Can We Make a Dent in Genotype and Environment Interaction in this High Throughput Era?

Jianming Yu

Department of Agronomy, Iowa State University
Quantitative Genetics and Maize Breeding

Position Description

• “to combine maize breeding with cutting-edge genomic technologies to address significant questions in quantitative genetics ... develop and improve contemporary breeding methods”
Overarching Questions

• **What is most efficient way to identify genes underlying complex traits?**
  Germplasm, breeding, genetic design, quantitative genetics, population genetics, experimental design, genomics, bioinformatics, statistics

• **How can we incorporate genomic/phenomic technologies to improve the current breeding process?**
  Genetic gain, breeding cycle, germplasm, trait complexity, logistic support
  Flatteners: Genomic technology, Method (G→P), Tool (DH, transformation, screening, high throughput phenotyping), Breeding methodology, and People
Genotype by Environment Interaction

Xin Li
Poster: *Genotype by Environment Interaction of Sorghum Flowering Time*
*Computer simulation in plant breeding*, Advances in Agronomy 116:219-264

Matt Dzievit
*Topics to Link Genotype to Phenotype at the Genome Level for Improving Genetic Gain*
Sotiris Archontoulis
My research interests include crop phenology, morphology, agronomy, physiology and simulation modeling of crop species. ... Understanding of complex agro-ecosystems concepts involving crops, soils and climate ...

Fernando E. Miguez
I use mathematical and statistical modeling approaches, develop crop databases, and manipulate weather and soil databases to scale up field plot level findings to larger regions.
Interaction

• When the overall effect of two variables is not additive. When the effect of one variable can not be accurately quantified without specifying the level of the other variable
  – “It depends.” – An annoying answer from experts
  – \[ y = A + B + AB + e \]

• Action occurs between two objects, and typically this action has an effect on both objects. A mutual or reciprocal action or influence
  – A plant biologist talking to an engineer
  – Systems biology studies the interactions between the components of biological systems
Genotype

• The unique identifier of an individual
  – Accession, inbred, line, cultivar, hybrid, testcross
  – Seeds representing a genotype can be grown in different places
  – Genotype by environment interaction

• The genetic constitution of an individual
  – Molecular marker fingerprinting data
  – Single nucleotide polymorphism (SNP) data
  – String of 1’s or 0’s
  – Genotype versus phenotype
Environment

• Classifiable environmental conditions
  – Combination of year and location
  – Regions of homogenous conditions in soil, weather, disease pressure, production pattern, etc.

• The physical and biological factors along with their interactions that affect an organism
  – Soil type, temperature, rainfall, wind, hail, heat, flooding, photoperiod, nutrients, insects, diseases, etc.
Genotype by Environment Interaction

- Performance of genotypes differs under different environments
  - Lack of a consistent ranking among genotypes
  - Recommendation and placement of different genotypes requires extensive testing across environment
  - Clustering of different environments and placing of testing sites need to be studied and revisited

- Identifying and deploying favorable genes are not that easy

- Predicting of performance under changing climate conditions is challenging

- Understanding the underlying mechanisms that give rise to G x E is challenging
A “Simple” Example of G x E

Flowering time (DAP)

P898012

Tx430

PR
KS
IA

Flowering time (DAP)

PR
KS
IA
Current Trends and State-of-The-Art

Drowned in next generation sequencing data

HELP!
Gene Regulatory Network (GRN)

Vegetative → Flowering

Murphy et al. PNAS 2011, 108:16469-16474

Crop Modeling

The prediction power and versatility of APSIM is due to its holistic inclusion of all major factors in environment, management, and biology.

Four key panels (dotted boxes) include soil carbon, soil nitrogen, soil water, and plant model.
Processes simulated by the model

**Agricultural Production System Simulator - APSIM**

**Soil**
- Water balance
- Nitrogen cycling
- Carbon mineralization
- Soil temperature
- Root development
- Residue
- Manure
- Erosion
- Biochar (under develop.)

**Crop**
- Phenology
- Biomass production
- Leaf development
- Dry matter partitioning
- Dry matter redistribution
- Crop senescence
- Plant death
- Crop transpiration
- Nitrogen uptake

**Atmosphere**
- ET demand
- N2O emissions
- CO2 emissions
- Climate change

**Management**
- Fertilizers
- Rotations
- Tillage
- Intercropping
Ongoing research direction

Global assessments:
Soil scientists
Climatologists
Computer scientists

Collaborate with scientists to bridge the gap between different research disciples

Link crop model parameters to genetic traits

Soybean rainfed yield (Dec 6, 2013)

Adaptability & productivity

Soybean phenology

Target = \( \frac{(T_a - T_b) \cdot 0.2}{1 - P_{e}n \cdot (P - P_{crit})} \)
Bottleneck I – Genetics & Genomics

\[ P = G + E + G \times E + e \]

Phenotype = Genotype + Environment + Genotype x Environment + Error

• Prediction accuracy across-environment can be increased slightly by considering G x E

• Much of the G x E component cannot be sufficiently be captured by a few abiotic and biotic factors if complex geographic and weather patterns exist

• Projection of untested genotype and untested environment
Bottleneck II – Crop Modeling

- Further capacity improvement to accommodate genetics-related inputs, processes, and outputs
- Lack of capacity to predict genotypes without any performance data
Bottleneck III – Systems Biology

• From detailed pathways and networks under specific tissue type, environmental condition, treatment, time point, genotype to practical applications

• Distance between biological and agricultural reality
Bottleneck

= Lack of interaction
= Lack of joint projects

Crop Modeling

Systems Biology

Genetics & Breeding
Opportunities

- Genomics-to-Field Initiative
- ISU Plant Science Institute and Presidential Initiative
- DOE-USDA Feedstock Genomics Program
- Many others?

![Graph showing predicted vs. observed flowering time with a correlation coefficient of r = 0.93.](image)
Integrated Modeling Approach for Performance Prediction (IMAPP)

Combine the merits of complementary approaches

• Genome-wide prediction to connect all individuals (tested or untested) through genetic principles

• Optimize the prediction through gene regulatory network research

• Crop modeling framework to connect all environments (tested or untested) through the ecophysiological modeling modules and to generated integrated performance prediction

The link is genotype-specific, across-environment, cultivar parameters that can be obtained for all genotypes (individuals) so that we can “predict plant performance under varied conditions”