Seeds of Discovery: Characterizing and Utilizing Maize Genetic Resources for Germplasm Diversification

Kevin Pixley*, Sarah Hearne, Martha Willcox, Charles Chen, Juan Burgueño, Huihui Li, Jose Crossa, Gilberto Salinas, Terence Molnar and Peter Wenzl

International Maize and Wheat Improvement Center (CIMMYT), Mexico
*Corresponding author; Email: k.pixley@cgiar.org

Introduction

"To eat or not to eat" is a daily question for nearly one billion people today. The near-certain expectation of population growth and challenges to agricultural production in the coming decades, suggest that hunger will increase unless we make difficult, bold decisions now. Difficult decisions require the commitment of valued and highly sought after resources to support research into the effective use of crop biodiversity. Bold decisions involve the use of genotype information from the entire maize and wheat seed collection in the CIMMYT gene bank and at other gene banks, in order to discover and validate strategies and tools to make better use of them and level the playing field to ensure equity in the access to and use of genomic diversity to feed this and future generations.

Within the next few decades, farmers will need solutions to increasingly frequent crop wilting due to declining water tables and rising maximum temperatures, or premature defoliation from attacks by diseases or pests arriving in their fields for the first time, or declining health of their livestock or their own families due to the diminished nutritional value of the grain they harvest. We know, beyond any reasonable doubt, that these challenges are coming and the time to act is now.

The *Seeds of Discovery* (SeeD) initiative may be mankind's best chance to learn how to develop solutions to the challenges that climate change and widely declining natural resource bases will pose to producing the wheat and maize needed by the world's growing population within this or the next generation. Beyond maize and wheat, "SeeD is a 'foundational initiative' and 'blueprint' for similar, broader efforts encompassing other crops" according to the Global Crop Diversity Trust (http://www.croptrust.org).

SeeD will enable breeders worldwide to identify and use genetic variation, which currently is shelf-ready in gene banks, for use in developing high-yielding, climate-resilient maize and wheat varieties with livelihood-enhancing quality traits. SeeD will achieve this by making available a comprehensive maize and wheat 'genetic resources-utilization platform' consisting of detailed, genetic maps together with descriptive data for high-priority traits for tens of thousands of accessions (seed collections) in gene banks and bioinformatic tools that empower researchers to customize searches and identify the biodiversity that will enable them to develop varieties for changing farmer and consumer needs.

Innovative science is needed to use genetic resources

The global demand for agricultural production will increase by 60 to 70 percent by 2050. The impacts of changing climate will seriously challenge reaching these targets, and most climate models indicate that the production potential of many developing countries, including food-insecure countries in Sub-Saharan Africa and South Asia, may be disproportionally and negatively affected. Achieving food security in coming decades will require anticipating and quickly adapting to climate change, plus a determined focus on resource-poor countries, many of which have rapidly growing populations.

Increases in yields will contribute most of the projected growth in crop production in developing countries. Whereas breeding for higher yield typically progresses by gradually pyramiding beneficial alleles with incremental effects, there are examples demonstrating that substantial progress can be made by introgressing alleles from genetically distant materials (Gur and Zamir, 2004; Imai et al., 2014). Currently, under-utilized genetic resources may offer one of the greatest opportunities for accelerating yield gains and overcoming emerging productivity bottlenecks resulting from a changing climate (McCouch et al., 2012). CIMMYT and sister institutions hold in-trust, the world's most diverse collections of maize and wheat. These collections are priceless; yet, the huge complexity of finding and strategically using these assets in these gene banks means that most of the would-be building blocks for new varieties remain on shelves, at a time when challenges to crop production are unprecedented.

SeeD applies recent breakthroughs in DNA-sequencing technologies, now 10,000 to 100,000 times more powerful than those used to map the first human genome, to help maize and wheat breeders accelerate genetic gains through more effective use of genetic resources. SeeD will characterize and build a "catalogue" or "atlas" describing the native diversity

of maize and wheat, at a cost similar to that of developing and commercializing one transgenic cultivar!

CIMMYT's mission is to sustainably increase the productivity of maize and wheat systems to ensure global food security and reduce poverty. Its ample experience in field-based germplasm evaluation and diverse network of partners uniquely position CIMMYT to lead this innovative project focused on benefiting resource-poor farmers. SeeD was launched in 2011 by CIMMYT and the Government of Mexico and has made important progress to date.

What has been learned?
A very efficient method to genotype maize populations was developed
Most maize gene bank accessions are landrace populations of genetically heterogeneous individuals. SeeD, partnering with Diversity Arrays Technology (DArT), tailored a *genotyping-by-sequencing* (GBS) method to evaluate 30-plant DNA bulk samples, which enables scientists to estimate allele frequencies within populations (accessions) and survey genetic distances among them. This method will reduce the cost of genotyping CIMMYT's entire maize collection of 27,000 accessions by a factor of at least 10.

Zeroing-in on sources of stress tolerance
Figure 1 illustrates the genetic relationships between CIMMYT's elite maize lines (CMLs) and accessions collected from drought- and/or heat-prone areas. The clear separation of elite lines from the stress-tolerant accessions suggests that there may be stress-tolerant genes in the accessions, that have not yet been captured in elite lines.

Validated a novel strategy to identify genetic variation in maize
SeeD embarked on a comprehensive genome-wide association study (GWAS) to establish the relationship between genetic "fingerprints" and important agronomic traits. The objective was to search for useful genes in 4,000+ gene bank accessions representing a substantial part of the natural diversity of maize. Test-cross hybrids of these accessions have been characterized for tolerance to a wide range of stresses and quality traits across 38 trials conducted in collaboration with the *Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias* (INIFAP) and other partners in Mexico. The parents of this panel were genotyped using two complementary GBS platforms from Cornell University and DArT.

Preliminary analysis of flowering time confirmed the pertinent genes previously published and new genome regions of relevance to tropical maize germplasm. This flowering time analysis provided proof-of-

concept for the experimental approach and demonstrated the utility of the SeeD GWAS panel.
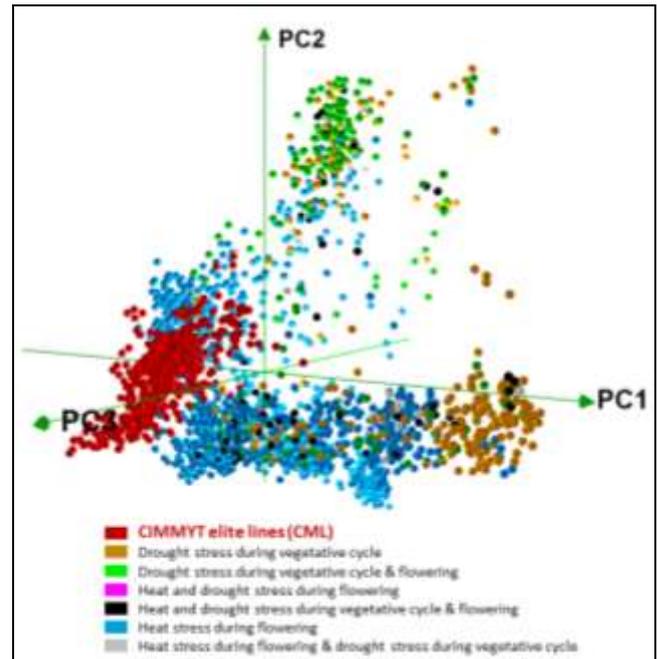

Figure 1. Genetic distance between and among CIMMYT maize genebank accessions from heat- or drought-prone environments and CIMMYT's elite lines (CIMMYT Maize Lines, CMLs)

Maize germplasm bank accessions found with strong resistance to tar spot complex (TSC)
Since 2008, farmers across the Mexican states of Oaxaca, Chiapas and Guerrero, as well as northern Guatemala have experienced yield losses of up to 90 percent due to movement of an aggressive strain of the TSC disease into previously unaffected areas. Testcross hybrids of two landraces from Oaxaca and Guatemala, evaluated for TSC resistance as part of the GWAS trials, were found to have strong resistance to the disease (Figure 2). Farmers in the village adjacent to the trial site were eager to help and offered land and labor to further any breeding process. These accessions have been confirmed to be resistant in subsequent trials and information and seed have been distributed to breeders in the endemic area. Crosses have been made with CIMMYT lines and populations will be tested during summer 2014 which have the potential to be distributed to farmers within the next few years.

Additionally, testcross progeny from GWAS trials with tolerance to low-nitrogen fertility and stalk rot resistance have been identified and families derived from these accessions are currently in trials to form resistant populations for farmer use as well as for lines used in breeding programs.

Genome regions controlling more than a dozen high-priority breeding traits have been identified with high statistical confidence

Examples include analyses of grain-protein content, a nutritionally and industrially valuable trait and of Grey leaf spot resistance, a disease of global importance. Follow-up *in silico* and experimental validations are required to confirm marker-trait associations and to enable the use of markers and new maize germplasm sources for (pre)breeding efforts.
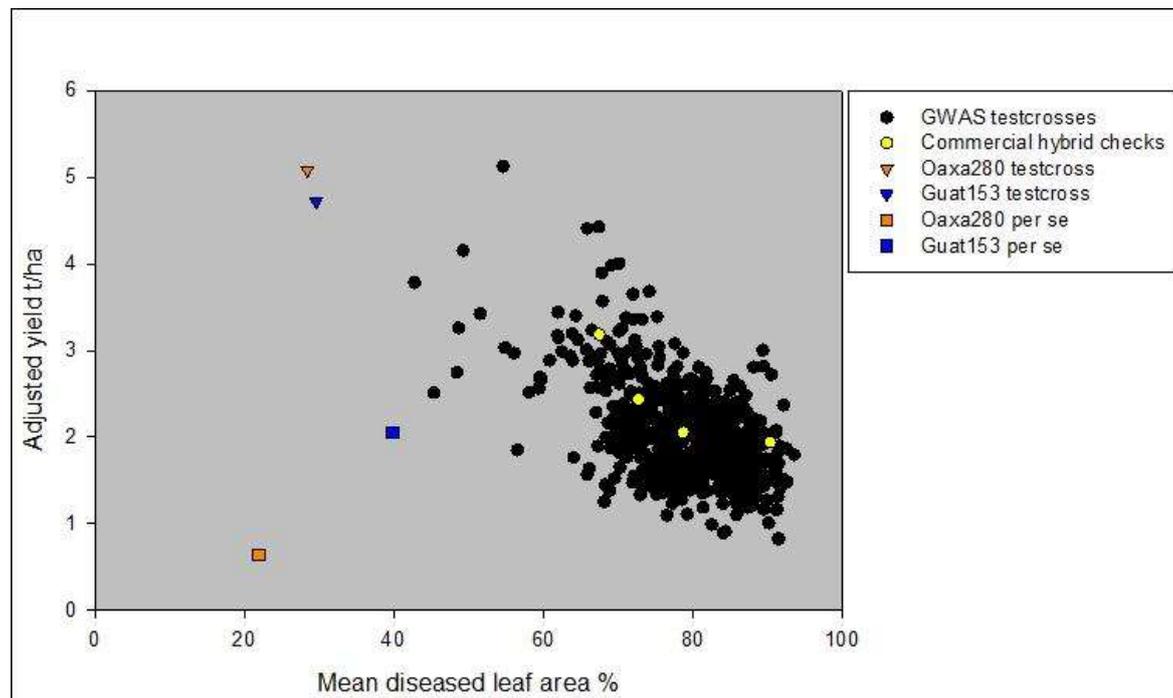

Figure 2. Grain yield and mean percent diseased leaf area of (three-way) topcross hybrids of landraces and four commercial hybrid checks evaluated under severe tar spot disease pressure. The two square points indicate the performance of the two resistant landraces per se, while the two triangles indicate the topcross hybrids of these two landraces.

Building tools and capacities

Open access informatics tools to empower others to multiply SeeD impacts

SeeD's strategy, which combines cutting-edge genetic-characterization technologies with extensive and high precision field and laboratory evaluations of traits, generates huge volumes of complex data. High-performance data management and software tools, therefore, are essential for success. A partnership with DArT is refining a database and application programming interface (API) to fit the data-management requirements of SeeD's beneficiaries. The API already communicates with programs to collect field data, to manage seed inventories and to pipeline routine statistical analyses. It is currently being linked to the breeding management software tool developed by the Integrating Breeding Platform (IBP). In a complementary effort, the James Hutton Institute (JHI) is adapting its data warehouse, web portal and visualization tools such as Flapjack and CurlyWhirly to fit the requirements of SeeD's current and future users.

Researchers worldwide will be able to add and extract value to or from this global resource

Priorities are often specific to locations and groups of users. Regardless of their trait of interest, researchers will be able to evaluate a core set of genetically diverse accessions and, empowered by the tools and know-how assembled by SeeD, tap into SeeD's comprehensive genetic atlas to identify and obtain biodiversity likely to meet their needs. SeeD's research will 'open the black box of genetic diversity.' The resulting 'molecular atlas' will help breeders hone into complex traits, such as heat and drought tolerance, disease resistance etc.

Nearly 130 beneficiaries – 38 percent of them women – have participated in workshops or graduate theses

Capacity-strengthening efforts designed to assist breeders and researchers to fully use the new genotyping platforms, databases and software, are an integral part of the project. SeeD has conducted a social network analysis of maize- and wheat-breeding professional networks in Mexico to enable better targeting of workshops and training opportunities to key beneficiaries to maximize the impact and reach of the training events sponsored by the project.

Future directions and SeeD milestones

SeeD is currently on a fundraising drive to enable the expansion of the project beyond Mexico, making it truly global in scope and impact. Figure 3 maps major outputs of SeeD to a timeline from 2015 to 2021, contingent on obtaining the required funding.

The main proposed milestones include:
- A publicly available database with molecular characterization of maize and wheat collections.
- Phenotype at least three priority traits each for maize and wheat in Africa and/or Asia, to extend the immediate geographic impacts of SeeD while demonstrating the utility of the platform.

- Identify promising accessions and alleles/haplotypes for at least three priority traits for maize and three for wheat and make them available to breeding programs worldwide.
- Avail bridging germplasm for at least two priority traits each for wheat and maize.
- Make publicly available a validated basic informatics tool kit.
- Build scientific expertise and experience in using the platform.
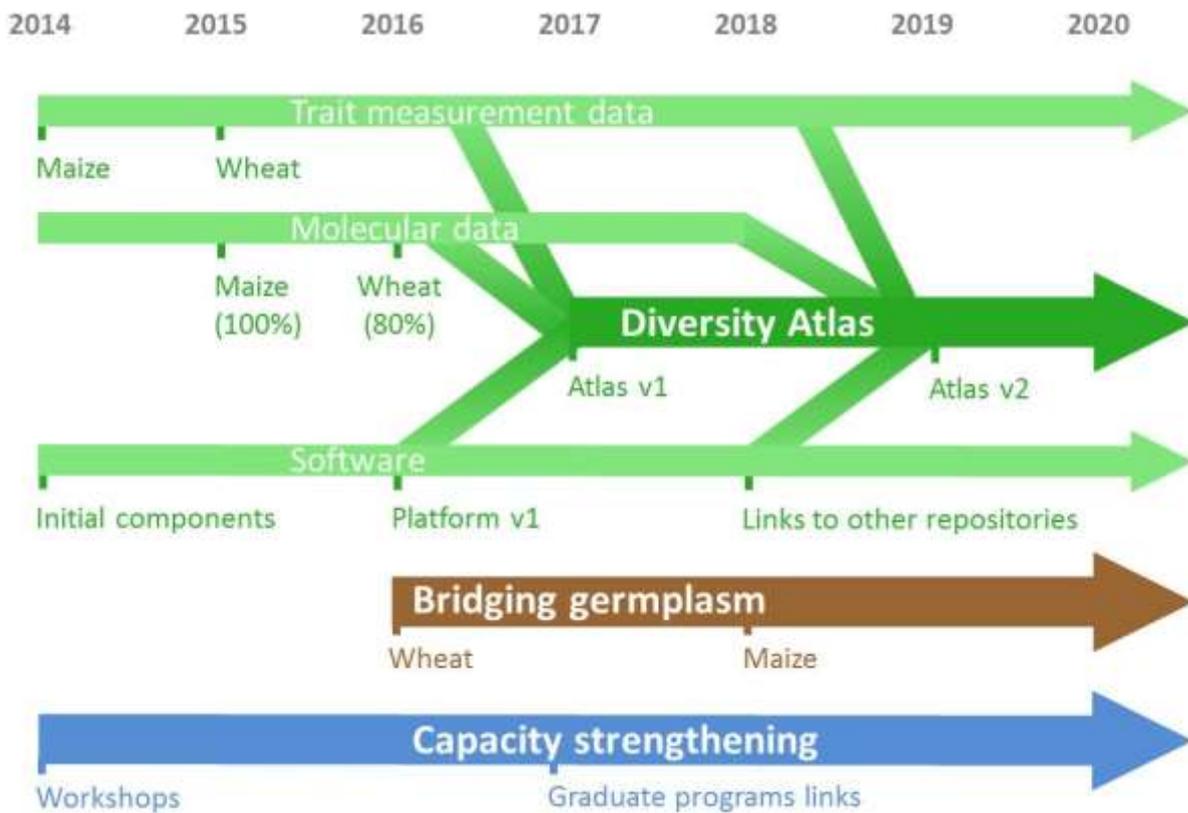- Ensure continued system maintenance and support beyond the life of the project.

Figure 3. Timeline illustrating the proposed milestones for the Seeds of Discovery (SeeD) project, contingent on obtaining the required funding.

What others say about SeeD

"Indian wheat breeding programs will be immensely benefitted through collaboration with the SeeD initiative for improvement of drought and Nutrient Use Efficiency (NUE), yield and quality improvement in both maize and wheat and terminal heat tolerance in wheat." – **H.S. Dhaliwal, Vice Chancellor, Eternal University, Himachal Pradesh, India**

"The project is yet another great opportunity and platform for both CIMMYT and AATF to expand and strengthen existing partnerships in crop improvement towards food sufficiency in Africa." – **Denis T. Kyetere, Executive Director, African Agricultural Technology Foundation (AATF)**

"The *Embrapa portal for Maize and Sorghum expects that this "new wave of research," encoded by the SeeD Project, will be as useful for Brazilian maize breeding programs as the first wave of germplasm introductions in the country... These materials were the basis of the modern maize tropical genetics in Brazil.*" – **Sidney Parentoni, Associate Director of Research and Development, EMBRAPA**

"This project is a tool that will allow us to move forward at a speed that would be impossible to reach through traditional means." – **Maria Esther Rivas, Director of BIDASEM, a Mexican seed company**

"...Your work in this field, will become the backbone of international efforts in this field. I also think that you will contribute breakthroughs in developing breeding methods..." "I am happy to confirm a strong positive opinion on SeeD scientific organization and vision, and my strong interest for further cooperation with your project." – **Alain Charcosset, INRA Research Director, Project Lead, Maize Integrated Genomics (Amaizing)**

"SeeD is a "blueprint" and "foundational initiative" for the Diversity Seek (DivSeek) initiative for other crops." – **Global Crop Diversity Trust**

"We are fully supportive of this project... these genotyping approaches with accurate phenotyping and pre-breeding will allow to speed up our CGIAR breeding programs for the smallholders' benefit..." – **CG Consortium Office**

References

Gur A, Zamir D (2004). Unused natural variation can lift yield barriers in plant breeding. *PloS Biol* 2:e245

Imai I, Kimball J, Conway B, Yeater K, McCouch S, McClung A (2013). Validation of yield-enhancing quantitative trait loci from a low-yielding wild ancestor of rice. *Molecular Breeding* 32:101-120

McCouch et al. (2012). Feeding the future. *Nature* 499:23-24