Advancing Plant Science with Predictive Models and Large-scale Phenotyping

Patrick S Schnable

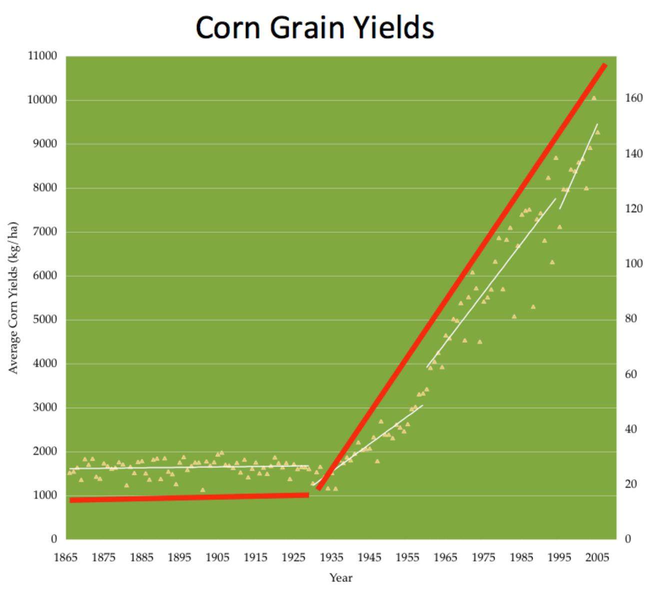
IOWA STATE UNIVERSITY
Plant Sciences Institute

"Using Computing to Sustainably Feed a Growing Population" Panel
AAAS
14 February **2020**

USDA United States Department of Agriculture
National Institute of Food and Agriculture
IOWA CORN® PROMOTION BOARD
NSF
arpa·e

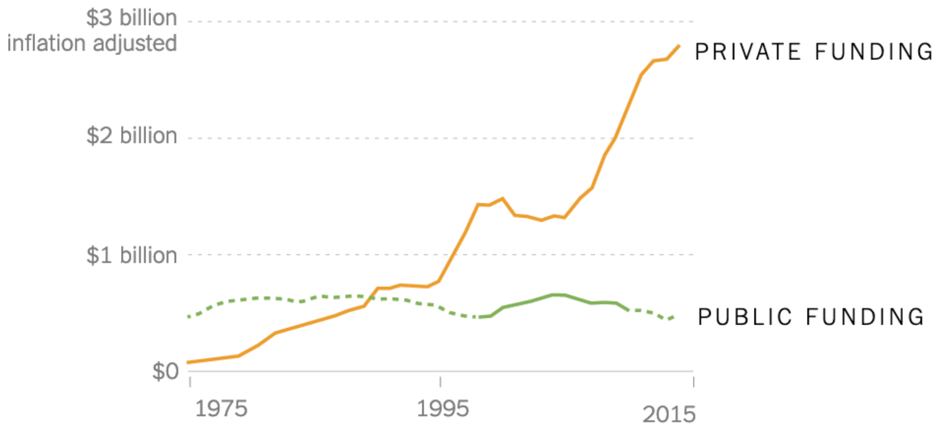
PSS has IP and/or equity interests in Data2Bio, Dryland Genetics, EnGeniousAg & Hi-Fidelity Genetics

Some Agricultural Successes & Challenges



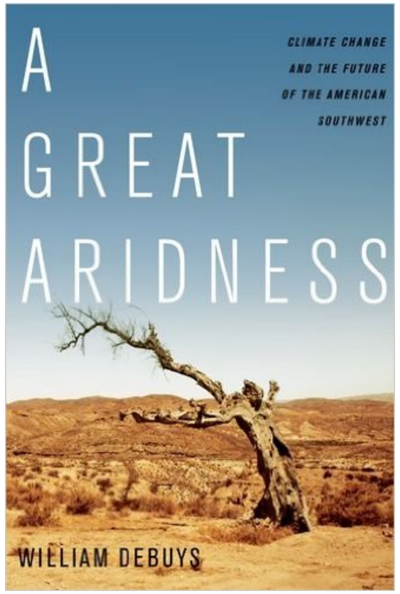
Troyer (2006) Crop Sci. 46:528-543

~$3B/yr ***inflation-adjusted*** investment in PB



Source: USDA and NYT

- Increasing demands for food, feed, fiber and fuel; decreasing amount of arable land
- Agricultural inputs: increasing costs (e.g., nitrogen), reduced availability (e.g., water), and undesirable ecological impacts
- Achieving yield stability is challenging due to climate change-driven weather variability



The Promise of Predictive Models

- Statistical approaches
- Artificial Intelligence
- Process–based Crop Growth Models

Phenotype (P) = Genotype (G) + Environment (E)* + GxE

Given sufficient genotyping, phenotyping and environmental data we can develop models that predict traits/phenotypes of specific genotypes, under defined environments and crop management systems

*Includes crop management

Predictive Models Will:

- Improve accuracy of selection in plant breeding programs, thereby increasing rate of genetic gain per year
- Enhance our ability to efficiently breed crops to withstand the increased weather variability associated with global climate change, and more generally breed crops for future environments
- Improve our ability to provide farmers with evidence-based recommendations for the appropriate varieties to plant in a given field, under a particular management practice in a given year, leading to increased yields and enhanced yield stability (thereby potentially reducing the cost of crop insurance)
- Enable daily national and global yield predictions, thereby facilitating early responses to food emergencies and avoiding market failures

Turbocharged "Shovelomics"

Lisa Coffey

Zaki Jubery

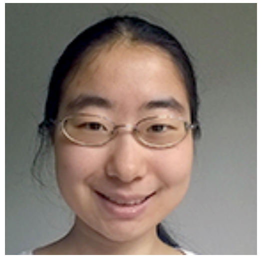
Baskar Ganapathysubramanian

Requirements for 60 Core Root Systems

	Traditional Shovelomics	Turbo
Excavation	1 h	1 h
Prep & Cleaning	3h 5 min	32 min
Resources	120 gal of water	1 gal diesel fuel

Zihao Zheng
(郑子豪)

Stefan Hey

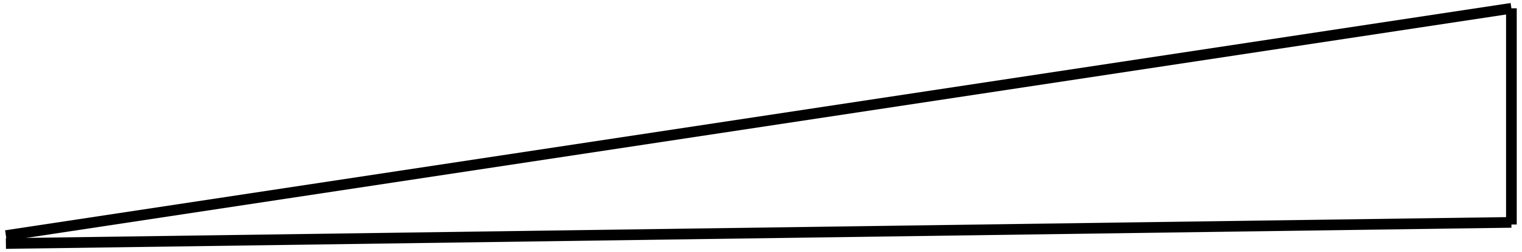


Huyu Liu
(刘祜宇)

Substantial Variation in Root System Architecture (N=380 inbreds)

Root Angle

Root Volume

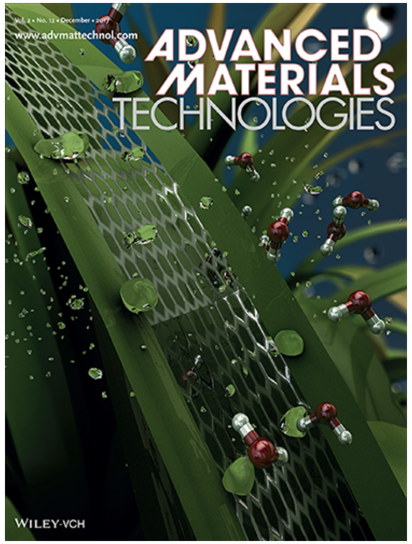


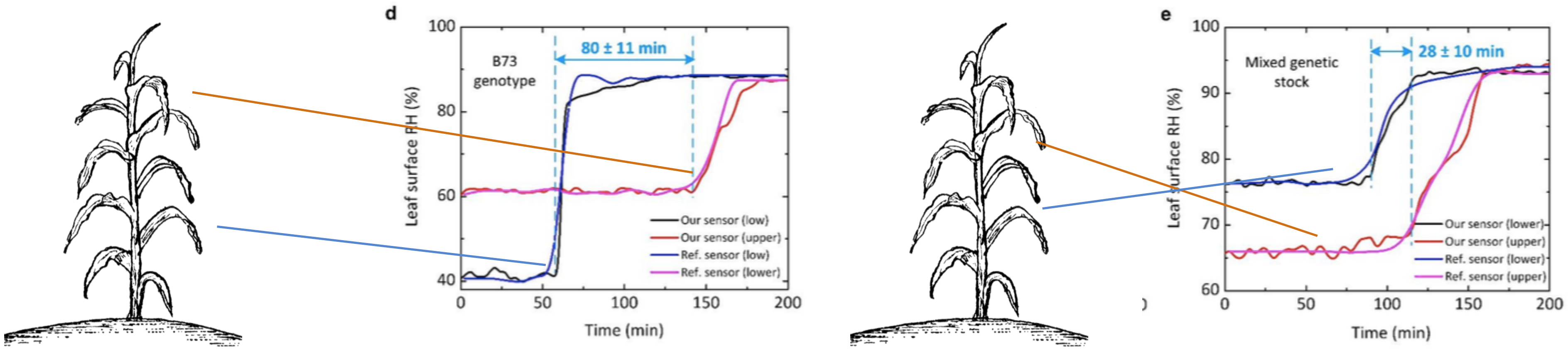
16% (N=22) of 139 maize RSA-associated genes (or their homologs) identified via GWAS are known to affect RSA in maize or other species

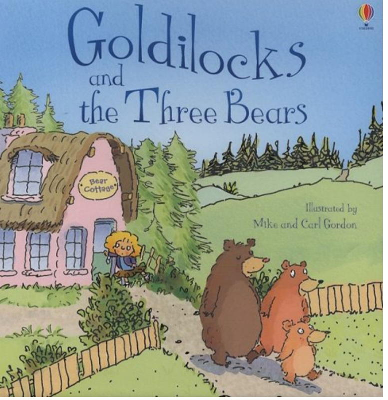
Zheng et al. (2020) Plant Physiology

Liang Dong

Plant "Tattoos": Detecting Relative Humidity at Leaf Surfaces

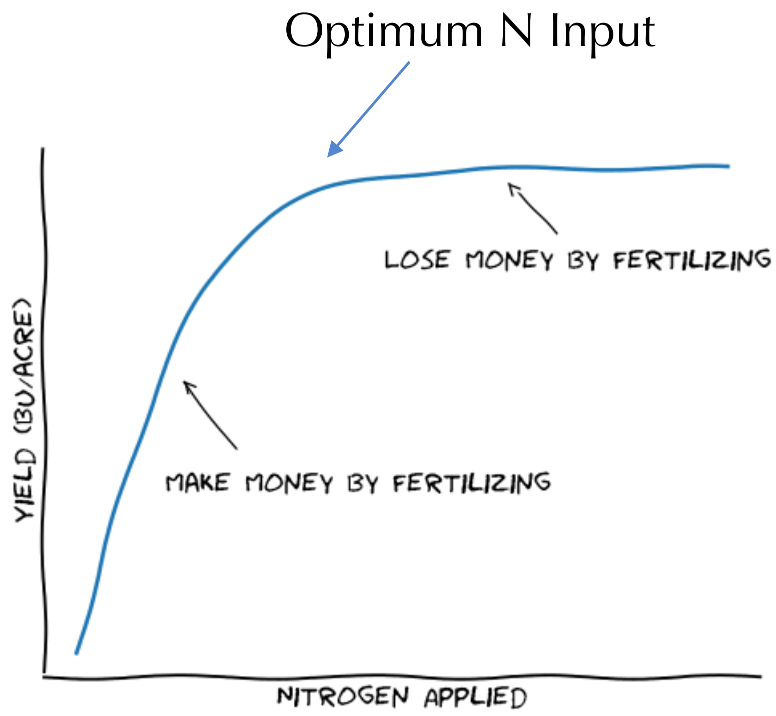






N Application is a Goldilocks Problem

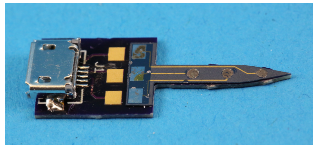
- *Under* application -> yield losses
- *Over* application -> wasted input costs & environmental impact

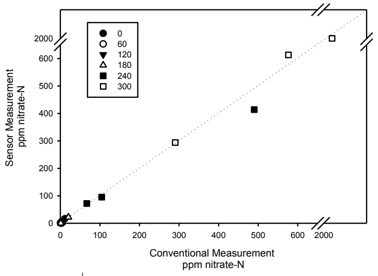


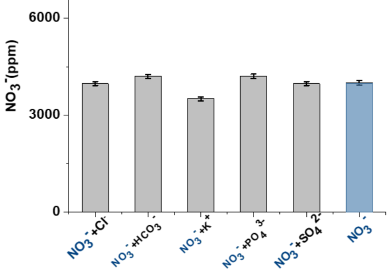
- Nitrogen is 2nd most expensive input for rain fed corn (after seed)
- Selecting appropriate N application rates is complicated by substantial year-to-year variation in N production from soil organic matter and field losses of N
- *Predicting the optimal level of Nitrogen is currently difficult to impossible*
- *35% of fields exhibit NO response to N*
- *$1.67B of wasted N fertilizer per year*

Nitrate-specific Instant Read Sensor

Liang Dong
(ISU)

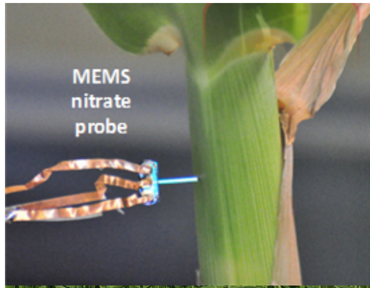






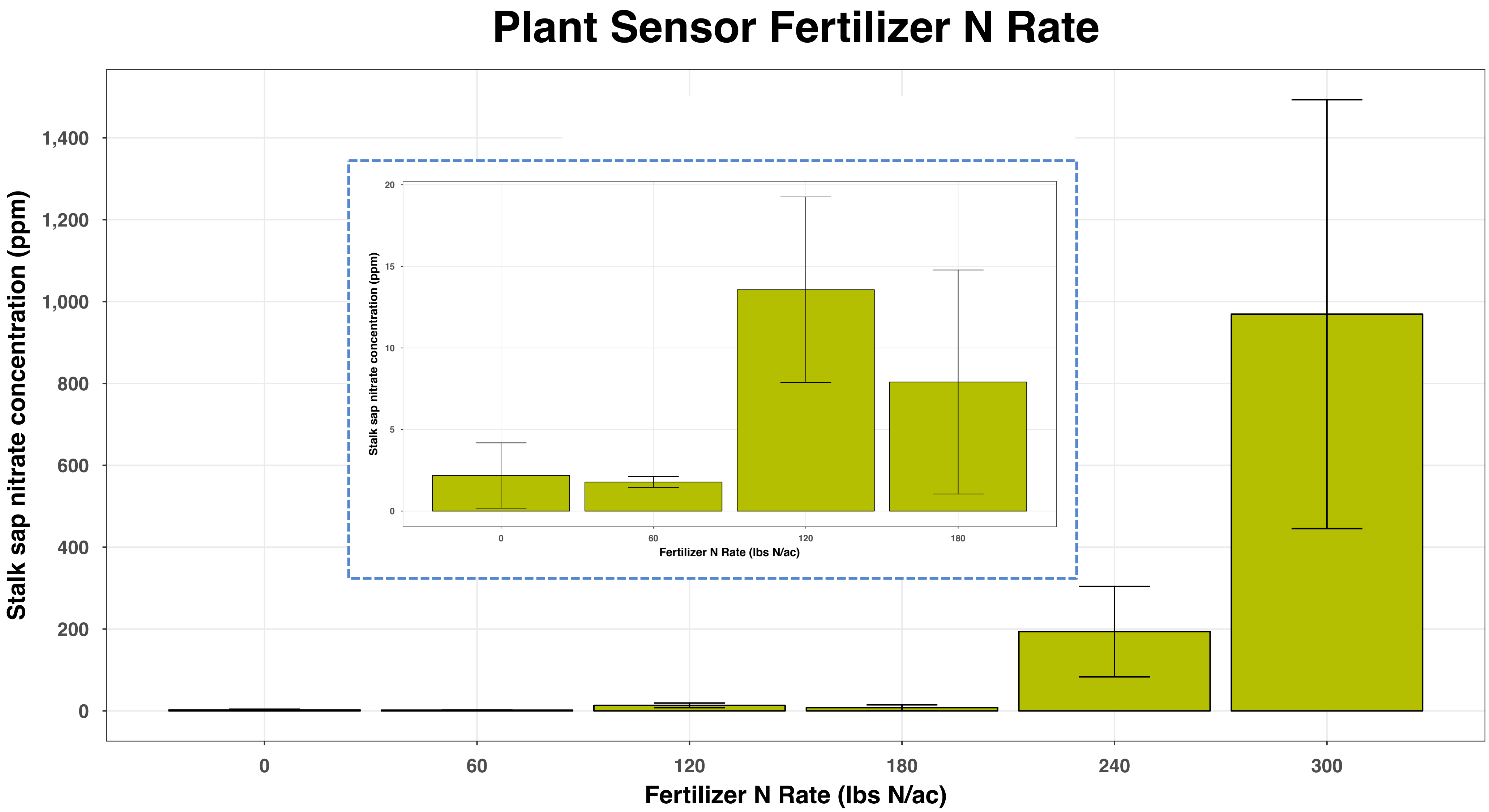
James
Schnable
(UNL)

Mike
Castellano
(ISU)



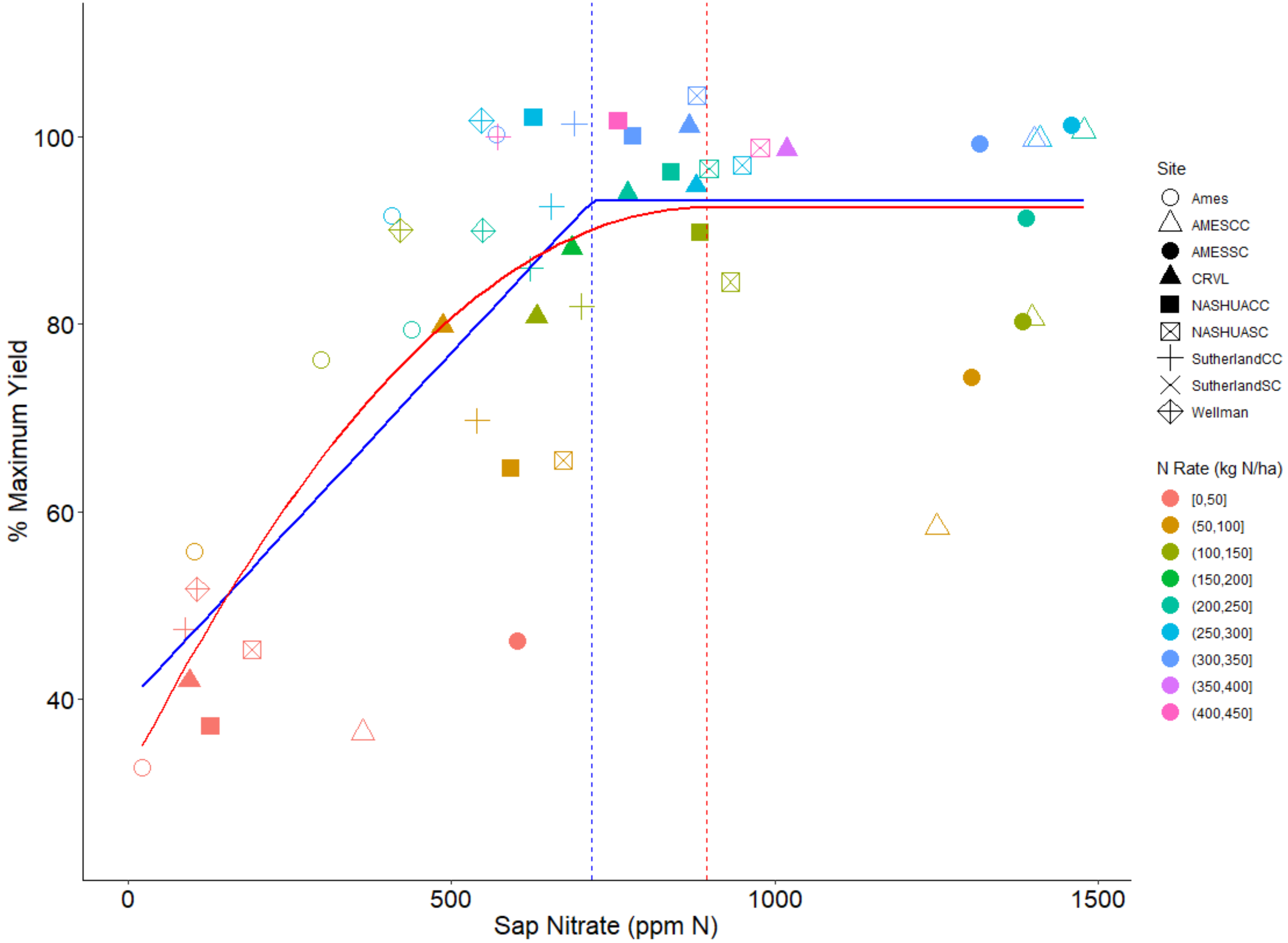
EnGeniousAg
NUTRIENT SENSOR

Potentially Actionable Data



Mike Castellano (ISU)

% Max Yield vs Sap [N]



the Genomes to Fields initiative

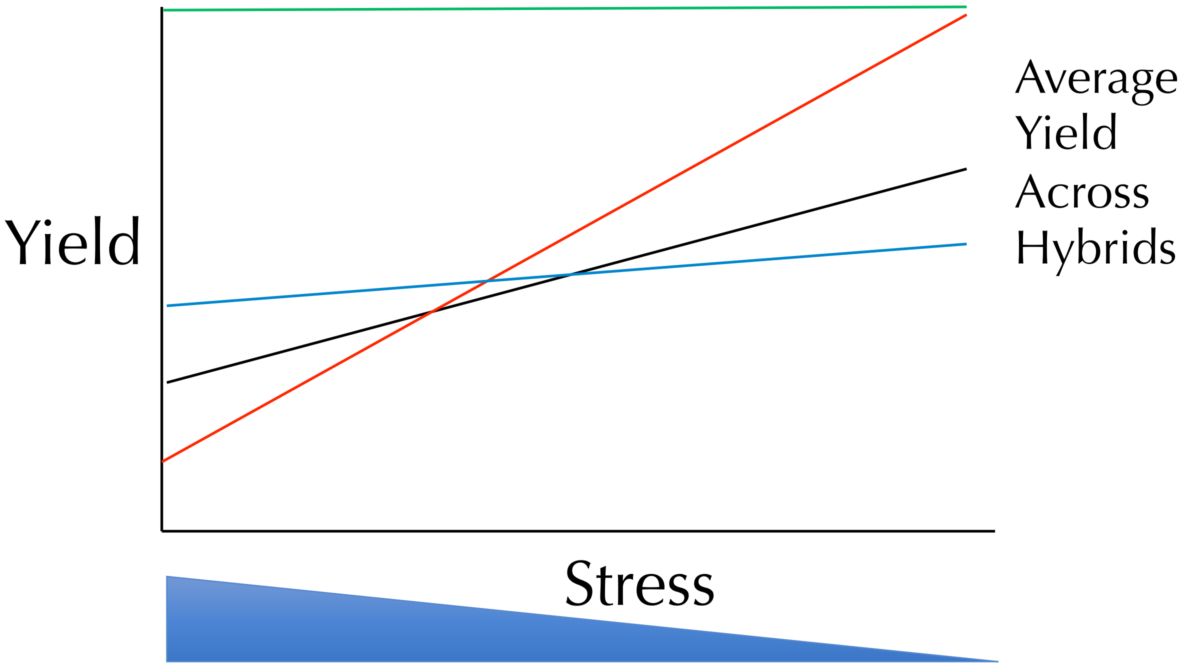
The Genomes to Fields Initiative

- Multi-state partnership led by ISU, Univ of WI and Iowa Corn to develop a “living laboratory” focused on understanding the genetics of maize performance across environments
- A large, multi-state, multi-year dataset of maize phenotype data has been made publicly available and is enabling researchers to identify relationships among genes, the environment and crop performance

Aaron Kusmec

Distinct Genetic Architectures for Phenotypes and Plasticity (GxE)

- **Allelic Sensitivity Model** (Via and Lande, 1985)
  - *Same genes* responsible for performance & plasticity
- **Structural Gene Model** (Scheider and Lyman, 1989)
  - Performance (intercept) & plasticity (slope) controlled by *different genes*



GWAS for ~25 traits on ~5,000 NAM genotypes grown in ~20 environments

nature plants
ARTICLES
DOI: 10.1038/s41477-017-0007-7

Distinct genetic architectures for phenotype means and plasticities in *Zea mays*

Aaron Kusmec[1], Srikant Srinivasan[2,4], Dan Nettleton[2,3] and Patrick S. Schnable[1,2]*

Some of our research goals:

- Use data-driven approaches to identify specific combinations of developmental windows and environmental parameters that are most impactful to grain yield
- Identify genes which confer resistance to these developmentally sensitive environmental stresses
- Use future predicted weather patterns to conduct simulated genomic selection/breeding experiments to assess the feasibility of breeding for future environments that do not currently exist on Earth, but unfortunately will in the not too distant future

What's Required to Understand GxE, Develop Predictive Models and Innovate our Way Out?

- A highly interactive community of plant scientists, engineers and data scientists
- New phenotyping technologies
- Data (phenotypes, genotypes and environment/management practices), ideally from coordinated, multi-location, multi-year projects
- New analysis methods
- Significant additional R&D investments