

SEEDS OF DISCOVERY

HARNESSING BIODIVERSITY FOR FOOD SECURITY



Crop biodiversity is a major underutilized resource to meet increasing demands for food, adapt to climate change and ensure a food-secure and sustainable future for humankind. CIMMYT's Seeds of Discovery (SeeD) project is the most significant initiative worldwide to make the biodiversity contained in the vast gene bank collections of the world available for much wider use. The publicly available products of SeeD will enable breeders to develop the new maize and wheat varieties needed to meet the demands of a growing population in a changing climate. SeeD's pioneering approaches are recognized as a guide for other initiatives to similarly secure and enhance the future productivity of other crops. You can be part of SeeD's vision to ensure crop productivity that feeds future generations.

Summary

The initial outputs from the SeeD project, briefly described below, offer a game-changing platform for the exploration and use of gene bank collections, enabling the identification, capture and incorporation of novel value into breeding programs. The resulting more resilient, nutritious and productive maize and wheat varieties give new options to farming families whose livelihoods are increasingly threatened by climate change (for example, hotter temperatures and new crop diseases) and by declining access to resources (ground water, arable land, fertilizer, etc.).

SeeD builds on opportunities generated by other projects, particularly the human genome sequencing project, to release the potential of genetic diversity to address the challenges of feeding the world's growing population. SeeD's strategy includes: 1) characterizing and developing atlases – at molecular-genetic and phenotypic levels – for the maize and wheat collections in CIMMYT's and other gene banks; 2) developing tools and methods to extract knowledge from these atlases; 3) placing the atlases and tools in the public domain; 4) training scientists worldwide how to use these tools to address the specific needs of their clients; and 5) developing and publicly availing maize and wheat lines with useful genetic variation for high-priority traits.

Without the information, tools and capacity-building legacy of SeeD the world's gene banks may remain museums of natural diversity. More importantly, current rates of productivity gains for maize and wheat are insufficient to keep pace with demand. Delays or failure to seize the opportunities available today through the SeeD project may leave breeders unable to address the intensifying and new challenges to crop production.

Further work is needed to capitalize on the initial investment and realize the potential sustainable flow of impacts from the SeeD project.

Why SeeD is Needed?

The earth's population will exceed 9.5 billion by 2050. Global food production needs to rise by 70 percent to meet the increased demand. Over the last century, food production increased primarily through increased land area cultivated and improved yields. Although grain yields in both developed and developing countries are still increasing, the rate of increase is slowing and a serious shortfall in staple grain availability is projected by 2050. In addition, despite the increase in global yields the World Health Organization estimates that more than 3 billion people are malnourished (deficient in intake of calories, protein, micronutrients). This is the largest number and proportion of malnourished people ever reported and is disproportionally represented by those born and living in developing nations, significantly limiting the development capacity of these countries.



The need to produce more food, improve dietary nutrition and ensure food security is critical to the maintenance of healthy, productive populations and stable governments. The food riots of the last decade are testament to the interplay between the food availability of a populace and social unrest.

Increases in global cereal staple production cannot be met by planting larger areas of land or adding more inputs to maximize potential yields. Additional crop land is very limited and continued deforestation or cultivation of historic pasture lands to enable cereal production is not sustainable.

Increasing costs and competition for inputs are exacerbated by well-known global climate variability (for example, La Niña/El Niño cycles) and emerging impacts of climate change. Increased incidences of stresses resulting from drought, heat, flooding, frost and wind damage, coupled with shifting dynamics and the spread of pests, diseases and weeds to new areas, all converge to decrease the stability of crop production systems. While the heaviest impact of climatic variability is felt in the developing world – where doubling of yield is required to feed the population of 2050 (Figure 1), the grain belts of the world will also be negatively affected, squeezing food production capacity globally.



Figure 1. Observed rates of percent yield changes per year (1961-2008) for (a) maize, (b) rice, (c) wheat and (d) soybeans. Red areas show where yields are declining whereas the fluorescent green areas show where rates of yield increase – if sustained – would double production by 2050. From Ray *et al.*: Yield trends are insufficient to double global crop production by 2050. PLoS ONE, 2013

There is no silver bullet to address the challenges which lay ahead. But we know that continuing the same approaches of recent decades will fail to meet humanity's needs.

SeeD addresses the needs for more resilient, nutritious and productive varieties by unlocking and mapping the vaults of gene banks. SeeD will provide breeders with new materials and tools to accelerate the development of the varieties that are more able to flourish under future conditions.

Genetic Diversity and Crop Productivity "Insurance"

Improved varieties better able to withstand the variability and pressures of current and future agricultural production environments are a primary input to a suite of technologies and practices which can contribute to meeting the food production needs of future generations.

Breeders need genetic variation to be able to select favorable genes and develop varieties. New genetic variation can be obtained by creating mutations, transgenics or using natural crop variation.

- Mutation breeding has been used for decades and has been most successful for simple traits. Mutation breeding has become less common with the mainstreaming of transgenics which are considered a more targeted methodology.
- Transgenics have been developed in many crops with increasing commercialization in some, including maize. They offer the possibility of using genes from other species to address objectives that are otherwise impossible using natural variation in a crop. Adoption of transgenic products is constrained by consumer perceptions in some societies. The costs of development of transgenic products are US \$70 to \$500 million per trait.
- Use of native genetic variation does not face the challenges of regulation or public perception faced by transgenic technology. Use of native variation can be achieved at a fraction of the cost of a commercial transgenic product; the entire SeeD initiative can be funded at similar cost to the development of one or two commercial transgenic products.

The genetic diversity present in the gene pools of maize, wheat and their associated wild relatives is a massive and valuable resource for plant breeders. The genetic diversity used in modern varieties and breeding efforts, however, represents only a small part of the gene pool diversity. Most breeders do not have the resources needed to sample the whole gene pool. The few breeders who have accessed the broadest genetic diversity have, during selection of the best materials for their target environment, discarded many genes from breeding pools. Although discarded genes have undesired effects in a target environment, they may have a positive impact on performance in a different environment. This process is a type of genetic bottlenecking (Figure 2).

Maize and wheat gene banks house thousands of accessions (varieties). These collections contain a wealth of genetic variants which are no longer represented in the modern breeding germplasm used to make new varieties for farmers' fields. These unused variants have potential value to farmers; for example, tolerance to the parasitic weed *Striga*, which devastates maize production in many regions of Africa, was found and introduced into maize varieties from an accession of teosinte obtained from CIMMYT's gene bank¹. Another example is the excellent levels of drought and heat tolerance found in maize lines developed over years of selection from accessions². Nonetheless, despite the extensive genetic potential contained in gene bank collections, the valuable variations remain mostly unused by plant breeders.

Lack of information relevant to breeding questions – for example, which of the materials are drought- or disease-tolerant – results in breeders facing a "black box" of information when considering which of the many thousands of accessions in gene banks to use.

Consider walking into a supermarket to buy a can of beans and finding that all the cans in the supermarket have had their labels removed; the information needed to select the can most appropriate to your needs is not available. This is the conundrum facing breeders. This lack of description of genetic resources, coupled with the difficulties of "taming" materials obtained from gene banks (which in addition to the genes of interest may carry many negative traits that breeders

Page **3** of **8**



¹ Amusan et al. 2008. New phytologist 178:157-166.

² Cairns et al. 2013. Crop Science 53:1325-1346

have spent time removing from their breeding programs), constrains the use of valuable genetic resources and effectively turns many gene banks into museum.



Figure 2. Genetic bottlenecking of a population. The original population has genes with a favorable effect in the target environment, genes with a deleterious effect in the target environment and genes with a beneficial effect in a different environment. Breeders create new varieties by selecting for beneficial genes while discarding deleterious genes for their target environment. Genes which are beneficial in another environment are not selected and are lost.

To bring high value genetic variation to breeders, SeeD will describe the materials in the gene banks and re-package and concentrate the valuable genetic variation into new germplasm products termed "bridging germplasm." Bridging germplasm helps span the common productivity chasm between unimproved gene bank materials and modern varieties.

Breeders need "unique–value" bridging germplasm, coupled with statistical approaches and informatics tools to enable its efficient use in their breeding programs. By combining the novel genetic variation in bridging germplasm with variation already present in their breeding programs, breeders can produce varieties which are tolerant to current and future stresses and have improved food, feed or industrial quality. These varieties will carry a form of insurance of productivity in adverse environments, a "genetic insurance" that new maize and wheat varieties will be able to feed and support the societies inherited by our children and grandchildren.

SeeD's Strategy and Achievements to Date

SeeD's strategy encompasses five main objectives (figure 3). Objectives 1, 2 and 3 are similar for maize and wheat except for differences in the relative importance of different traits (for example, crop-specific diseases). Objectives 4 and 5 focus on information management, training and capacity strengthening, which are cross-cutting for both the wheat and maize research communities.

SeeD's strength is rooted in worldwide partnerships that gather expertise ranging from cutting-edge DNA profiling technologies to large-scale field phenotyping to address objectives. The DNA-sequencing techniques developed by partners and used by SeeD are up to 100,000 times more powerful than those used to map the first human genome. The traits evaluated to date for each crop were prioritized by panels of experts including breeders, physiologists, pathologists and nutritionists, and include characteristics of global importance such as tolerance to drought, heat, low soil fertility and diseases prevalent in Latin America. The informatics tools developed, adapted and customized by SeeD are oriented to the exploration and use of crop genetic resources.



Figure 3. Seeds of Discovery research strategy.

A sample of outputs produced by SeeD to date includes:

Maize:

- DNA profiles for more than 20,000 landraces; by the end of 2014 all materials available to the world through CIMMYT's gene bank will have been genetically profiled.
- This information has already been used to identify accessions of high potential value for drought and heat tolerance which are not represented by modern breeding lines, enabling scientists to focus attention on more detailed assessment of these materials.
- An extensive study has been implemented to pinpoint relationships between DNA information and traits including drought tolerance and disease resistances. High value genetic variants have been identified for disease resistance, nutritional quality, flowering time and more. The best variants will be deployed in bridging germplasm.
- Six multi-accession breeding populations have been formed to develop bridging germplasm for stresses such as drought, low soil fertility and diseases. These populations capture important novel genetic variation for use by breeders to benefit farmers and consumers.

Wheat:

 50,000 genetic profiles of wheat accessions, with an additional 110,000 samples well under way. This information has been used to identify a core set of a few hundred representative accessions which will be characterized in detail for important traits.

Page **5** of **8**



- The largest search for heat and drought tolerance for any crop in history was initiated by evaluating over 70,000 wheat varieties in field trials. Heat- and drought-tolerant accessions were identified for further evaluations to select the best for use in bridging germplasm.
- More than 200 populations are in development to enable the identification of important genetic variants and the development of bridging germplasm.

Technology and Capacity Building:

- Informatics tools (including Curly Whirly and Flapjack) have been adapted and are being used by researchers and breeders to explore, select and use valuable genetic variation in the gene bank.
- Information management tools to enable processing and handling of large data volumes including a DNA sample tracking system and a hand-held computer tablet-based phenotypic data capture tool have been developed and are being used by SeeD researchers and others in the broader wheat and maize communities.
- An intellectual property (IP) framework to enable public data dissemination in an IP-sensitive manner.
- More than 100 researchers have been trained on genetic and information technology skills for breeding. Developing the next generation of scientists who will use the bridging germplasm and knowledge outputs from SeeD is a key goal.

Seeds of Discovery: Food for the Future

SeeD requires an investment of US \$10 million per year to achieve its aims in seven years. With this funding, SeeD will build on the initial and preliminary outputs achieved to date, exploiting synergies to reach project goals. Five of the envisioned main outputs for the future of SeeD are described in figure 4 with complementary value descriptions in table 1.

Translation of outputs to impacts will begin through partnerships and continue for decades through the efforts of clients of SeeD's genetic resources utilization platform. Building of capacity for use of SeeD's tools and establishing a structure that ensures continued maintenance and support for use of these tools after the SeeD project ends are essential to the vision of success.

SeeD will develop outputs that create a self-sustaining, enabling environment for the use of an efficient genetic resources utilization platform by the maize and wheat breeding communities.

Already the value and potential of SeeD are recognised globally:

"...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)

"Embrapa 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

The technologies are ready and the time is right to leverage the value stored in the world's gene banks for the benefit of farmers and future generations.

SeeD is supported by the Mexican Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food (SAGARPA) as one of the components of the MasAgro initiative (http://masagro.mx/index.php/en)



Figure 4. SeeD major output schedule and exit scenario based on a \$10M per year investment portfolio from 2015 to 2021. Balance of activities between the objectives of molecular atlas (), novel alleles (), allele donors and bridging germplasm development () and capacity building () changes over the lifespan of the project empowering breeders to use information, knowledge, germplasm and tool packages during and beyond the lifetime of the initiative.

Page **7** of **8**



| Output | Value to primary clients | Additional/secondary benefits |
|--|---|--|
| 2016 Maize | Improved access to international genetic resources. The generation and presentation of a standardized | Rationalization of germplasm collections. |
| | information "adds labels" and detailed information to the "cans" of germplasm present in the international | For example, identification of potential |
| | collections. A maize breeder in Zimbabwe is now able to compare the genotype of her maize with that of | duplicates. |
| Iviaize | accessions present in the gene bank and make more informed selections and requests for germplasm | Evaluation of gene bank activities and |
| atlas onlino | | assistance in priority setting. For example, |
| atias online | | identification of well-represented and rare |
| | | genetic variants. |
| 2017 New | Ability to use multiple information resources and select accessions of highest potential value for tailored | Gene banks are better able to respond to |
| | applications. The development and refinement of approaches to combine genotypic, passport and other | client requests. For example, "I want wheat |
| methods to | characterization data and use a synergy of information resources to select appropriate germplasm for | accessions with late maturity and drought |
| idontify | defined applications. A wheat breeder in India would like drought-tolerant early maturing wheat which | tolerance." |
| | brings new genetic diversity to his breeding program. Combining genotypic and passport information, he is | Breeders are better able to identify |
| of high value | able to identify and obtain accessions which are early maturing, genetically complement his germplasm | complimentary germplasm by adopting |
| | and originate from environments with high frequency of drought during the growing season. | methods available through SeeD. |
| | 2018: Sources of genetic variation for heat, drought and resource use efficiency. Genetic variation not | Improved use and renewed interest in the |
| | available within breeding pools made available in easily adoptable package for use in the development of | exploration and conservation of crop |
| 2018 and | more durable varieties for farmers. | genetic resources. |
| 2020 | 2020: Sources of genetic variation for heat, drought and resource use efficiency, disease resistance and | |
| New | improved nutritional quality. Broadening of the base of genetic variation for target traits. Genetic | |
| bridging | variation not available within breeding pools made available in easily adoptable package for use in the | |
| germplasm | development of more durable and nutritious varieties for farmers. | |
| | A maize breeder in the Lake Victoria region of Kenya is able to request lines which introduce novel genetic | |
| | variation for heat tolerance, maize streak virus tolerance and starch quality to his breeding program. | |
| | Improved capacity for marker-based selection of parents and breeding germplasm within and across | Cross-species applications to leverage |
| 2019 Information, knowledge and | breeding programs. More effective breeding programs with more targeted and accelerated incorporation | information from wheat and maize and |
| | of new genetic elements into breeding pools and faster genetic gain. | apply in close and more distantly related |
| | Use of information from many different genetic backgrounds, studies and environments will facilitate the | species for research and breeding. |
| | timely and cost-efficient selection of germplasm and tracking of the genetic elements most favorable to | Use of informatics tools in different |
| informatics | specific breeding programs. A maize breeder in Mexico is able to select new bridging lines which have | breeding systems. |
| tools | known novel alleles for drought tolerance. Using genomic selection coupled with background selection, she | |
| | is able, in two years, to transfer the drought tolerance from the white grain bridging germplasm to her | |
| | drought-sensitive blue maize lines, something that would take at least four years without using markers. | |

Table 1. Value of SeeD outputs to primary clients and additional, spill-over benefits.