



CSS 2016 & SEED EXPO

DECEMBER 5-9, 2016 | CHICAGO, IL

**AMERICA'S LARGEST SEED
INDUSTRY CONFERENCE**

The Genomes to Fields Initiative: Progress & Perspective

Natalia de Leon, Joe Gage, Diego Jarquín, Cinta Romay, Aaron J. Lorenz, Ed Buckler, Shawn Kaeppler & G2F Consortium

www.Genomes2Fields.org



Genomes to Fields (G2F) Initiative:

✧ Goal: Catalyze and coordinate research linking genomics and predictive phenomics to expand our understanding of the interacting roles of crop genomes and crop environments (weather and management practices) to improve prediction ability



Schnable, Ware
et al. (2009) Science



Computer scientist

Engineers

Mathematicians

Plant Scientists

Statisticians

Crop Modelers

Soil Scientists

Biologists

Bioinformaticians

Enviromtypers

GENOMES TO FIELDS (G2F)

www.genomes2fields.org

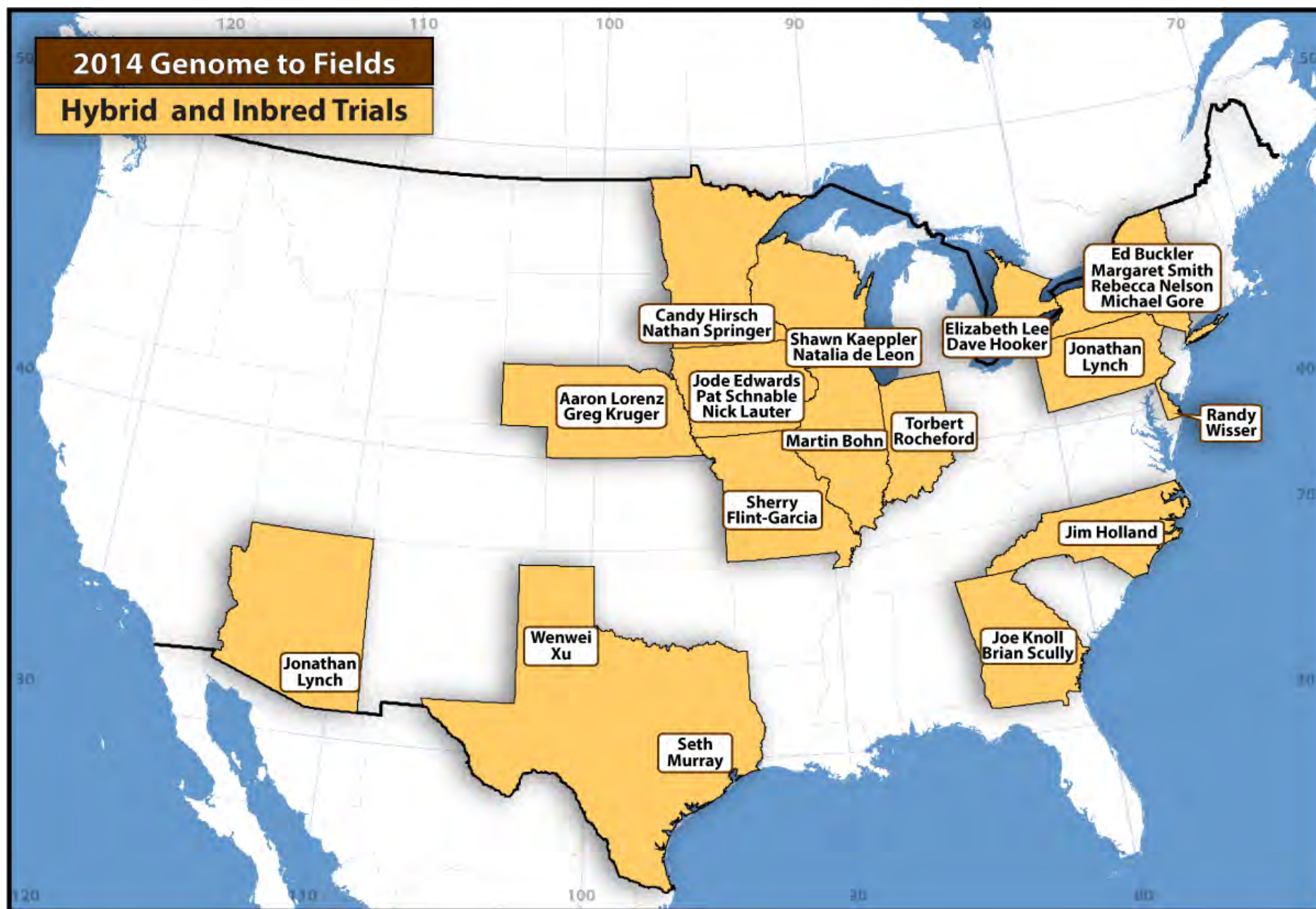
Modeling and
analytical
tools

HT phenotyping
tools

Database
development

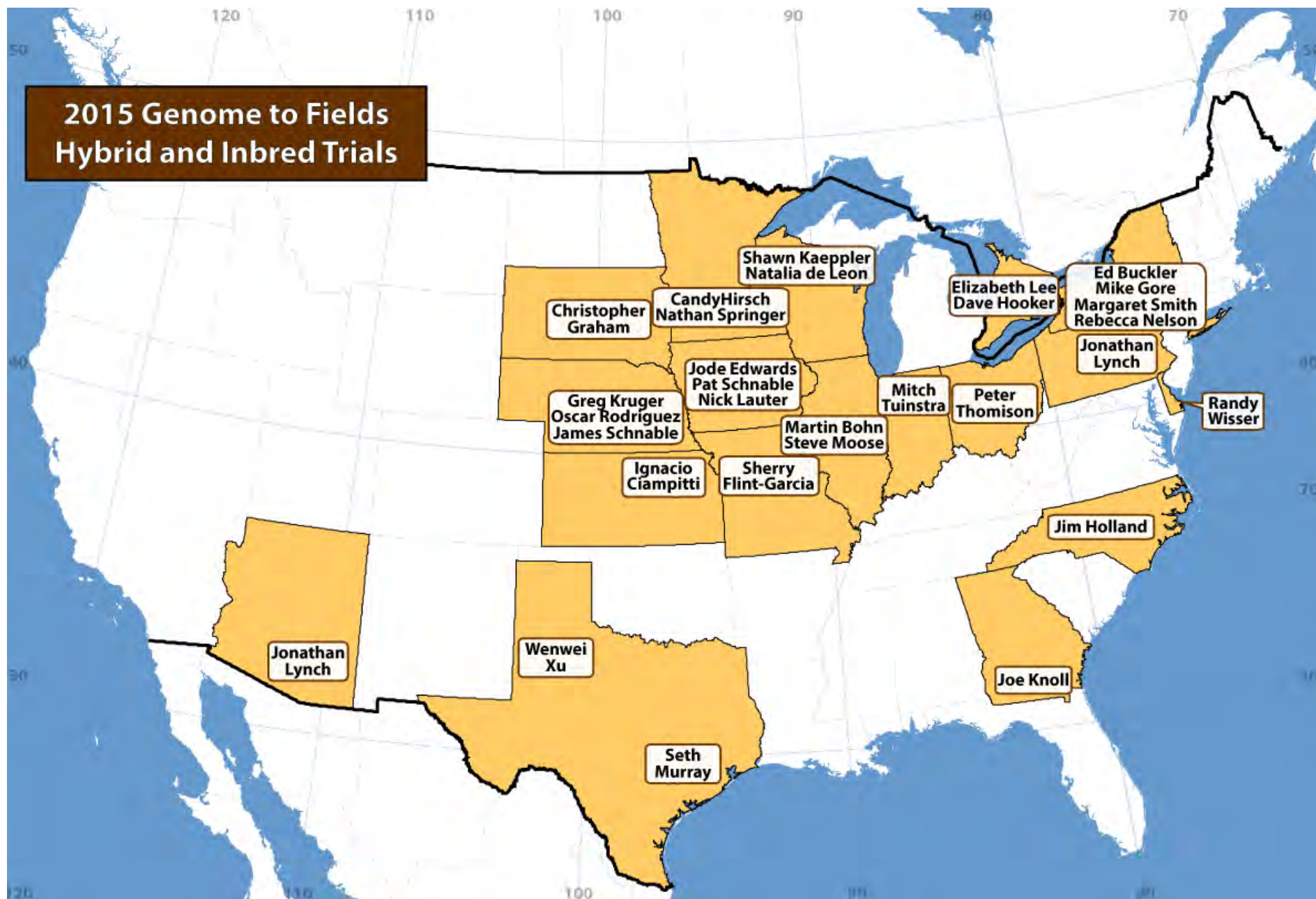
**G X E
Project**

Student Training



Figures courtesy of Darwin Campbell

2015 Genome to Fields Hybrid and Inbred Trials



✧ Additional locations in 2016



✧ Plan for 2017



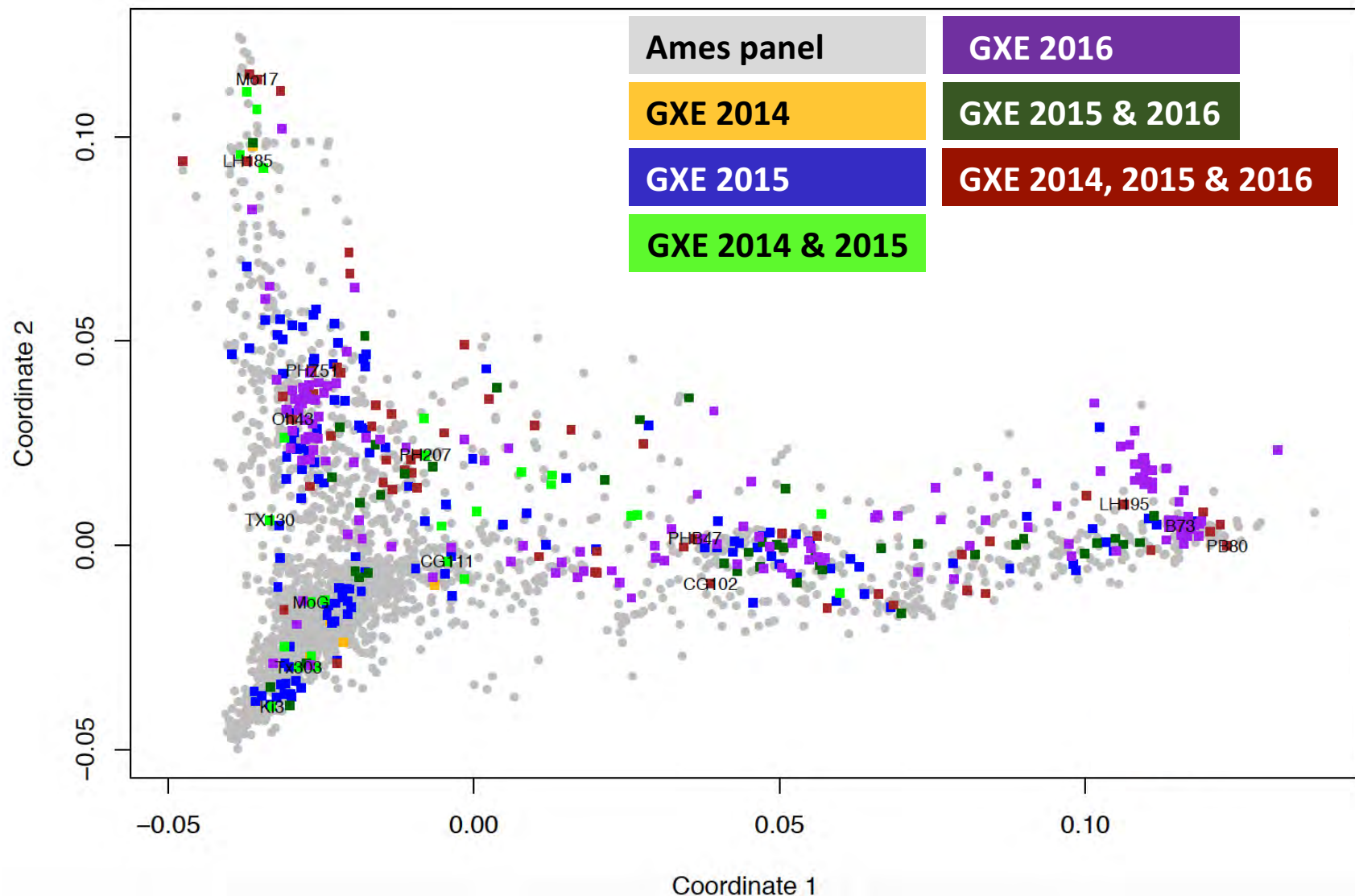
G X E Project – Experimental Description:

- ✧ Each location included ~500 plots at various levels of replication
- ✧ Hybrids involve diverse lines, ex-PVPs, RILs from biparental populations and current breeding materials crossed by relevant testers to move materials from south to north across North America
- ✧ Subset of hybrids were common across all locations

Types of Data - Genotypic Diversity:



Cinta Romay
Cornell Univ



Types of Data - Phenotypic:

✧ Evaluation across all environments included:

- ❖ Days to anthesis
- ❖ Days to silk
- ❖ Ear height
- ❖ Plant height
- ❖ Plot weight
- ❖ Test weight
- ❖ Moisture
- ❖ Grain yield
- ❖ Stand count
- ❖ Root Lodging
- ❖ Stalk Lodging

✧ Additional traits measured at specific locations



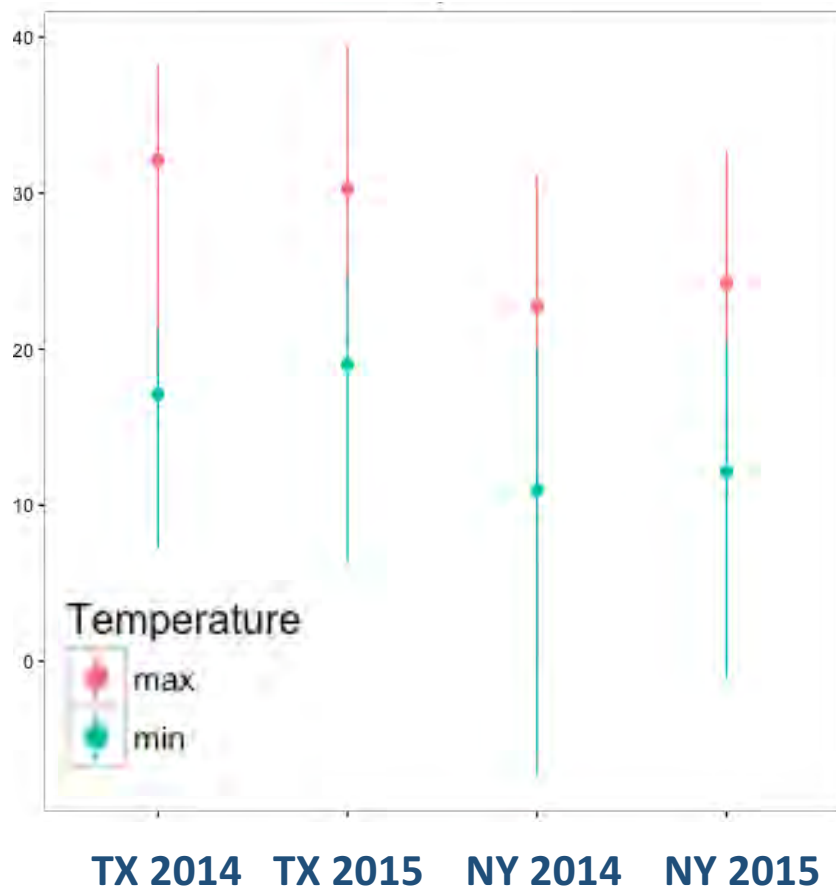
Types of Data - Environmental:

- ✧ Spectrum watchdog at each site – temperature, precipitation, solar radiation, wind speed, and related info + soil characterization

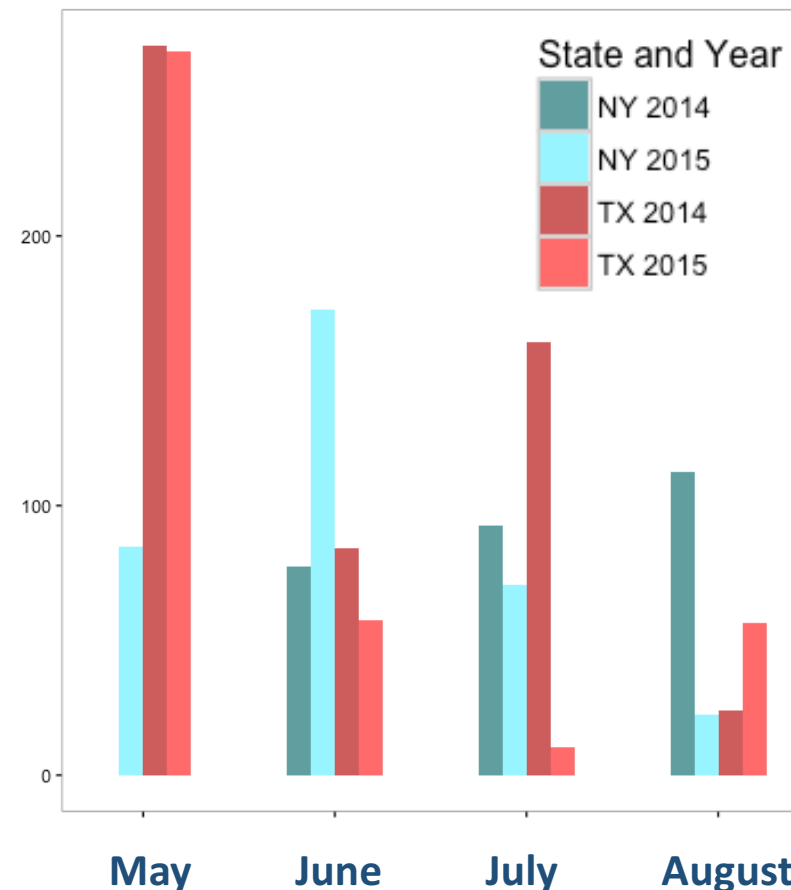


Renee Walton
ISU

Daily Max and Min Temperatures



Accumulated Rainfall



Scale and Scope:

Year	2014	2015	2016
No. of trials	22	27	34
No. of unique locations	21	25	30
No. of states/provinces	13	16	18
No. of Principal Investigators	19	25	30
No. of plots	12,678	13,650	17,040
No of unique inbreds:	380	553	310



The Effect of Selection on GXE – 2014 Data Example

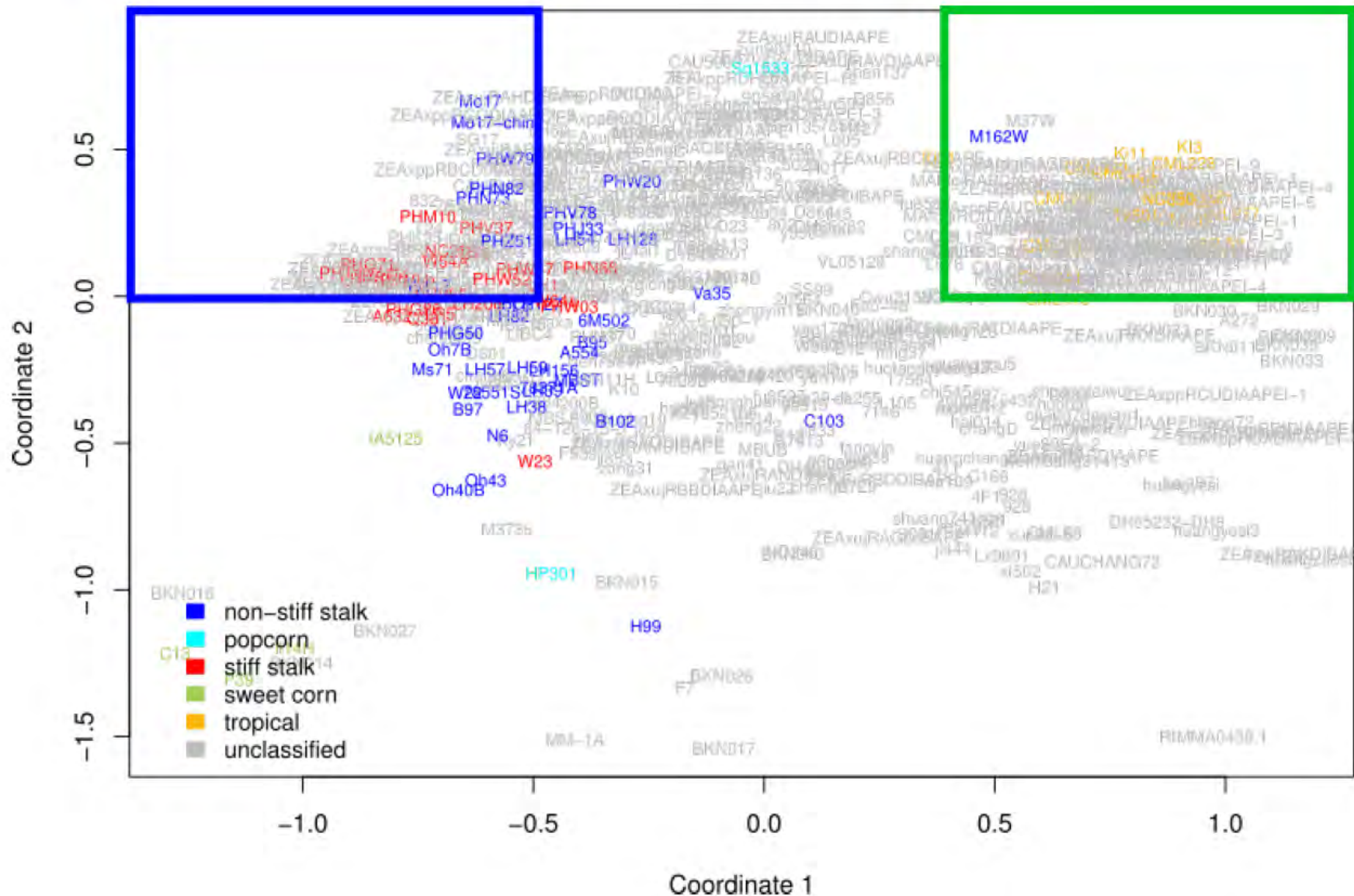
- ✧ GXE has a genetic foundation and therefore is affected by selection
- ✧ Has selection for high productivity decreased available genetic variation that controls the ability of crop species to adapt to variable environments?
- ✧ Also, understanding the types of genetic architectures and modulation mechanisms controlling G X E will provide opportunities to enhance prediction ability on variable environments



Identify regions selected for temperate climates based on Fst between Temperate and Tropical materials:



Cinta Romay
Cornell Univ



Analysis – Model Details:

✧ Use genotypic scores in **high** and **low** Fst loci to calculate GxE attributable to those regions:

$$y = \mu + L + E + g + g_H \times E + g_L \times E + \varepsilon$$

where:

- μ : overall mean
- y : phenotype: Grain Yield and Plant Height
- L : line (hybrid) effect
- E : environmental effect
- g : overall genotypic effect
- $g_H \times E$: High Fst by Environment effect
- $g_L \times E$: Low Fst by Environment effect
- ε : residual error

✧ **Low Fst SNPs** (n ~200k) sampled to eq. **low High SNPs** (n ~9k)

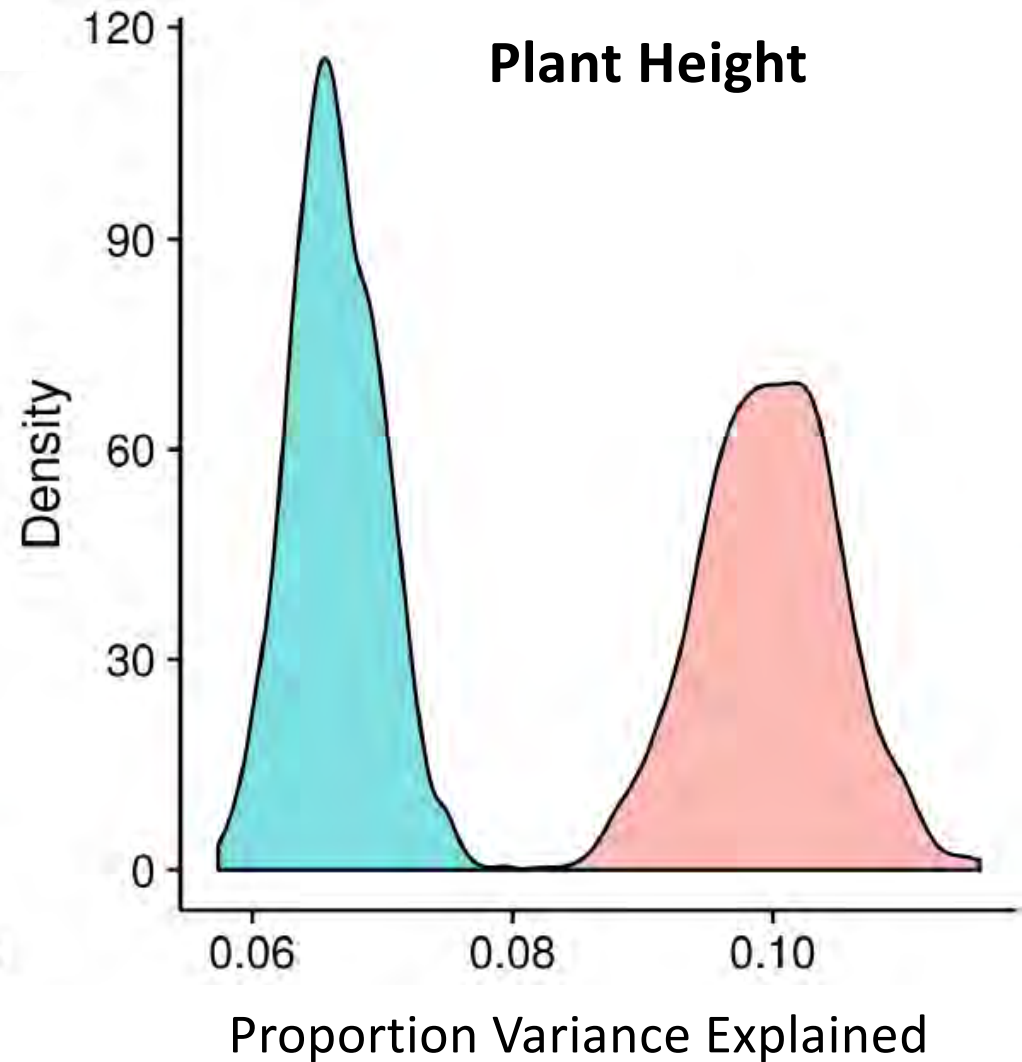
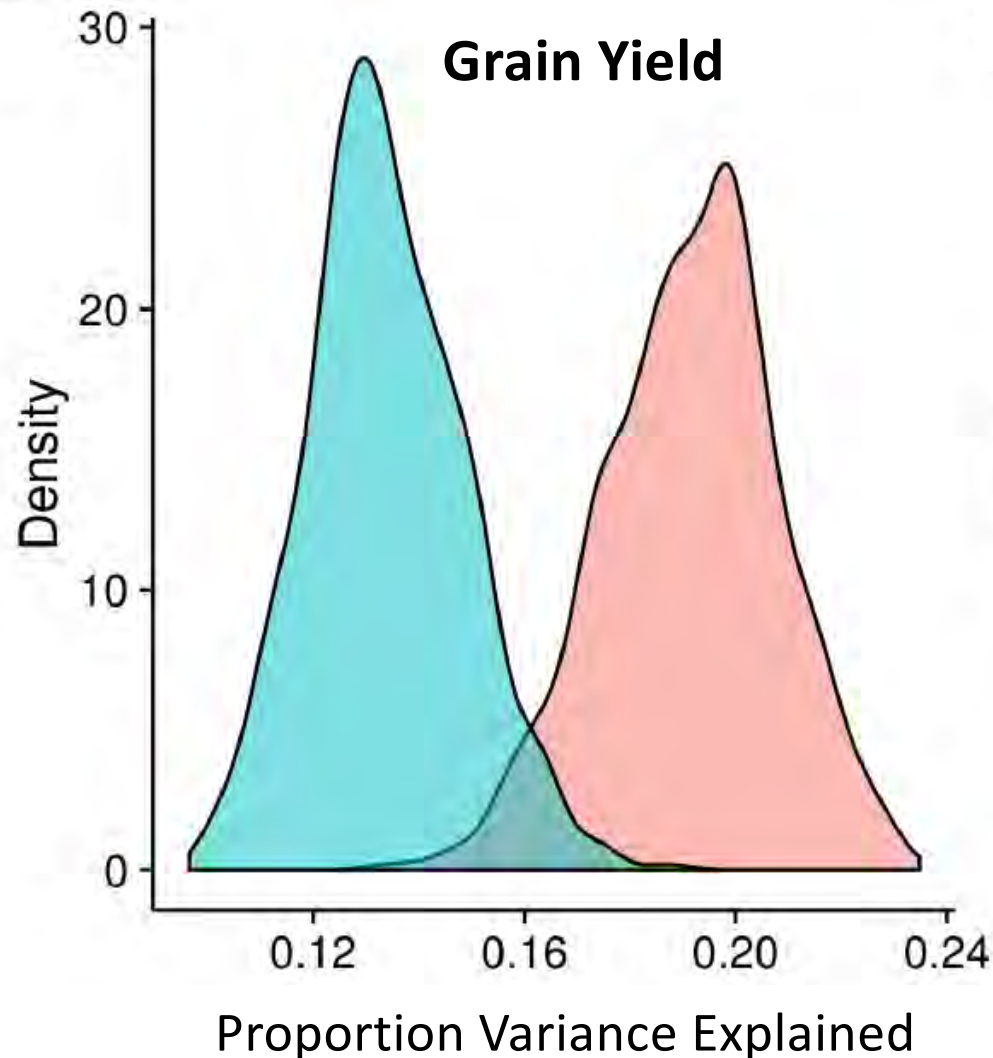


Diego Jarquin
Univ of NE



Results – Assessment of Effect of Selection:

High Fst Regions - Low Fst Regions



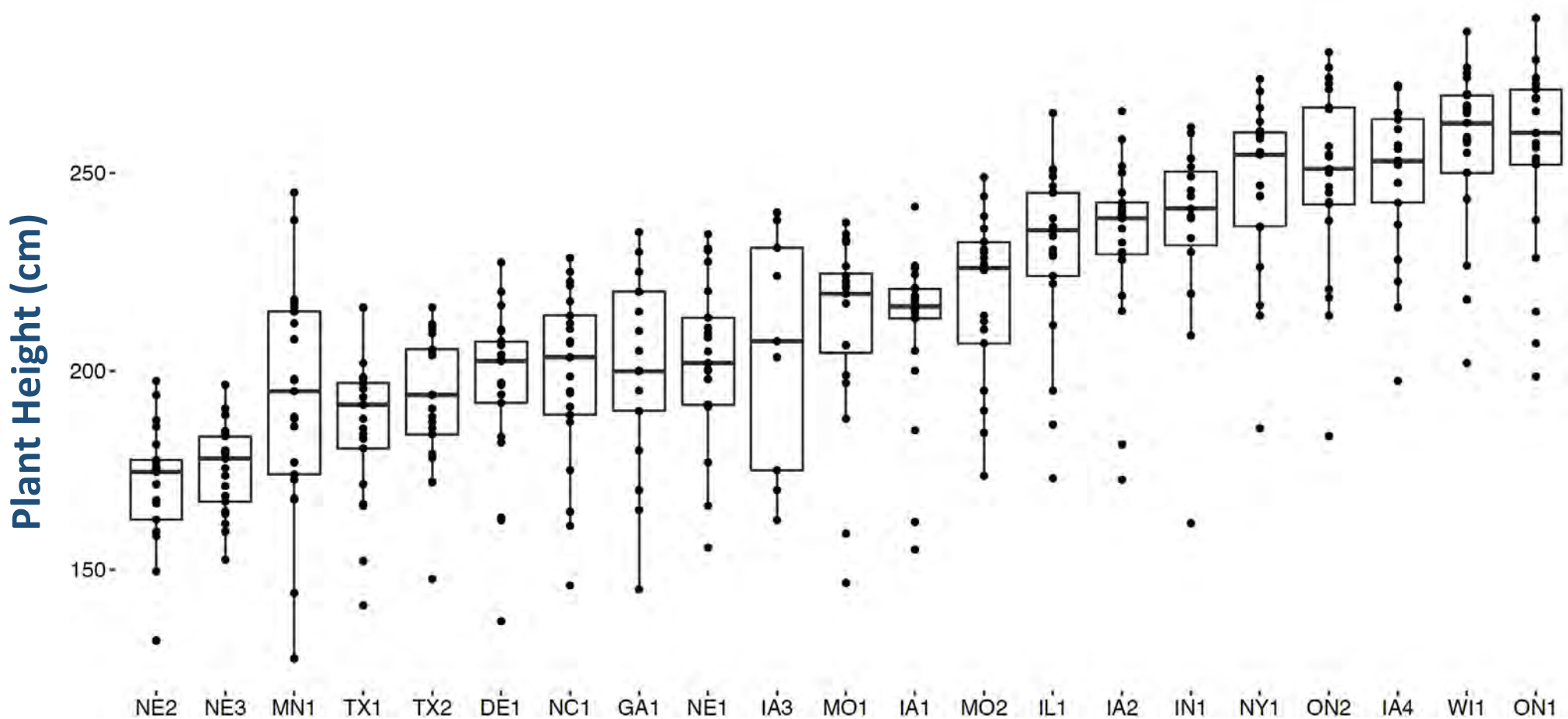
Control & Modulation of GXE – 2014 Data Example

- ✧ Is G X E modulated by the genetic architecture of the trait or are there alternative sources of control?
 - ❖ If G X E is mostly due to the genetic architecture of the trait → current models that focus on assigning values to polymorphisms for prediction would work
 - ❖ If G X E is due to more complex regulatory mechanisms → information needs to be included in the model to account for G X E
 - ❖ It is important not only to make decisions of what to select, but also of what type of variation is important to keep in the germplasm for long term breeding gains



Quantifying Stability:

- ✧ 20 checks evaluated at 21 environments
- ✧ Mean of checks at each location → reference value for location

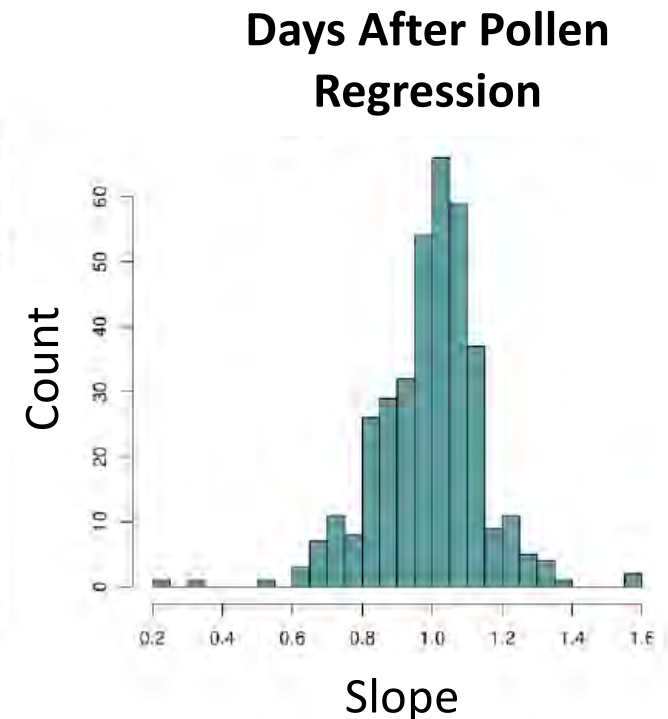
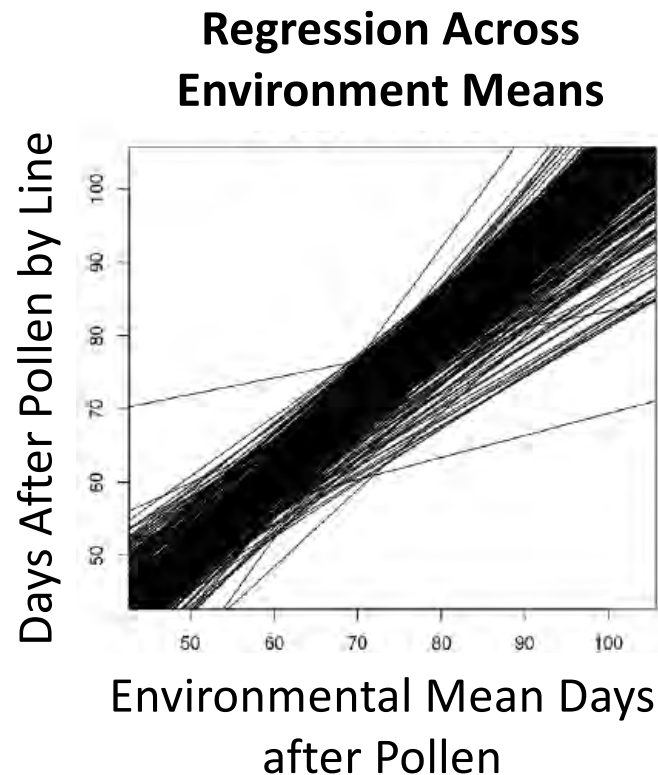
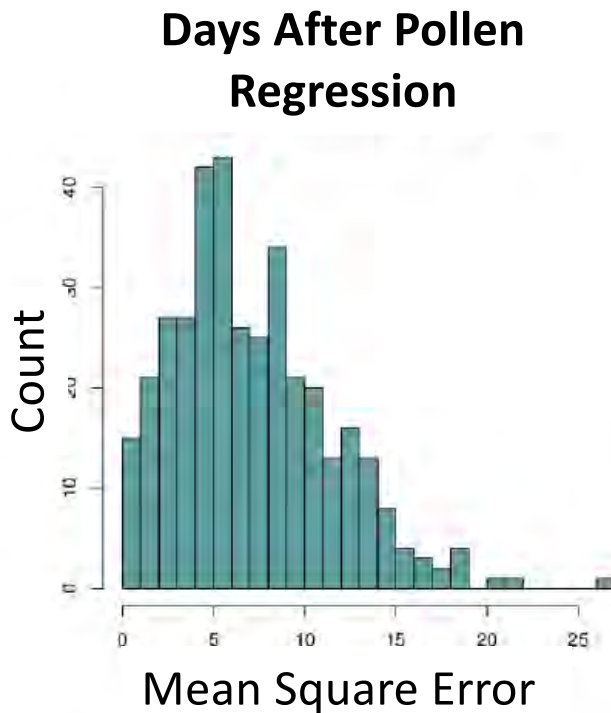


Definition of Stability Parameters:



Joe Gage
Univ of WI

- ✧ For each unique hybrid at 4 locations or more:
 - ❖ Plot hybrid value at each location against the location's reference value (average check performance)
 - ❖ Fit a line
 - ❖ Extract regression parameters: slope and MSE



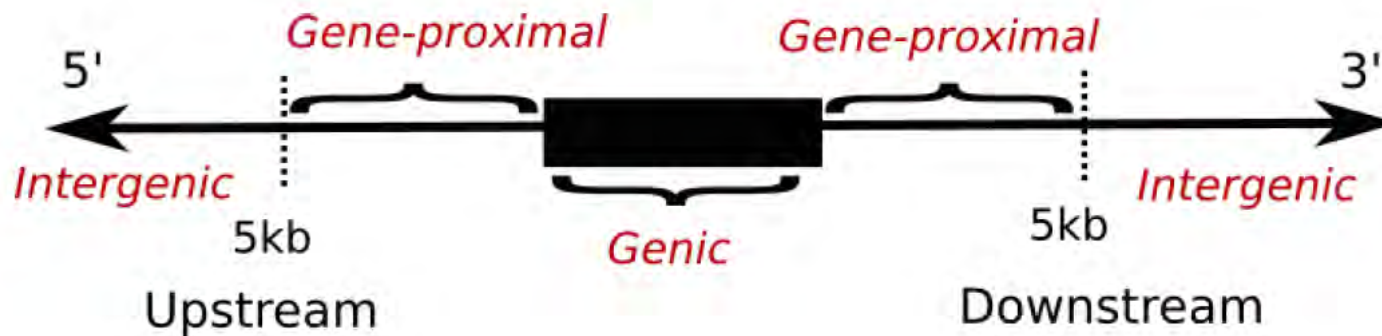
Relationship between Regression Parameters and Stability:

- ✧ Type II Stability: Environmental response has a slope = 1 → quantified by slope
- ✧ Type III Stability: Variation around the regression line → quantified by mean square error (MSE)
- ✧ Each regression parameter (slope & MSE) and trait per se were evaluated for the following traits separately:
 - ❖ Ear height (cm)
 - ❖ Plant height (cm),
 - ❖ Number of days to silk
 - ❖ Number of days to anthesis
 - ❖ Yield

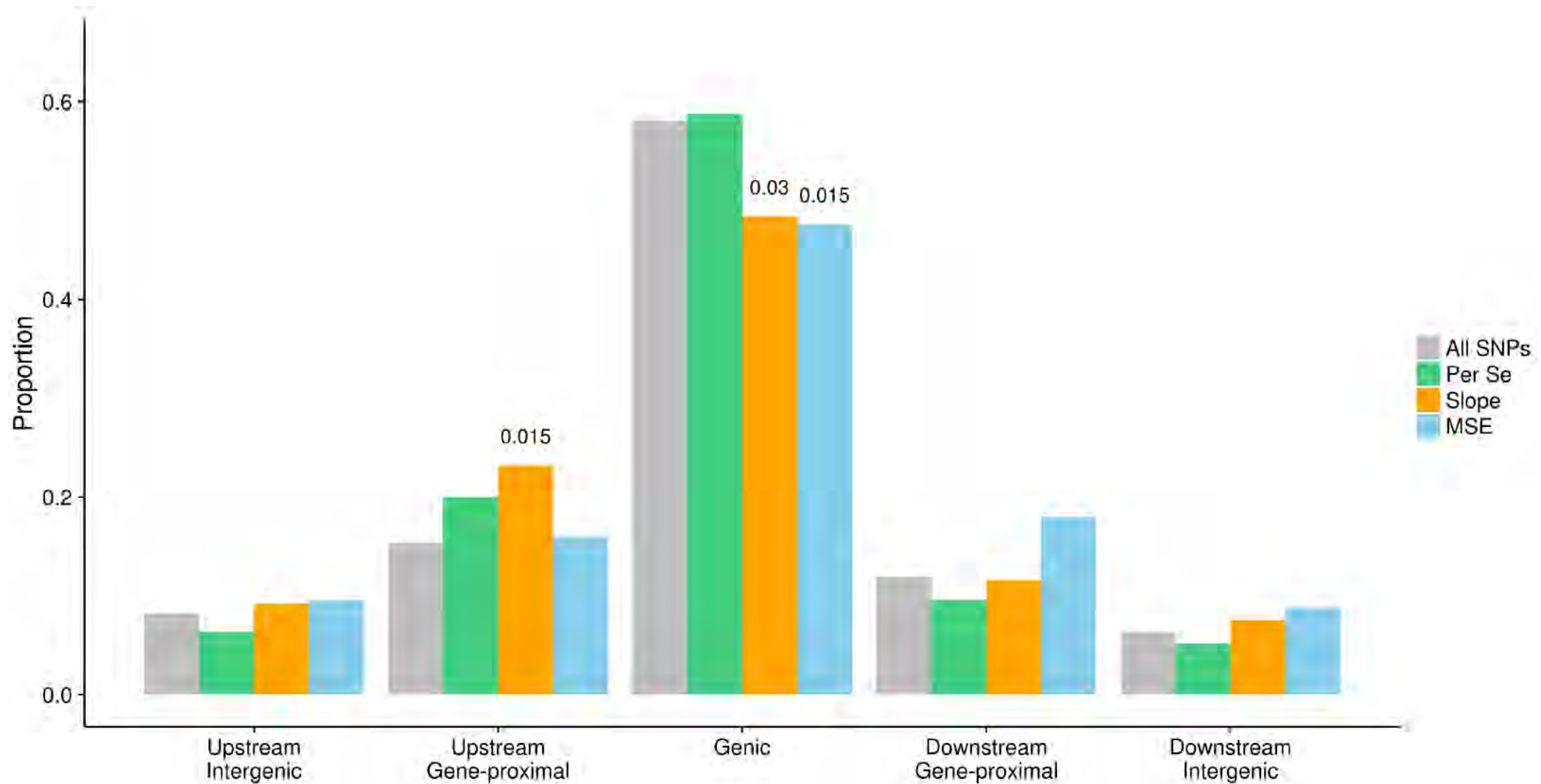


GWAS Analysis:

- ✧ The 50 SNPs with the lowest p-values from each GWAS for slope and MSE were pooled and also GWAS of traits per se
- ✧ Classify associated SNPs as Upstream/Downstream, Intergenic/Gene-proximal/Genic and compare to null distribution of classification for all SNPs (~ 400k)



Patterns of Functional Variation:



Final Remarks:

- ✧ The G2F G X E Maize Project provides a useful platform to test important hypotheses related to the differential response of cultivars across environments
- ✧ Proportion of the phenotypic variance explained for G X E appears to differ for regions that have undergone differential selection for adaptation to temperate regions
- ✧ Loci associated with type II stability show enrichment in upstream gene-proximal regions and depletion in genic regions



GxE Consortium: Data Usage Disclaimer

This presentation includes data analysis and interpretation conducted by the presenter and does not necessarily reflect the observations and conclusions of the GxE Consortium.



G X E Cooperators

Principal Investigators who grew GxE trials in 2014-2016

- | | | |
|------------------------------|------------------------------|------------------------------|
| ✧ Martin Bohn (UIUC) | ✧ Joe Knoll (ARS - UGA) | ✧ Oscar Rodriguez (UNL) |
| ✧ Ed Buckler (ARS - Cornell) | ✧ Judith Kolkman (Cornell) | ✧ Cinta Romay (Cornell) |
| ✧ Ignacio Ciampitti (KSU) | ✧ Greg Kruger (UNL) | ✧ James Schnable (UNL) |
| ✧ Jode Edwards (ARS - ISU) | ✧ Nick Lauter (ARS - ISU) | ✧ Pat Schnable (ISU) |
| ✧ Sherry Flint-Garcia (ARS) | ✧ Liz Lee (Guelph) | ✧ Brian Scully (ARS) |
| ✧ Christopher Graham (SDSU) | ✧ Natalia de Leon (UW) | ✧ Rajandeep Sekhon (Clemson) |
| ✧ Mike Gore (Cornell) | ✧ Sanzhen Liu (KSU) | ✧ Margaret Smith (Cornell) |
| ✧ Candy Hirsch (UMN) | ✧ Argelia Lorence (AR-State) | ✧ Nathan Springer (UMN) |
| ✧ Jim Holland (ARS - NCSU) | ✧ Aaron Lorenz (UMN) | ✧ Kurt Thelen (MSU) |
| ✧ Elizabeth Hood (AR-State) | ✧ Jonathan Lynch (PSU) | ✧ Peter Thomison (OSU) |
| ✧ David Hooker (Guelph) | ✧ Steve Moose (UIUC) | ✧ Mitch Tuinstra (Purdue) |
| ✧ Fiona Goggin (Univ AR) | ✧ Seth Murray (TAMU) | ✧ Jason Wallace (UGA) |
| ✧ Shawn Kaeppler (UW) | ✧ Rebecca Nelson (Cornell) | ✧ Randy Wisser (UDel) |
| | ✧ Torbert Rocheford (Purdue) | ✧ Wenwei Xu (TAMU) |



Genomes To Fields Collaborators

† G2F Executive Committee members *GxE Coordinating Groups ~G2F co-lead

- | | | |
|------------------------------|--------------------------------|------------------------------|
| ✧ Naser Alkhalifah (ISU) | ✧ Diego Jarquin (UNL)* | ✧ James Schnable (UNL) |
| ✧ Martin Bohn (UIUC)* | ✧ Shawn Kaeppler (UW)† | ✧ Pat Schnable (ISU)†~ |
| ✧ Ed Buckler (ARS)† | ✧ Joe Knoll (ARS) | ✧ Brian Scully (ARS) |
| ✧ Darwin Campbell (ISU)* | ✧ Judith Kolkman (Cornell) | ✧ Rajandeep Sekhon (Clemson) |
| ✧ Ignacio Ciampitti (KSU) | ✧ Greg Kruger (UNL) | ✧ Kevin Silverstein (UMN) |
| ✧ James Clohessy (Cornell) | ✧ Nick Lauter (ARS) | ✧ Margaret Smith (Cornell) |
| ✧ Liang Dong (ISU) | ✧ Carolyn Lawrence-Dill (ISU)* | ✧ Edgar Spalding (UW) |
| ✧ Jode Edwards (ARS)* | ✧ Liz Lee (Guelph)* | ✧ Nathan Springer (UMN)† |
| ✧ David Ertl (IA Corn)*† | ✧ Zhizhai Liu (TAMU) | ✧ Srikant Srinivasan (ISU)* |
| ✧ Sherry Flint-Garcia (ARS)* | ✧ Natalia de Leon (UW)*†~ | ✧ Yiwei Sun (ISU)* |
| ✧ Joseph Gage (UW) | ✧ Alex Lipka (UIUC) | ✧ Kurt Thelen (MSU) |
| ✧ Jack Gardiner (ISU)* | ✧ Argelia Lorence (AR-State) | ✧ Peter Thomison (OSU) |
| ✧ Fiona Goggin (Univ AR) | ✧ Aaron Lorenz (UMN)* | ✧ Kelly Thorp (ARS) |
| ✧ Byron Good (Guelph) | ✧ Jonathan Lynch (PSU)† | ✧ Mitch Tuinstra (Purdue) |
| ✧ Mike Gore (Cornell) | ✧ Nathan Miller (UW) | ✧ Jason Wallace (UGA) |
| ✧ Christopher Graham (SDSU) | ✧ Steve Moose (UIUC) | ✧ Renee Walton (ISU) |
| ✧ Patricio Grassini (UNL) | ✧ Seth Murray (TAMU)* | ✧ Rod Williamson (IA Corn) |
| ✧ Jerry Hatfield (ARS) | ✧ Rebecca Nelson (Cornell) | ✧ Randy Wisser (UDel)* |
| ✧ Candy Hirsch (UMN) | ✧ Torbert Rocheford (Purdue) | ✧ Wenwei Xu (TAMU) |
| ✧ Jim Holland (ARS) | ✧ Oscar Rodriguez (UNL) | ✧ Cheng-Ting Yeh (ISU) |
| ✧ Elizabeth Hood (AR-State) | ✧ Cinta Romay (Cornell)* | ✧ Jianming Yu (ISU) |
| ✧ David Hooker (Guelph) | ✧ Emily Rothfusz (UW)* | |



Genomes To Fields Sponsors



United States Department of Agriculture
National Institute of Food and Agriculture

