The Effect of Artificial Selection on Phenotypic Plasticity: The Genotype by Environment Interaction Project in Maize

de Leon, N., D. Jarquin, M.C. Romay, J. Gage, S.M. Kaeppler, E.S. Buckler, A.J. Lorenz, G2F Consortium

PAG CSSA: Translational Genomics Workshop
January 9th, 2016

www.Genomes2Fields.org

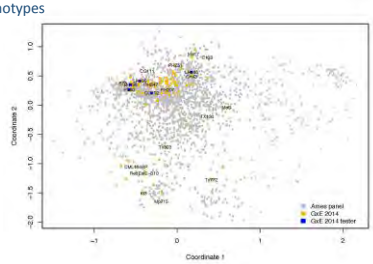
Premise:

- Phenotypic plasticity (or G X E) → ability of a single genotype to produce different phenotypes in response to different environments.
- Plasticity has a genetic foundation and therefore is affected by selection
- This adaptation capacity is directly related to the extent and nature of standing genetic variation
- Has selection for high productivity decrease available genetic variation that controls plasticity in crop species?
- Also understanding the types of genetic architectures and modulation mechanisms controlling G X E will provide opportunities to enhance prediction ability on variable environments
- Using the platform of the **Maize G X E project** that is part of **Genomes to Fields Initiative** to test these hypotheses

Experimental Description - Genotypes:

Cinta Romay
Cornell Univ

- 853 hybrids from 374 inbreds crossed by 4 testers and evaluated across 22 environments (243 hybrids per location)
- GBS data collected for all lines
- Synthetic genotypes of hybrids created from inbred parent genotypes



Experimental Description – Environments in 2014:

Figure generated by Darwin Campbell

Experimental Description - Phenotypes:

- Plant Height
- Days to Pollen
- Days to Silk
- Yield (bushels/acre)

Hypothesis 1:

- Selection tends to reduce variation in genomic regions that are relevant for that trait, therefore regions that underwent reduction in variability due to selection should explain less of the G X E variability observed compared to regions under no selection

Analysis – Definition of Differential F_{ST} Regions:

Cinta Romay
Cornell Univ

- Categories were defined by comparing F_{ST} on 20 SNP windows for 60 inbred lines from the HapMap (30 on each extreme tropical vs temperate) based on MDS distance (confirmed by Jeff Ross-Ibarra)
- Contrasting F_{ST} regions: Low (F_{ST} <0.15) and High (F_{ST}>0.5) SNPs in the genome across all hybrids and environments
- Set of 9,194 SNPs with high F_{ST} values and random samples of 9,194 SNPs from low F_{ST}

Analysis – Variance Components Analysis:

Diego Jarquin
Univ of NE

- Use variance components approach in Gusev et al., 2014 (Am. J. of Hum Gen, 95:535-552)
- Method calculates r^2 between the phenotype and the prediction for SNPs from specific functional category (in our case differential F_{ST})
- We analyzed variance components to evaluate if interaction between the studied regions (i.e.: differentially selected regions) and environment show significant differences in the amount of phenotypic variability explained

Analysis – Model Details:

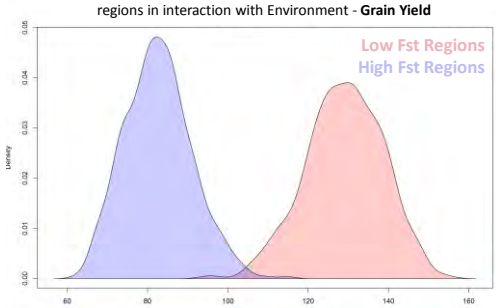
- Random effects model:

$$y = L + E + g + g_L \times E + g_H \times E + e$$

- where:
  - y = is the vector of phenotypes
  - L & E = vectors of hybrid and environmental effects, such that L~N(0, $I\sigma_L^2$) and E~N(0, $I\sigma_E^2$)
  - g = the vector of genomic values following a multivariate normal distribution such that g ~ N(0, $G\sigma_g^2$), G being the Genomic Relationship Matrix computed using all the available markers
  - The interaction terms were included through covariance structures (following Jarquin et al., 2014). Thus g_L X E~N(0,Z_g G_L Z′ g #Z_g Z′$_g$) and g_H X E~N(0,Z_g G_H Z′ g #Z_g Z′$_g$)
  - G_H computed using the 9,194 SNPs with high F_{ST} values and G_L computed taking random samples of 9,194 SNPs from the low F_{ST}
  - analysis repeated 1,000 time

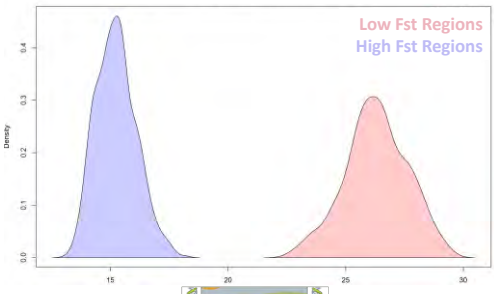
Results – Assessment of Effect of Selection:

Posterior distribution of estimated variance components for regions in interaction with Environment - **Grain Yield**



Results – Assessment of Effect of Selection:

Posterior distribution of estimated variance components for regions in interaction with Environment - **Plant Height**

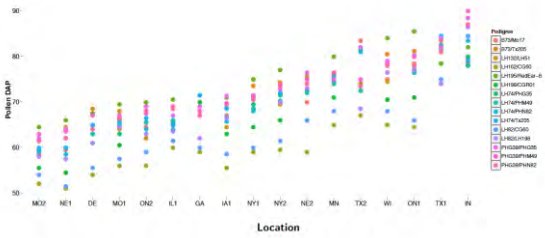


Hypothesis 2:

- 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 (combination of alleles, epistasis, etc) → 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
  - 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

Analysis – Quantifying Stability:

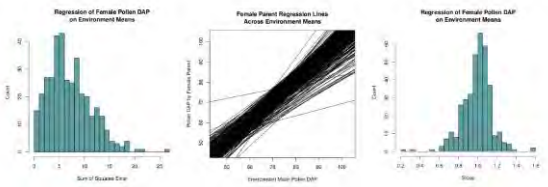
- 12 checks evaluated at 21 environments or more
- Average check at each location → reference value for location



Analysis – Definition of Parameters:

Joe Gage
Univ of WI

- For each unique hybrid at 3 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 SSE

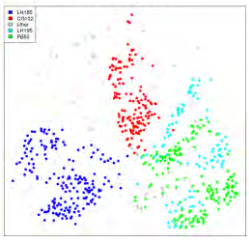


Relationship between Regression Parameters and Stability:

- Type I/II Stability: Environmental response is parallel to mean response → quantified by slope
- Type III Stability: Small among-environment variance → quantified by SSE
- Each regression parameter was evaluated for each trait, resulting in eight parameter-trait combinations to be mapped separately

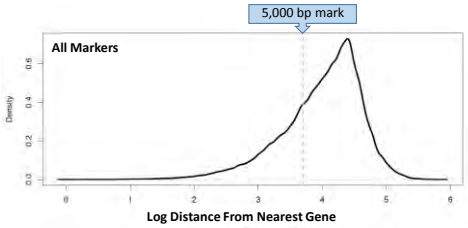
Analysis - GWAS:

- Each phenotype and parameter-trait combination mapped separately using TASSEL 5.0
- Evaluated both additive and dominance effects
- Kinship, as well as 3 principal components covariates used to compensate for strong population structure (four males)

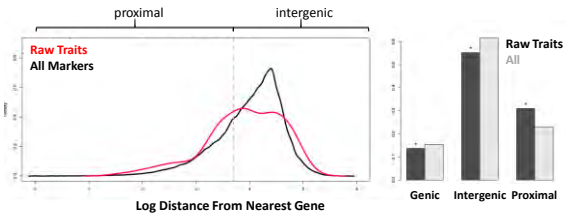


Results - Patterns of Functional Variation:

- Classified SNPs based on their distance to the nearest gene model (following Wallace et al., 2014 – Plos Genetics):
  - Genic: Inside a gene
  - Proximal: 1-5,000 base pairs from closest gene
  - Intergenic: >5,000 base pairs from closest gene
  - Use all markers (~420k) to form a null distribution

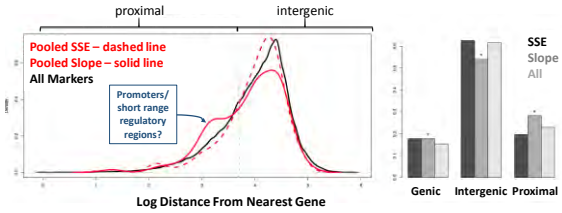


Results – Patterns of Functional Variation:



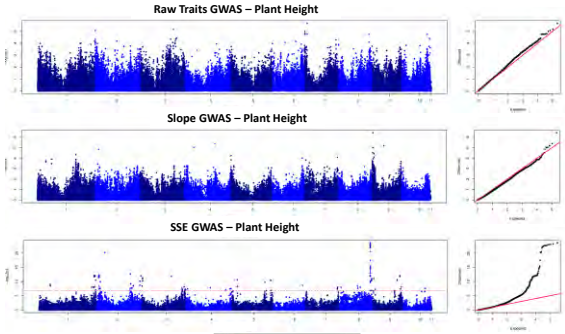
→ 100 SNPs with the lowest p-values for each trait per se were classified as intergenic, gene-proximal or genic

Results – Patterns of Functional Variation:



→ 100 SNPs with the lowest p-values for each trait-parameter combination were classified as intergenic, gene-proximal or genic

Results – Are Traits *Per Se* and Stability Controlled by Similar Regions?



Take Home Messages:

- ✧ Proportion of the phenotypic variance explained for G X E differs for regions that have undergone differential selection in this dataset
- ✧ Control of stability for the traits evaluated appears to involve variants located mostly in intergenic regions
- ✧ No obvious co-localizations of stability and phenotype *per se* associations
- ✧ Need further analysis of the ontology of potential regions associated with control of plasticity

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

Researchers who grew GxE trials in 2014 and 2015

- ✧ Martin Bohn (UIUC)
- ✧ Ed Buckler (ARS)
- ✧ Ignacio Ciampitti (KSU)
- ✧ Jode Edwards (ARS)
- ✧ Sherry Flint-Garcia (ARS)
- ✧ Christopher Graham (SDSU)
- ✧ Candy Hirsch (UMN)
- ✧ Jim Holland (ARS)
- ✧ David Hooker (Guelph)
- ✧ Shawn Kaeppler (UW)
- ✧ Joe Knoll (ARS)
- ✧ Judith Kolkman (Cornell)
- ✧ Greg Kruger (UNL)
- ✧ Nick Lauter (ARS)
- ✧ Aaron Lorenz (UMN)
- ✧ Liz Lee (Guelph)
- ✧ Natalia de Leon (UW)
- ✧ Jonathan Lynch (PSU)
- ✧ Steve Moose (UIUC)
- ✧ Seth Murray (TAMU)
- ✧ Rebecca Nelson (Cornell)
- ✧ Torbert Rocheford (Purdue)
- ✧ Oscar Rodriguez (UNL)
- ✧ Cinta Romay (Cornell)
- ✧ James Schnable (ISU)
- ✧ Brian Scully (ARS)
- ✧ Margaret Smith (Cornell)
- ✧ Nathan Springer (UMN)
- ✧ Peter Thomison (OSU)
- ✧ Mitch Tuinstra (Purdue)
- ✧ Randy Wisser (UDel)
- ✧ Wenwei Xu (TAMU)

Genomes To Fields Sponsors

Genomes To Fields Collaborators

Project planning, data management, phenotyping initiatives, and Executive Committee members

- Martin Bohn (UIUC)*
- Ed Buckler (ARS)
- Darwin Campbell (ISU)*
- James Clohessy (Cornell)
- Michael Coen (UW)
- Carolyn Lawrence Dill (ISU)*
- Liang Dong (ISU)
- Jode Edwards (ARS)*
- David Ertl (IA Corn)*
- Sherry Flint-Garcia (ARS)*
- Joseph Gage (UW)
- Jack Gardiner (ISU)*
- Byron Good (Guelph)
- Mike Gore (Cornell)
- Patricio Grassini (UNL)
- Jerry Hatfield (ARS)
- Diego Jarquin (UNL)*
- Shawn Kaeppler (UW)
- Liz Lee (Guelph)*
- Natalia de Leon (UW)*†
- Zhizhai Liu (TAMU)
- Aaron Lorenz (UMN)*
- Jonathan Lynch (PSU)
- Nathan Miller (UW)
- Jane Petzoldt (UW)*
- Seth Murray (TAMU)*
- Cinta Romay (Cornell)*
- James Schnable (UNL)
- Pat Schnable (ISU)†
- Nathan Springer (UMN)
- Edgar Spalding (UW)
- Srikant Srinivasan*
- Kelly Thorp (ARS)
- Rod Williamson (IA Corn)
- Randy Wisser (UDel)*
- Jianming Yu (ISU)

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

Genomes to Fields

Madison-Wisconsin, June 12-17, 2016

http://www.icqg5.org

Follow @ICQG2016

Questions?

Genomes to Fields