SeeD

A learning model towards effective and equitable use of genetic resources

El Batan – 15th February 2017
Seeds of Discovery (SeeD) (MasAgro Biodiversidad)

- Initiated September 2011
- Mostly funded by the Mexican government

Exploring the untapped biodiversity for maize and wheat
Four pillars of work

- Genotypic characterization of germplasm banks and public elite germplasm
- Phenotyping and Marker-trait associations
- Pre-breeding
- Capacity Development
  Data Management
Genotyping

DArTseq for both wheat and maize
GbS for select maize GWAS analysis
KASP for select wheat markers and samples
Genotyping wheat

- Genomic characterization of DArTseqSNP and SilicoDArT markers
- Wheat Diversity Analysis (CIMMYT/ICARDA)

- ~6,000 Wild relatives
- 21,000 Tetraploids
- ~63,000 Hexaploids

Diversity Core Set

Global diversity analysis
Diversity Profiles of Gene Bank Accessions

General genetic diversity
DArTSeq SNP

Elites

Exotics (FIGS, Australian hot, primary synthetics)

Functional diversity
~20 KASP
Diversity analysis within important groups using SNP D genome

930 Aegilops tauchii D genome

43 Aegilops ventricosa DN genome

CIMMYT
ICARDA
Genetic Core sets of Wild relatives

Full collection: 5000
Core set collection: 500

Full collection: 5000
Core set collection: 1000
Genomic regions on 21 wheat chromosomes having fixed alleles in Mexican landrace accessions of Chihuahua and Central Valley region.
Genotyping maize

- Genomic characterization of DArTseqSNP and SilicoDArT markers
- Modified protocol- composite samples

- 25000 Landraces
- ~2000 Pools, Pops, Varieties, Composites
- ~400 Wild relatives
- 570 CML

Diversity analysis
Breeder oriented cores
Selection sweeps
Diversity

15384 Landraces with GIS data
Breeder cores

$1M \text{ usd}$

$200k \text{ usd}$

$100k \text{ usd}$

$20k \text{ usd}$
Selection sweep

Environmental GWAS

Pcadapt
XP-CLR

9954 acc
Focus on field-based phenotyping

<table>
<thead>
<tr>
<th>Traits</th>
<th>Maize</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abiotic stresses</strong></td>
<td>heat</td>
<td>heat</td>
</tr>
<tr>
<td></td>
<td>drought</td>
<td>drought</td>
</tr>
<tr>
<td></td>
<td>low N</td>
<td>low P</td>
</tr>
<tr>
<td><strong>Biotic stresses</strong></td>
<td>tar spot, ear rot, stalk rot, Turcicum, Cercospora, MLN (MCMV &amp; SCMV)</td>
<td>tan spot, spot blotch, (Fusarium, blast), karnal bunt</td>
</tr>
<tr>
<td><strong>Grain quality</strong></td>
<td>hardness, starch, oil, amino acids, phenolics</td>
<td>hardness, color, protein, test weight, yellow berry, Fe, Zn</td>
</tr>
</tbody>
</table>

- **Wheat**: >1.5m data points from 30 trials across 9 locations
- **Maize**: >1.5m data points
Impact of heat on wheat

- ~10% yield loss per 1°C increase in temperature
- By 2050, 20-30% yield loss in South Asia alone, affecting over 1 billion people

SeeD: ~70,000 wheat gene bank lines screened under heat stress (2011-2013)
Exploring the Gene Bank for Heat Tolerant Wheat

Mexican landraces with grain yield >150 g m\(^{-2}\) under heat stress (Cd. Obregón, México)

- Early flowering
- Late flowering
- Medium flowering

Control: Elite bread wheat (Sokoll), 186 g m\(^{-2}\)

- Tolerant Mexican landraces (YELLOW)
- Tolerant Iranian landraces (RED)
- Elite lines (BLUE & GREEN)
Screening Gene Bank for Yellow Rust – adventitious phenotyping

Yellow rust resistant wheat landraces

Association mapping for yellow rust in wheat landrace

<table>
<thead>
<tr>
<th>Accession / Pedigree</th>
<th>YR (%) in Mexico</th>
<th>YR (%) in India</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHIH95.5.18</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>CHIH95.5.23</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>DGO95.3.8</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>OAX93.1.1.1</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Susceptible check</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
Maize: Grain protein content

→ Comprehensive **catalogue of >100 haplotypes** and associated markers which increase grain protein content
Identified 1367 accessions of value for acid soil tolerance evaluation and 353 accessions carrying favorable allele for acid soil adaptation.
Realizing promise - pre-breeding

• Development of inbred and semi-inbred germplasm which contains high value exotic introgressed alleles in elite backgrounds

• Good agronomic performance is needed in addition to novel genetic variation
Gene Bank to Breeding Pipeline

8000 Mexican landraces

Michoacán: Rare allele hot spot

(A) Diversity analysis

(GBS profiling)

(20K SNPs)

Phenotype (Heat, Drought, Quality)

(B) Core set development

(C) Core set evaluation

(D) Trait donor identification

Heat, Drought Diseases, Quality

(E) Pre-breeding

Large-scale efforts to move useful gene bank diversity to breeding pipelines
Wheat ‘bridging germplasm’

Linked topcross panel (LTP)
- Evaluate effects of exotic alleles in elite genetic backgrounds
- Large number of small, but expandable populations (ca. 10,000 lines in total)
- TC chains with partly overlapping parents to facilitate joint linkage/association mapping
Selections made from 600 populations under heat and drought stresses
Evaluations made from 8000+ advanced pre breeding lines at El Bataan and Toluca
Wheat ‘bridging germplasm’

- Over 1000 donors used in pre-breeding, generating >15,000 F6/7/8 fixed lines
- Preliminary selections of 2000 with acceptable agronomy, drought and heat tolerance
- Partners evaluating pre-bred germplasm
Maize ‘bridging germplasm’

Useful novel alleles & haplotypes

Early generation lines & pools enriched for favorable alleles

…using **multiple strategies** defined by trait complexity and breeder needs (desired input germplasm, demand for new sources)

<table>
<thead>
<tr>
<th>Breeder demand</th>
<th>Trait complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Urgent</strong></td>
<td>Monogenic (1-3)</td>
</tr>
<tr>
<td></td>
<td>Oligogenic (4-10)</td>
</tr>
<tr>
<td></td>
<td>Polygenic (&gt;10)</td>
</tr>
<tr>
<td><strong>Medium-term</strong></td>
<td>MABC</td>
</tr>
<tr>
<td></td>
<td>MARS &amp; prediction index</td>
</tr>
<tr>
<td></td>
<td>GS with MABC for BC1S2 development</td>
</tr>
<tr>
<td><strong>Long-term</strong></td>
<td>MABC &amp; GS</td>
</tr>
<tr>
<td></td>
<td>MARS, prediction index &amp; GS</td>
</tr>
<tr>
<td></td>
<td>GS with MABC for BC1S2 development</td>
</tr>
</tbody>
</table>
Assessment of options; simulations
GS options

Capacity Development

• 238 researchers, professors, and graduate students in courses and workshops 2012-2015.

• 33 PhD, MSc, & BSc students in SeeD to date.

• Scientists are conducting research projects to apply SeeD products in their own programs.
Targeting capacity-strengthening efforts

- Workshops
- Online learning modules
- Visiting scientists and projects
- Postgraduate program

maize breeders

wheat breeders
Putting knowledge together - Molecular Atlas
Components

- **Data** - genotypic, phenotypic, GIS, passport.

- **Knowledge** - marker-trait associations, germplasm panels, protocols, “how to”.

- **Tools** - data collection software, online query tools, data visualization tools and software, statistical analysis methods, training links.

KDSmart – Android based tablet and phone phenotypic data collection.

KDXplore – Computer based trial data management and data curation tool.

Germinate 3

Online data warehouse and knowledge center with data query, visualization and download capability germinate. seedsofdiscovery.org

Flapjack – Graphical genotype viewer helps identify germplasm and diversity of highest value.

CurlyWhirly – multidimensional data visualization to help understand genetic diversity and identify the most useful landraces.

www.seedsofdiscovery.org
• Deploy a computing hardware and software platform
• Integrate computing resources and expertise at Earlham, CIMMYT and NIAB
• Facilitate data-intensive bioinformatics analyses leveraging data generated by SeeD for the genetic improvement of wheat.
• Data will be available via public and open resources (Wheat Information System; Ensembl Plants browser).
• Add value to SeeD and IWYP, and enhance their benefits to the wheat research community.
EiB linkages

• **Data**- genotypic, phenotypic, GIS, passport.

• **Knowledge**- marker-trait associations, germplasm panels, protocols, “how to”.

• **Tools**- data collection software, online query tools, data visualization tools and software, statistical analysis methods, training links.

• **SAGA, training**
Thank you for your interest!