Published June 27, 2019

SCIENTIFIC PERSPECTIVES

Idea Factory: the Maize Genomes to Fields Initiative

Carolyn J. Lawrence-Dill,* Patrick S. Schnable, and Nathan M. Springer

ABSTRACT

Increases in agricultural productivity are necessary to meet anticipated demands, and these increases must occur in an environmentally sustainable fashion. However, it has become measurably harder to generate ideas and new approaches that result in real gains. The maize Genomes to Fields (G2F) Initiative seeks to bring together researchers across diverse disciplines to create situations that generate new ideas for research, development, and education. Contact from those interested to learn more about G2F is invited, and inquiries on how to initiate and support research and training at scale are welcome.

C.J. Lawrence-Dill, Dep. of Genetics, Development and Cell Biology and Dep. of Agronomy, Iowa State Univ., Ames, IA 50011; P.S. Schnable, Dep. of Agronomy, Iowa State Univ., Ames, IA 50011; N.M. Springer, Dep. of Plant and Microbial Biology, Univ. of Minnesota, St. Paul, MN, 55108. Received 1 Feb. 2019. Accepted 1 Apr. 2019. *Corresponding author (triffid@iastate.edu). Assigned to Associate Editor Manjit Kang.

Abbreviations: FACT, Food and Agriculture Cyberinformatics and Tools; G2F, Genomes to Fields; IMPACTS, Integrated Training PrograM in Plant And CompuTational Sciences; NIFA, National Institute of Food and Agriculture; NRT, National Science Foundation Research Traineeship.

It is not news to anyone that increases in agricultural productivity and efficiency are required to meet anticipated needs for food, feed, fiber, and fuel in the coming decades. However, efforts to improve crops are affected by two overarching challenges that are not nearly so widely appreciated:

1. Well-documented, large-scale, publicly accessible datasets measuring plant performance in diverse environments are critically lacking. The absence of accessible, large-scale, multiyear, multienvironment datasets limits our ability to test ideas for how genotypes and the environment interact to produce improved traits in one crop. The availability of such datasets also offers the possibility of determining whether single-crop findings are generalizable across diverse crop species. Put another way, for any given hypothesis, researchers begin by designing and executing their own field experiments and collect data in novel ways, which results not only in years of lag time, but in datasets that are difficult to aggregate and integrate across experiments, research groups, and species.

Published in Crop Sci. 59:1406–1410 (2019).
doi: 10.2135/cropsci2019.02.0071

2. Ideas for how to achieve increased crop yields have become measurably harder to come by. Many of the widely accepted, obvious ways to improve traits have already been identified and leveraged. What now? As we pour more resources into research, the return on investment decreases, with agricultural improvement serving as a case study for this phenomenon of diminishing returns (Bloom et al., 2017). Clearly, we must find ways to think differently if we are to continue to improve traits that allow us to grow high-yielding crops on less land with fewer inputs.

To address these challenges, a coalition that included researchers, engineers, economists, and others focused on measuring the performance of *Zea mays* ssp. *mays* (maize) across diverse environments in North America was assembled.

WHY MAIZE?

Maize is both a premier, long-standing research model for genetics and genomics and a primary staple for agricultural production worldwide (reviewed in Strable and Scanlon, 2009, and Nannas and Dawe, 2015). At the same time, per-hectare yields and global production of maize exceed those of both rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.) (Table 1).

Moreover, the maize research community is diverse, including geneticists, molecular biologists, physiologists, breeders, and others. In addition to the diversity of expertise, the maize research community is widely recognized as collegial, welcoming to outsiders, open to novel ideas, and focused on excellence in education and training. These qualities derive from nearly a century of concerted efforts to nurture and grow the community based on shared values that support a healthy environment for open scientific inquiry (reviewed in Kass et al., 2005).

GENOMES TO FIELDS: BEGINNINGS AND EARLY ACCOMPLISHMENTS

In 2014, Natalia De Leon, Patrick Schnable, David Ertl, and colleagues assembled a group of maize researchers and established the Genomes to Fields (G2F) Initiative. Together, we began collecting key datasets necessary to investigate how genotypes, variable environmental conditions, and field management practices affect traits including, but not limited to, phenotypic plasticity and yield stability (reviewed in Gage et al., 2017). The initiative's biggest investments have supported field trials and shared experiments conducted at >20 locations distributed across North America (reviewed in AlKhalifah et al., 2018; see Fig. 1). To enable the discovery of novel associations among traits, genes, and environmental variation, data are collected across these locations annually including phenotypes, genome sequences, and environmental conditions including weather, soils, and crop management practices. To enable scientists beyond G2F members to make use of these data for their own discoveries, data are released annually.

The group of researchers directly involved in G2F have diverse disciplinary backgrounds including engineers interested in sensor development and machine learning applications, scientists focused on understanding and improving efficiencies for maize production, and economists who determine how production, planting distribution, and access to data affect market conditions. This engagement of a broad diversity of experts is a necessary condition for novel idea generation. This strength is especially apparent when G2F researchers get together to discuss findings, consider opportunities, and plan new experiments.

PLANNING OUR NEXT STEPS

To assess the G2F Initiative and identify focus areas for the coming years, Patrick S. Schnable and colleagues conducted a 2-d meeting supported by the USDA National Institute of Food and Agriculture (NIFA) held at Iowa State University in early 2018. Seminars from scientists, economists, and engineers describing their collaborative activities, findings, and ideas provided the basis for subsequent discussions and brainstorming sessions, which were led by Carolyn J. Lawrence-Dill. Here, we share some outcomes and recommendations from that meeting because we believe they have relevance beyond the maize community—particularly for those planning to organize and deploy large-scale, multiyear, collaborative experiments focused on other crops.

1. Continue shared community experiments. Thus far, G2F has primarily been focused on using consistent methods to measure and document phenotypes in field environments. Expansion to support specific research goals by adding controlled environments, novel sensors, and site-specific areas of focus while continuing to release datasets annually will enable the G2F community to continue to grow in both breadth and depth of expertise.

Table 1. Worldwide maize, rice, and wheat production, yield, and area (FAOSTAT, 2018; search terms: world total production quantity for maize, rice [paddy], and wheat in terms of metric tons produced).

Crop	Production	% of maize	Yield	% of maize	Area	% of maize
	t	%	kg ha^{-1}	%	ha	%
Maize	1060,107,470	100	56,401	100	187,959,116	100
Rice (paddy)	740,961,445	70	46,366	82	159,807,722	85
Wheat	749,460,077	71	34,050	60	220,107,551	117

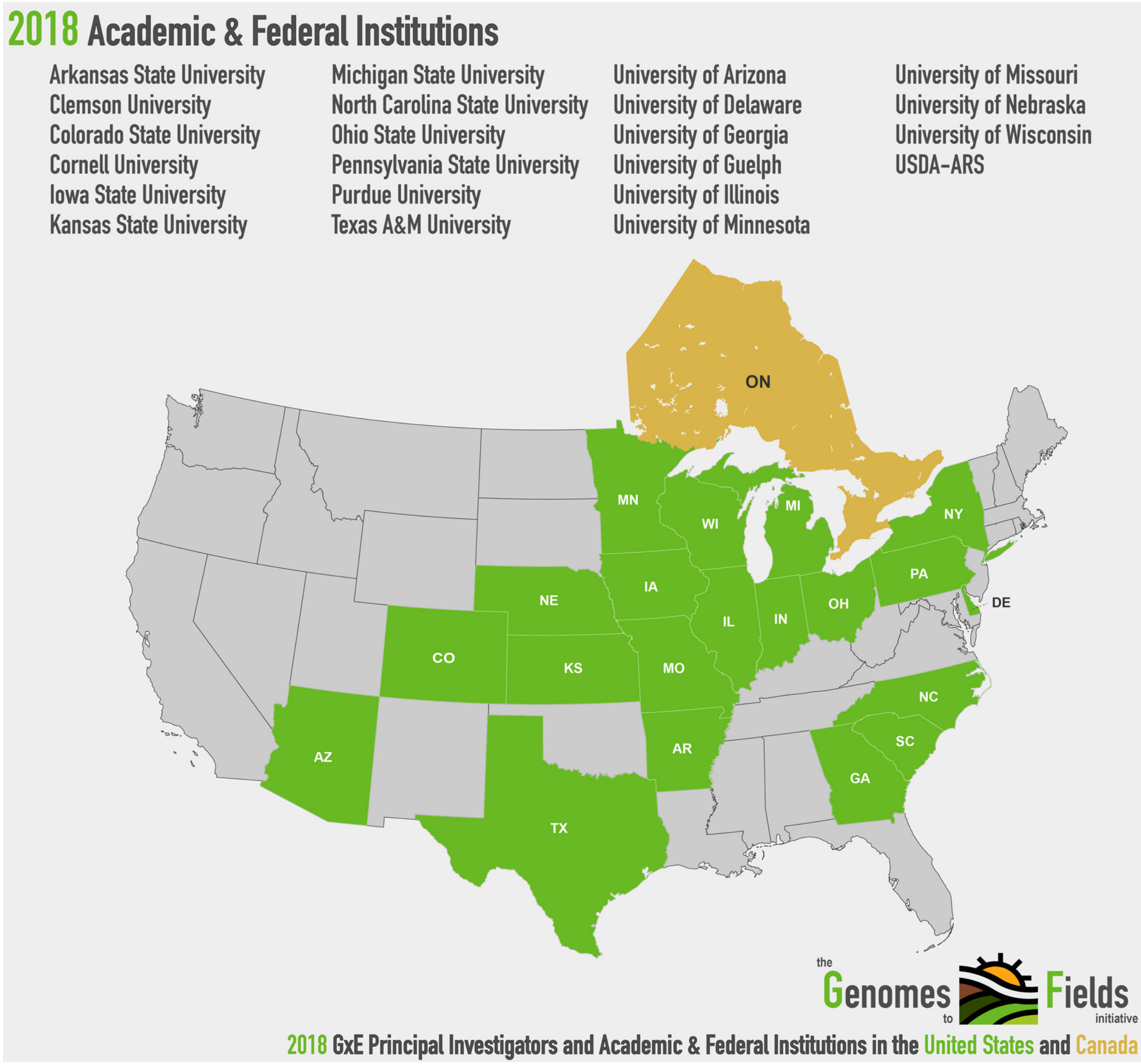

Fig. 1. Genomes to Fields (G2F) locations in 2018. Researchers managing G2F field sites come from 22 academic and federal institutions. The G2F fields are located in 19 states (green) in the United States and in southeastern corner of the province of Ontario, Canada (yellow).

2. Develop and test novel sensors to chart environmental and physiological conditions. Developing and deploying a standard set of sensors and sensor-equipped unmanned ground and aerial vehicles (UGVs and UAVs) to collect diverse data from single plants to fields across all sites at multiple time points will expand our capacity to extrapolate from single-plant measurements to field data and beyond when coupled with data already being collected and made available from satellite imagery.
3. Create data management, sharing, and analytics platforms. Streamlined data communication platforms and mechanisms are needed. Community-oriented data deposition, storage, access, and analysis portals are sorely needed, with connections to relevant existing data resources. Mechanisms for image-based data deposition, storage, access, and analysis are critically needed.
4. Engage additional scientific disciplines in research and training. Targeted efforts to engage scientists beyond the current group of plant scientists, breeders, and engineers currently involved in G2F are needed. Involving physiologists, biochemists, crop modelers, soil scientists, ecologists, climatologists, engineers, computer scientists, data scientists, bioinformaticians, mathematicians, statisticians, economists, and communication specialists, including policymakers, will broaden the scope of impact of G2F and support the development of new ideas on how best to achieve crop improvement.
5. Designate and create high-intensity phenotyping sites. New technologies including sensors and data collection protocols and analysis techniques must be developed and tested before widespread deployment. Local expertise at G2F sites will be matched to specific needs for technologies to be deployed across sites. From there, a number of high-intensity phenotyping sites will be designated to assess and optimize ease of use, accuracy, and value of novel technologies prior to widespread adoption at G2F sites.

Detailed information on each recommendation is available via Lawrence-Dill et al. (2018b).

EDUCATION IS KEY

Together, these five recommendations make it clear that novel training programs that aim to develop skills and expertise in idea generation, data management and analysis, and collaborative capacity are sorely needed. Examples of programs doing exactly that are beginning to appear, with significant funding provided by the National Science Foundation Research Traineeship (NRT) Program. At Iowa State University, the Predictive Plant Phenomics (P3) NRT combines expertise in plant sciences, engineering, and data sciences with skills in collaboration and entrepreneurship to support the development of a new type of researcher equipped to work across disciplines and institution types (https://www.predictivephenomicsinplants.iastate.edu; Lawrence-Dill et al., 2018a). This program is in its fourth year. At Michigan State University, the Integrated Training PrograM in Plant And CompuTational Sciences (IMPACTS) NRT aims to train students in computational plant sciences focused on tackling multiscalar research problems, from molecular to ecosystem levels (https://natsci.msu.edu/news/msu-nets-3m-nsf-grant-for-stem-graduate-education). The IMPACTS program is in its first year. These and other novel training concepts promise to expand how the next generation of plant scientists approaches hard problems to generate and test new ideas in support of increased crop productivity with minimal environmental impact and to represent excellent programs for educational partnerships with the G2F Initiative.

REQUEST FOR PARTICIPATION

We welcome inquiries on how to initiate and support research and training on broad scales and invite contact from anyone interested in partnering with G2F in novel ways to strengthen the initiative and to broaden participation of diverse scientists and engineers. This request for contact and broad engagement is especially aimed at economists and social scientists who may be interested to learn what it takes to create, conduct, and support large-scale, collaborative research projects. We also seek to expand application of principles and methods for data collection, aggregation, and access to other crops, thus enabling generalized hypothesis testing beyond single-crop models.

New ideas and approaches are needed to meet demands for agriculture into the future. To generate these new ideas, it is critical to expand our community and ways of thinking. We believe that readers of *Crop Science* are very likely to know and interact with researchers across diverse disciplines who have an interest in, and an appreciation for, needed agricultural innovation. We call on you to spread this message and request that you and your colleagues visit the G2F website at https://www.genomes2fields.org/. From there, you can reach out to members of the G2F Executive Committee to find out how to join the initiative or to learn how to apply methods, tools, and community-building perspectives developed for G2F to your favorite crop.

Conflict of Interest

The authors declare that there is no conflict of interest.

Acknowledgments

Funding for this project was provided by the USDA-NIFA, Food and Agriculture Cyberinformatics and Tools (FACT) program via Grant no. 2017-07799 to P.S. Schnable with additional support provided by Iowa State University's Plant Sciences Institute. Recommendations reviewed here derive from the efforts of many researchers involved in the G2F project as well as those who contributed to the 2018 NIFA FACT Workshop whitepaper (Lawrence-Dill et al., 2018b).

NIFA FACT Speakers and Participants

A.J. Ackerman, Bill Adamowski, Naser Alkhalifah, Matt Baseggio, Steven Anderson, Phillip Benfey, Jessica Bubert, Darwin Campbell, Mike Castellano, Parag Chitnis, Jinlyung Choi, Natalia De Leon, Liang Dong, Jode Edwards, David Ertl, Celeste Falcon, Brice Floyd, Baskar Ganapathysubramanian, Stefan Hey, Matt King, Matt Johnson, Shawn Kaeppler, Aaron Kusmec, Nick Lauter, Carolyn J. Lawrence-Dill, Bridget McFarland, John McKay, Francisco Munoz-Arriola, James McNellie, Daniel Nettleton, Danilo Moreta, Seth Murray, Duke Pauli, Nathalia Penna Cruzato, Guru Rao, Colby Ratcliff, Ray Riley, James Schnable, Patrick Schnable, Kevin Silverstein, Amritpal Singh, Edgar P. Spalding, Nathan Springer, Lie Tang, Addie Thompson, Ruth Swanson-Wagner, Lizhi Wang, Mark White, Adam Vanous, Jason Wallace, Justin Walley, and Jianming Yu.

We are also grateful to MaryAnn Grapp for coordinating logistics for the 2018 G2F NIFA FACT workshop and to Cheng-Ting Yeh for creating the 2018 G2F locations map.

References

AlKhalifah, N., D.A. Campbell, C.M. Falcon, J.M. Gardiner, N.D. Miller, M.C. Romay, et al. 2018. Maize genomes to fields: 2014 and 2015 field season genotype, phenotype, environment, and inbred ear image datasets. BMC Res. Notes. 11:452. doi:10.1186/s13104-018-3508-1

Bloom, N., C.I. Jones, J.V. Van Reenen, and M. Webb. 2017. Are ideas getting harder to find? Working Paper No. 23782. Natl. Bureau Econ. Res., Cambridge, MA. doi:10.3386/w23782

FAOSTAT. 2018. FAOSTAT database. FAO, Rome. http://www.fao.org/faostat/en/#data/QC (accessed 23 Sept. 2018).

Gage, J.L., D. Jarquin, C. Romay, A. Lorenz, E.S. Buckler, S. Kaeppler, et al. 2017. The effect of artificial selection on phenotypic plasticity in maize. Nat Commun. 8:1348. doi:10.1038/s41467-017-01450-2.

Kass, L.B., C. Bonneuil, and E.H. Coe. 2005. Cornfests, cornfabs and cooperation: The origins and beginnings of the maize genetics cooperation news letter. Genetics 169:1787–1797.

Lawrence-Dill, C.J., T.J. Heindel, P.S. Schnable, S.J. Strong, J. Wittrock, M.E. Losch, and J.A. Dickerson. 2018a. Transdisciplinary graduate training in predictive plant phenomics. Agronomy 8:73. doi:10.3390/agronomy8050073

Lawrence-Dill, C.J., P.S. Schnable, N.M. Springer, N. De Leon, J.W. Edwards, D. Ertl, et al. 2018b. G2F NIFA FACT Workshop: High throughput, field-based phenotyping technologies for the Genomes to Fields (G2F) Initiative. Figshare. doi:10.6084/m9.figshare.7074860.v1

Nannas, N.J., and R.K. Dawe. 2015. Genetic and genomic toolbox of *Zea mays*. Genetics 199:655–669. doi:10.1534/genetics.114.165183

Strable, J., and M.J. Scanlon. 2009. Maize (*Zea mays*): A model organism for basic and applied research in plant biology. Cold Spring Harbor Protoc. 2009(10):pdb.emo132. doi:10.1101/pdb.emo132