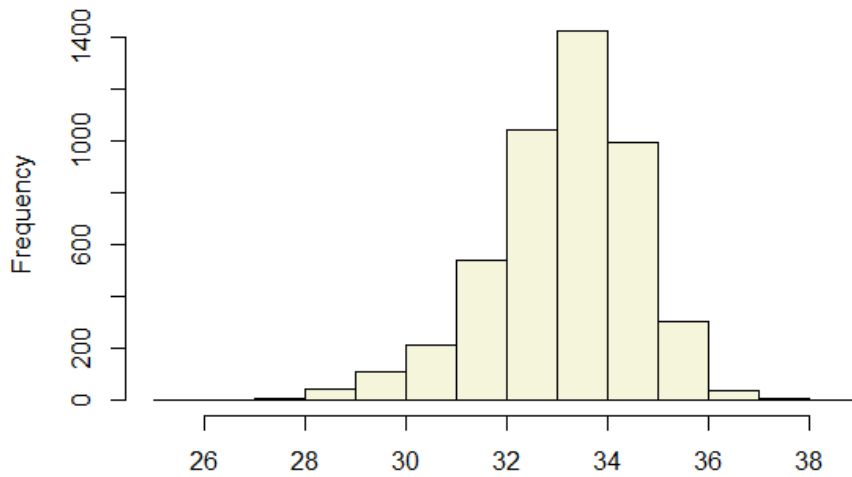
# Distributions of Individual Plot Data



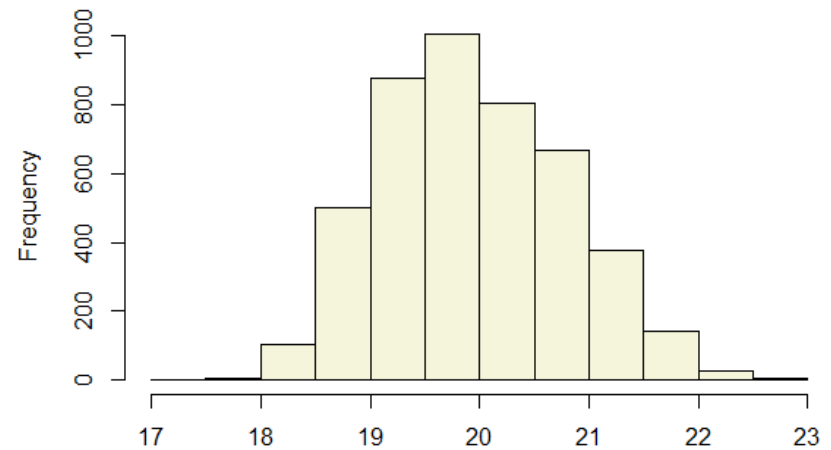
**Yield bu/ac**



**Maturity Date**

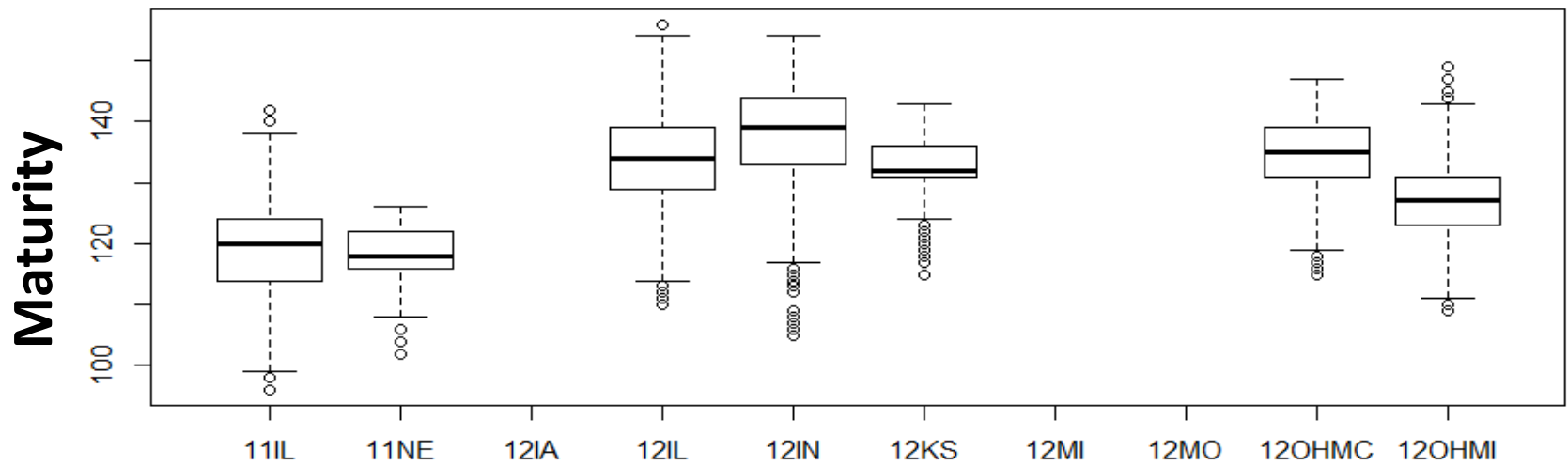
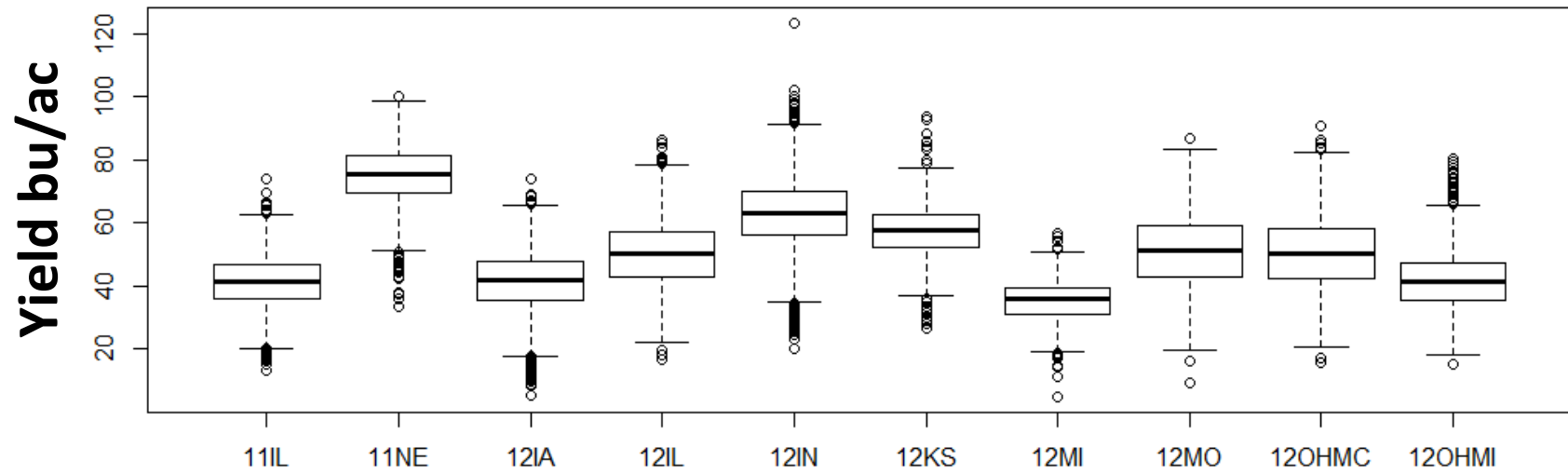


**Protein %**

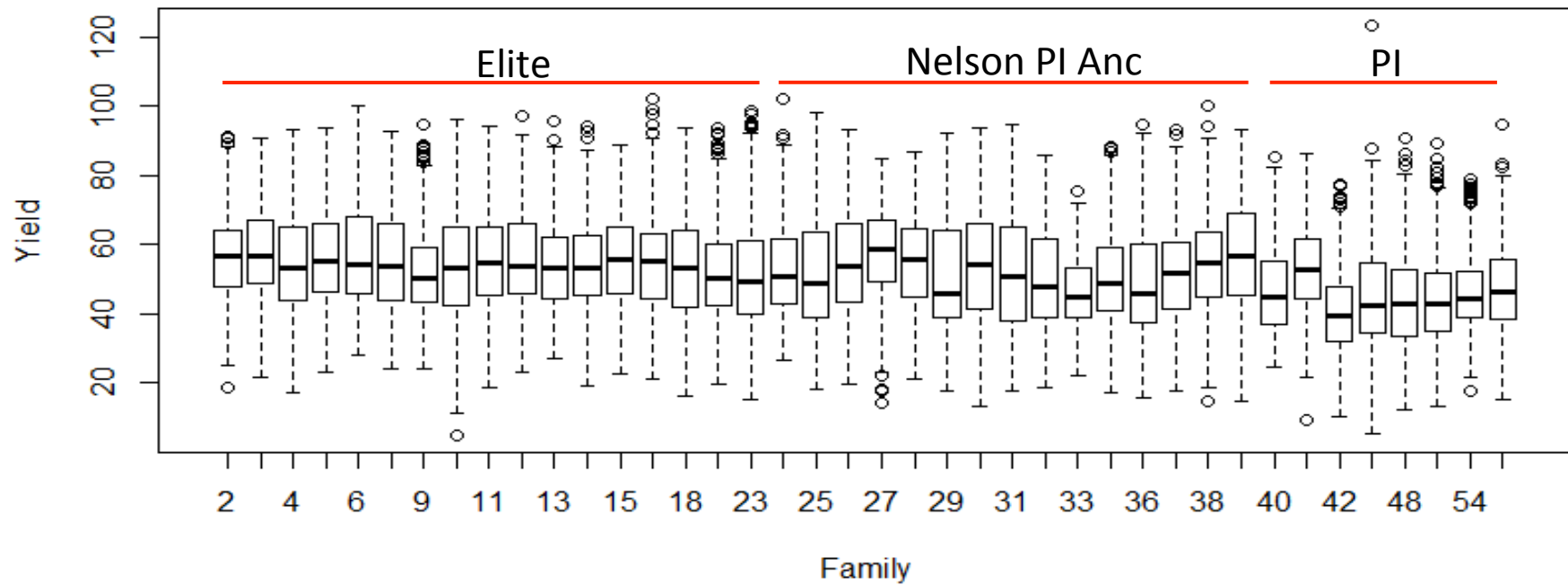


**Oil %**

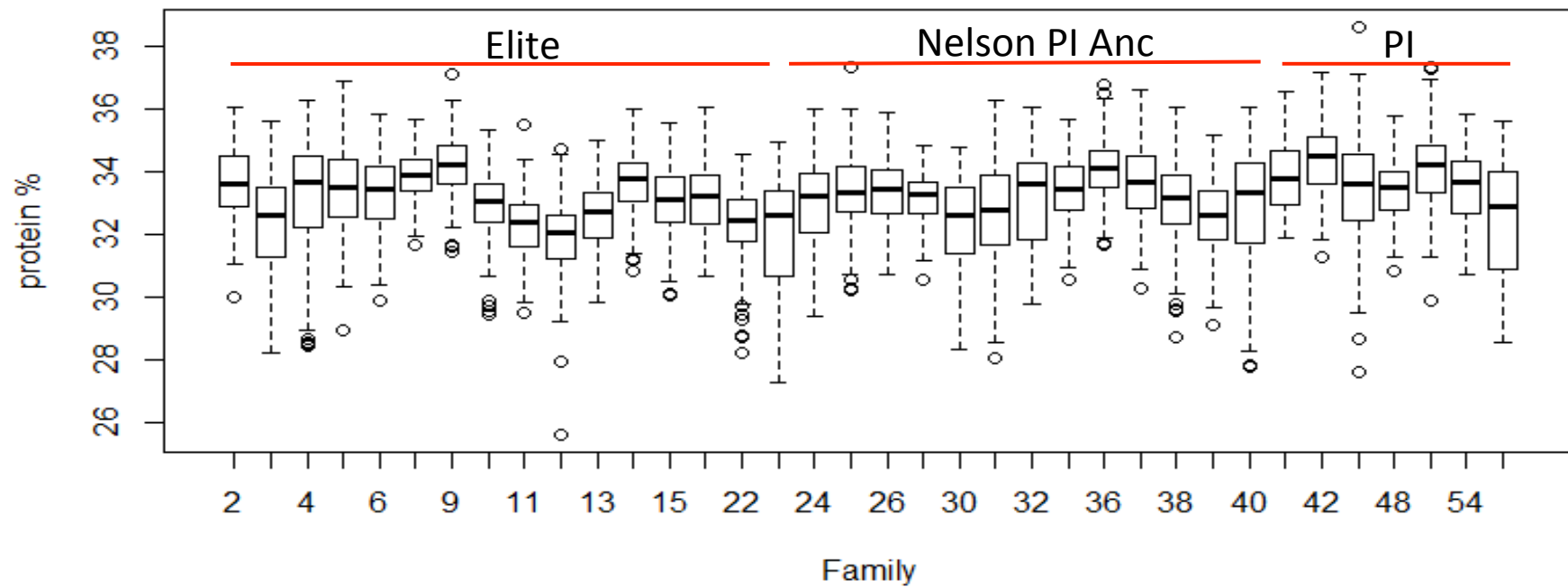
# Distributions of Yield and Maturity by Environment



# Distribution of Yield by Population

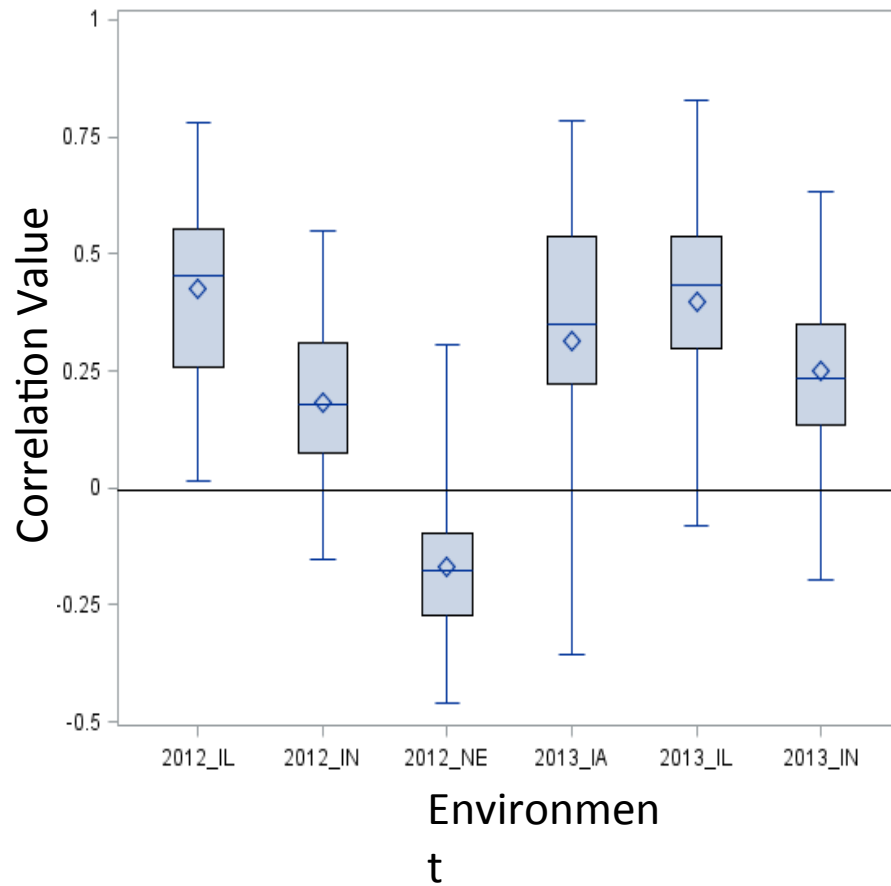


# Distribution of Protein by Population

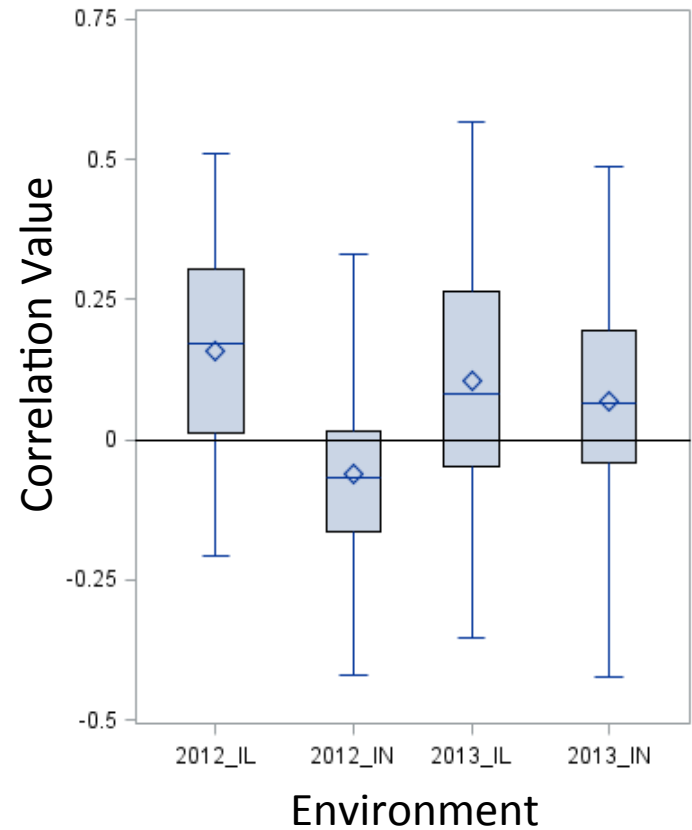


# Distribution of Population Correlations Between Yield and R8 and R1

## Correlation of R8 Date with Yield

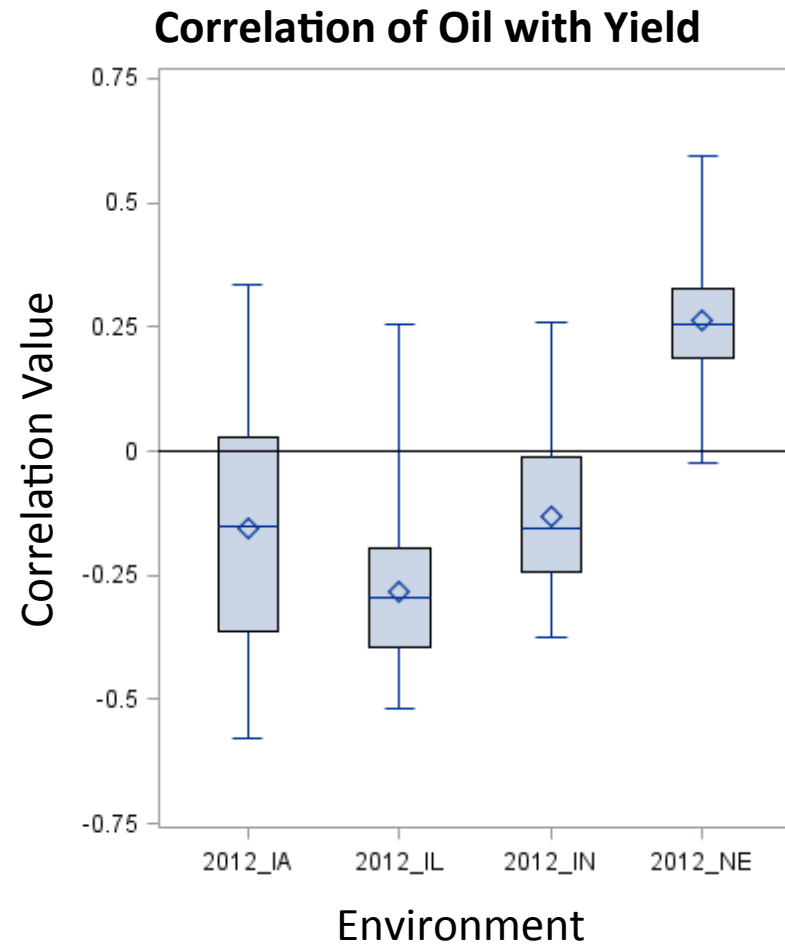
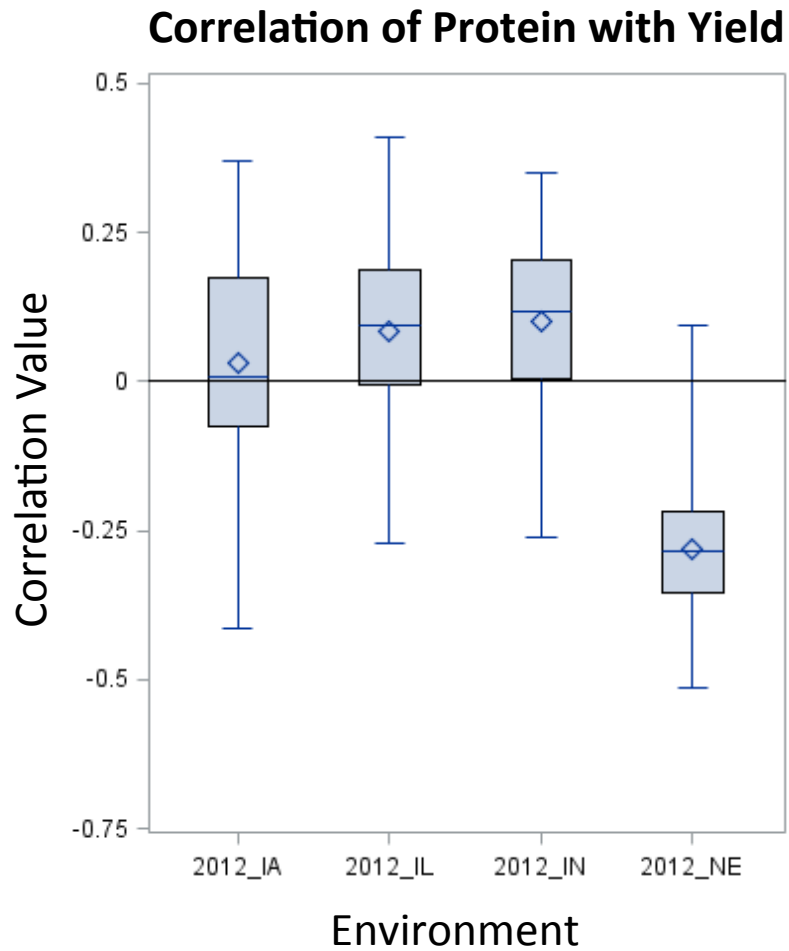


## Correlation of R1 Date with Yield





# Distribution of Population Correlations Between Yield and Protein and Oil

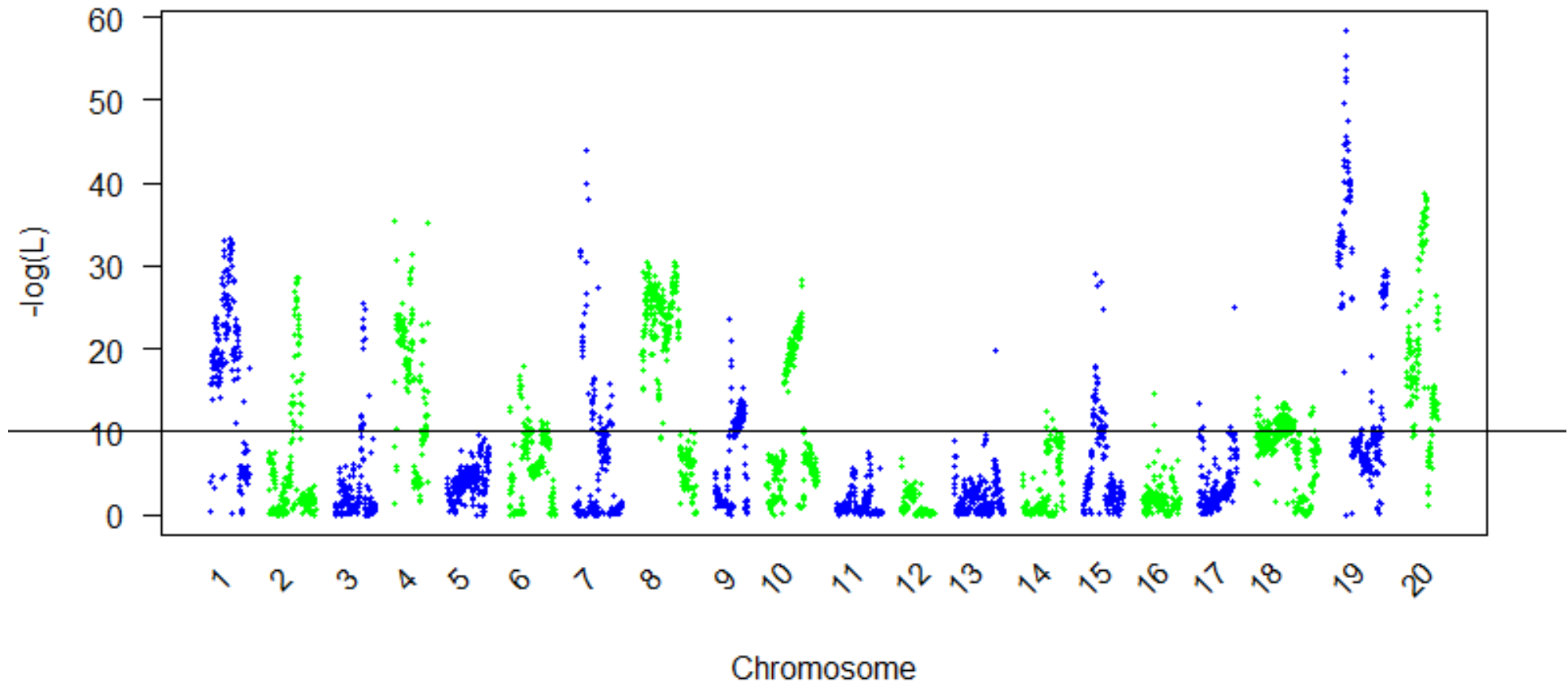


# QTL Scans

---

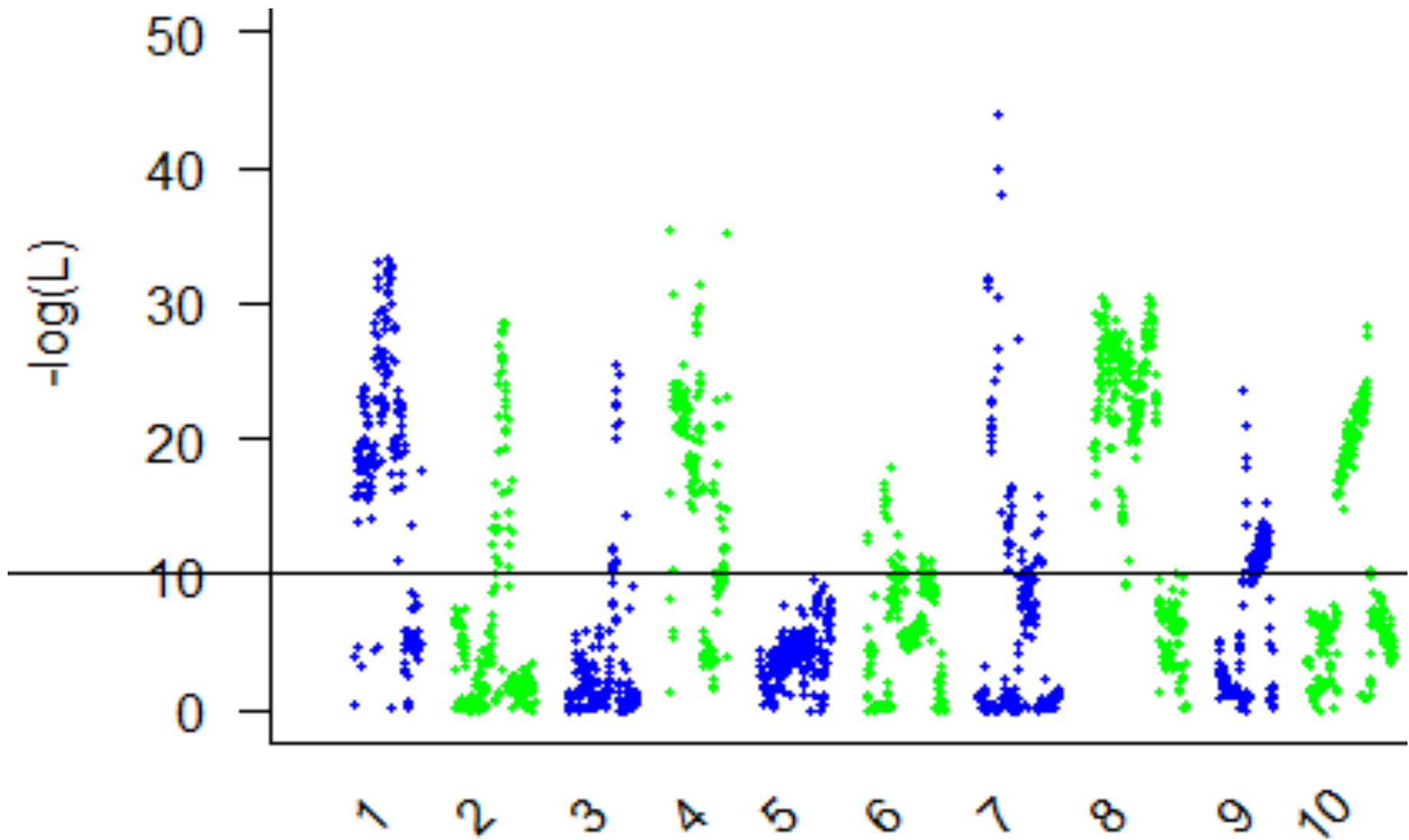
- BLUPs were calculated for RILs across environments from the phenotypic data.
- Standard single marker QTL scans done across populations.
  - The BLUP calculations resulted in all populations having the same mean.
- LOD threshold of 10 (FDR of 0.01).

# Significant Regions for Yield

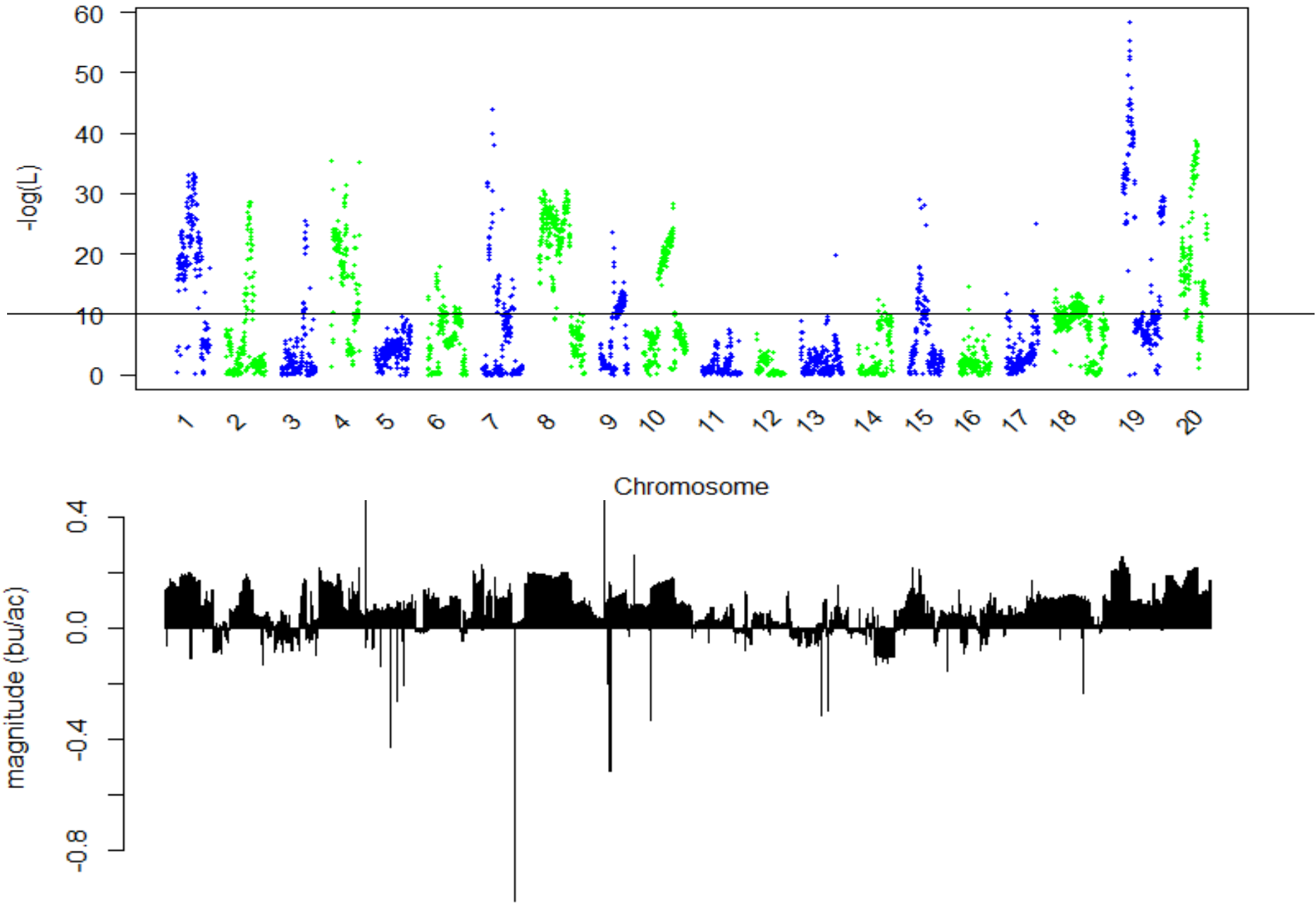


Predicted yields are BLUPs based on harvest yields at 2 locations in 2011 and 8 locations in 2012

# Significant Regions for Yield



# Effects of Allele Substitutions for Yield



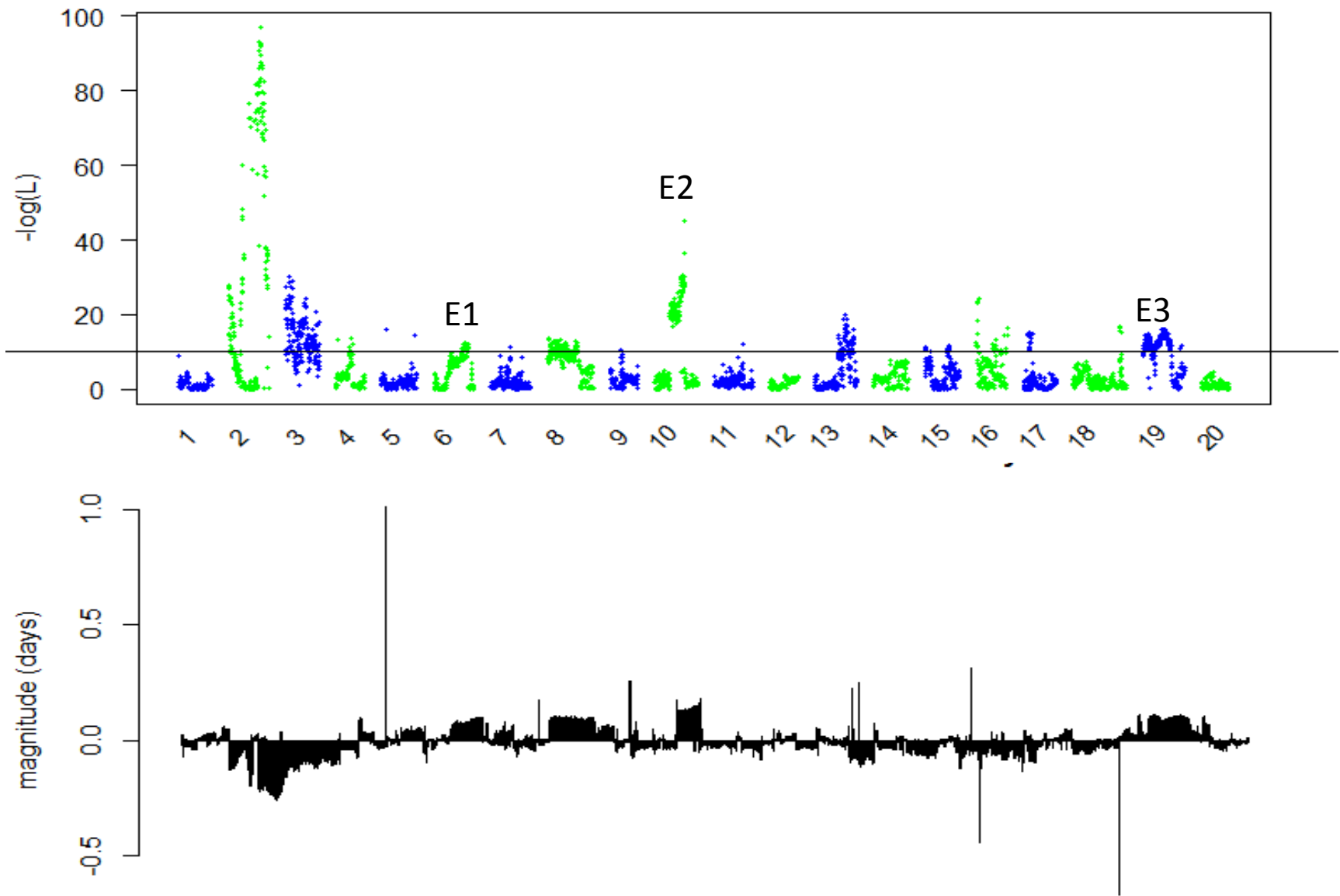


# Why Most Significant Effects from IA3023?

---

- Positive QTL alleles from IA3023 likely segregating in multiple populations.
- Positive effects in nonIA3023 parents are diluted across populations (allele in 2 of 40 populations but linked marker segregating in 20 of 40).
- Problem will be reduced when the 100,000+ markers from parents will be projected onto RILs, resulting in markers in linkage disequilibrium with QTL.

# Effects of Allele Substitutions for Maturity



# Next Steps

---

- Calculate BLUPs with the entire dataset.
- Project the all SNPs from the 50K SNP array and resequencing onto the lines.
- Complete re-genotyping of 287 RILs that failed in the SNP assay.
- Complete the data analysis using the entire data set and more sophisticated models.

# Release of Data

---

- All data and RILs from the experiment will be publically available May 2015.



# Funding





Thanks! Any questions?





# Population Specific Effects for Chromosome 19 QTL

