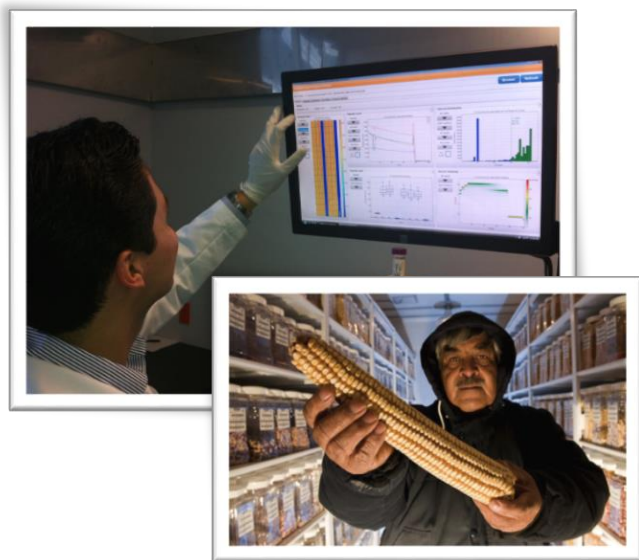


Letting the genie out of the maize genebank bottle



*Sarah Hearne,, Martha Willcox,
Juan Burgueño, Jiafa Chen,
Terence Molnar, Charles Chen,
Carolina Sansaloni, Cesar
Petrolli, Peter Wenzl, Alberto
Romero, Ed Buckler, Iain Milne,
Paul Shaw, Gordon Stephen,
Sebastian Raubach, Micha
Bayer, David Marshall, Andrzej
Killian, Andres Christian, Jorge
Franco, John Hickey, Gregor
Gorjanc and many, many others*

shearne@cgiar.org

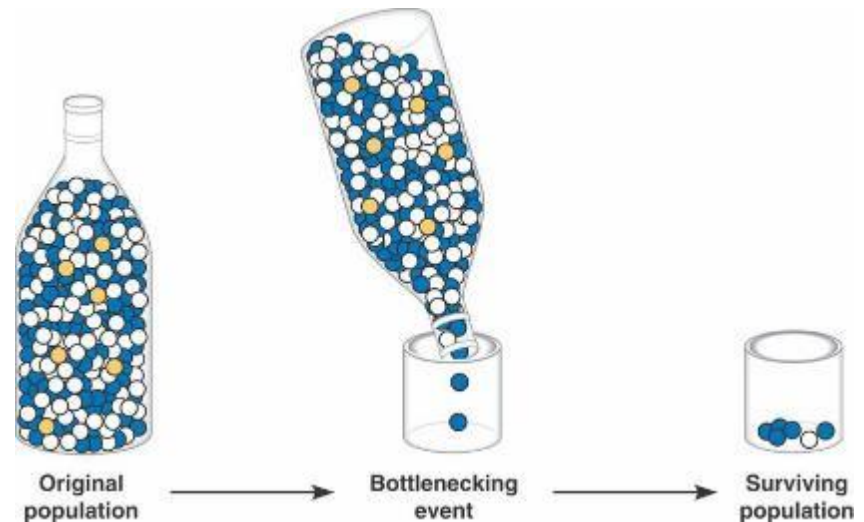
“Genebanks are not museums”

- Most of the requests to banks with maize and wheat germplasm holdings are for described elite lines or landraces / wild relatives with some publication history
- The majority of breeders would not “touch” an un-described landrace and would have to be “desperate” to use one even with very good characterisation



The genie in the genebank bottle

- Farmers and breeders have spent a long time removing unwanted genetic variation
- This bottlenecking was in many cases a good thing, the frequency of favourable traits was increased compared with deleterious traits.
- However, it carried the penalty of “throwing the baby out with the bathwater”- valuable genetic variation has been lost also during the selection process



<http://bio1151.nicerweb.com/Locked/media/ch23/bottleneck.html>

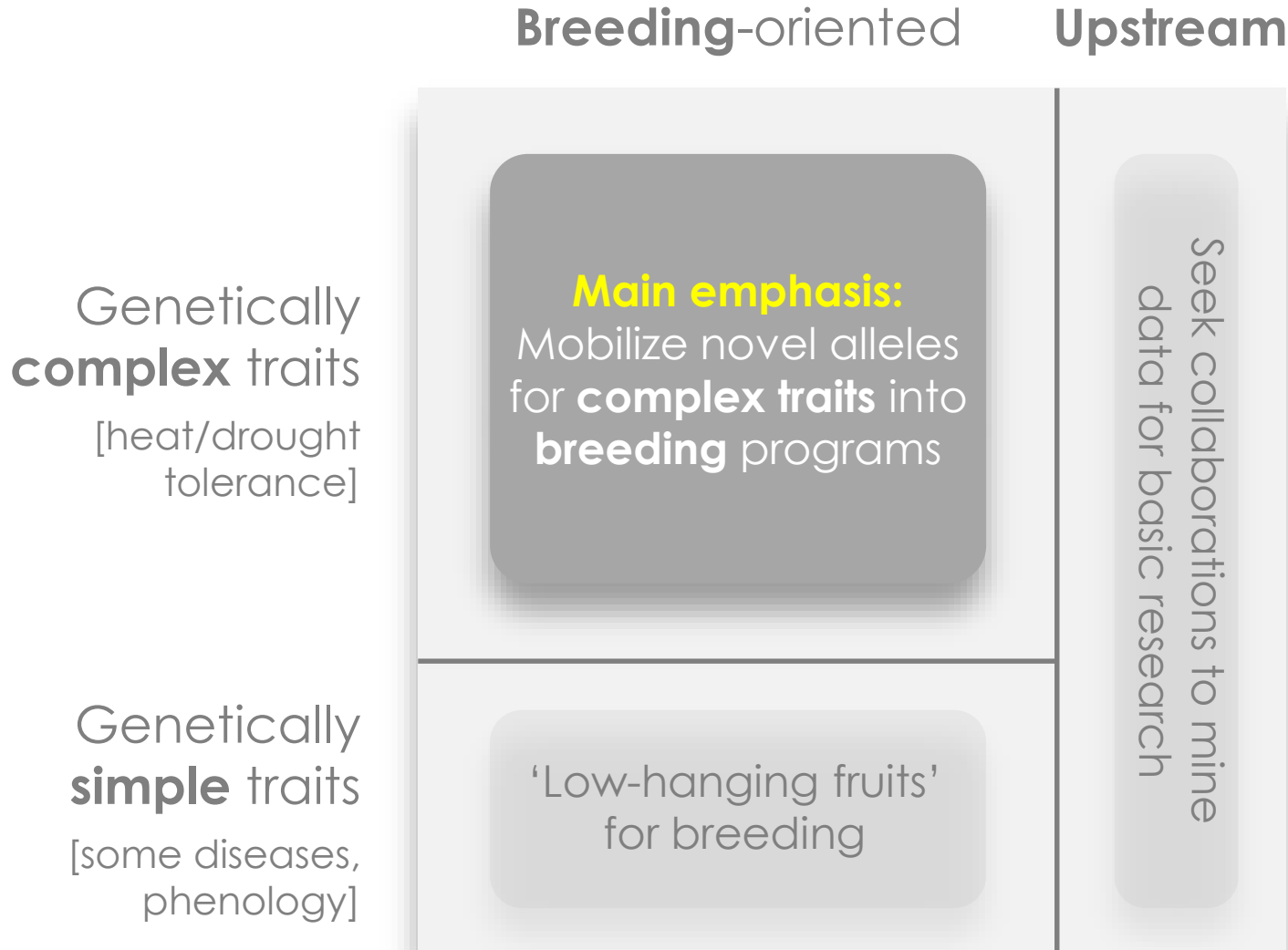
Bridging the divide- getting the needles from the haystack

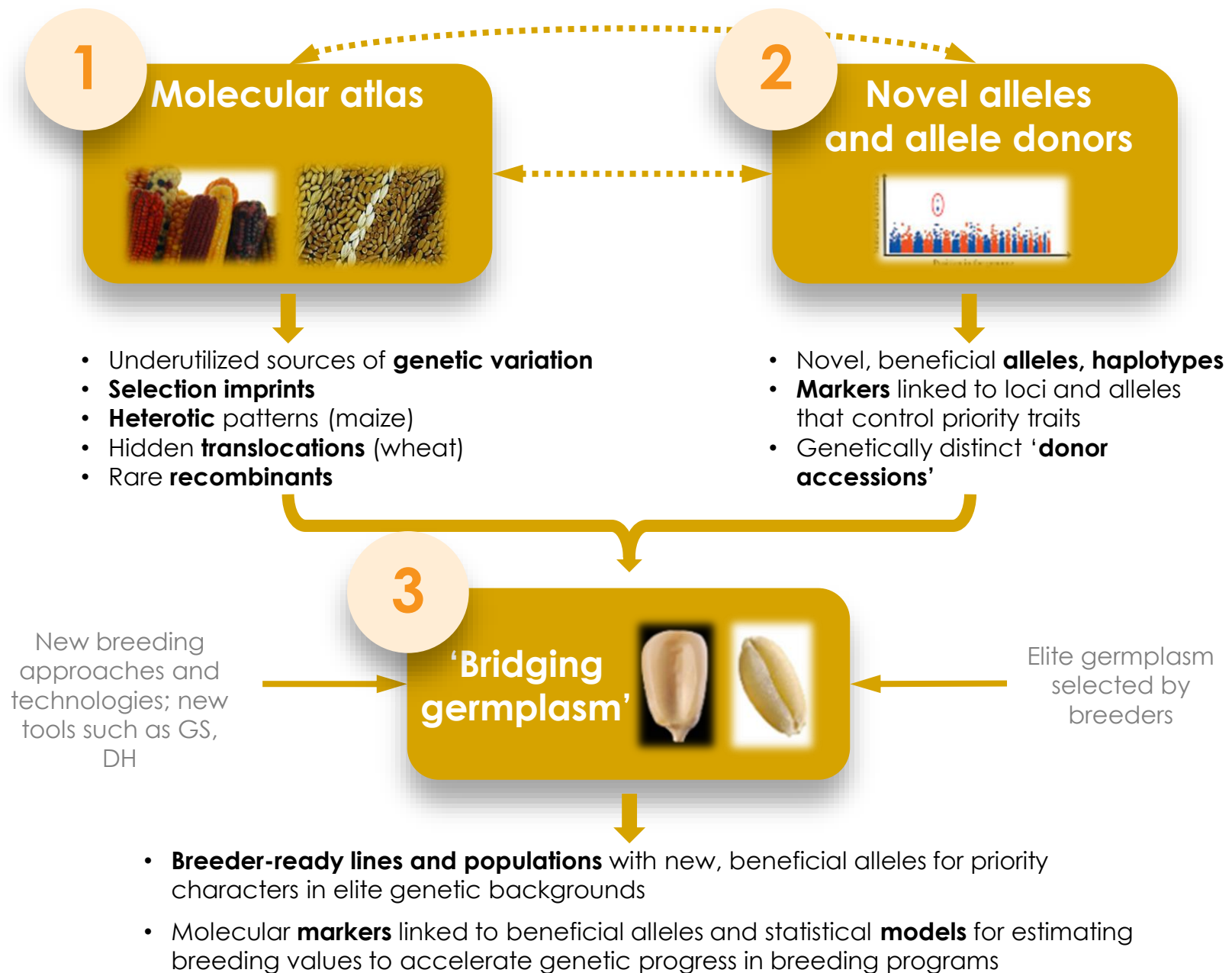


Information
Knowledge
Germplasm



Research emphasis



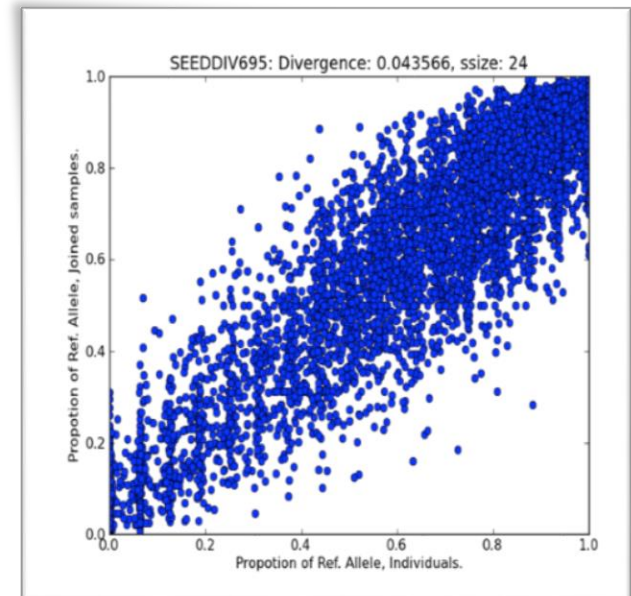




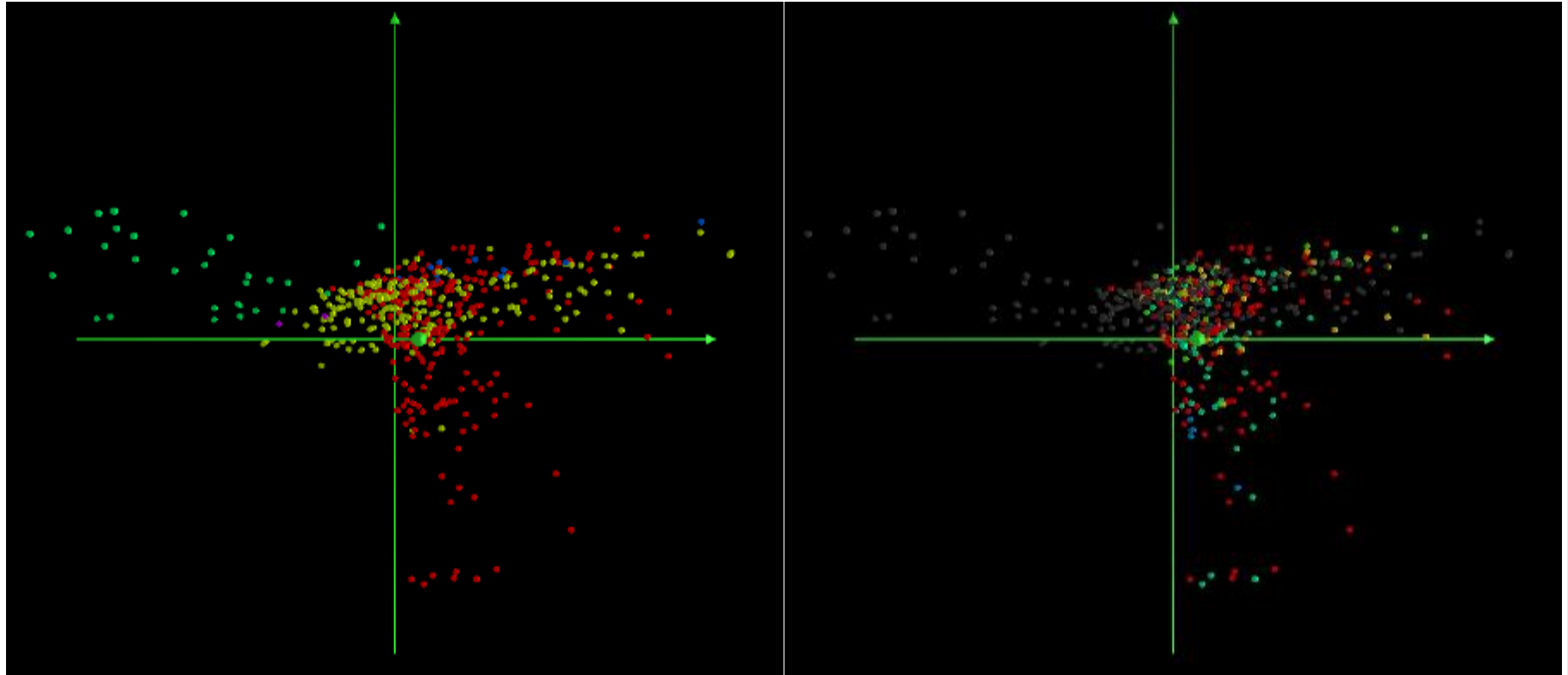
Molecular atlas

Molecular atlas

- **Maize**: ~27,500 accessions in CIMMYT's genebank further 15-20,000 unique accessions in Mexican genebanks
- Developed new GbS method for **composite DNA samples** (30 plants per accession); the method simultaneously
 - ▶ Quantifies **allele frequencies** within accessions (SNP), and
 - ▶ Estimates **genetic distances** among accessions (PAV)
 - ▶ 45,000 SNP per sample, 260k polymorphisms
- Completed the sequencing of **20,000 accessions** → diversity analysis in progress



Preliminary analysis



Analysis 2014

- Environmental selection footprints
 - ▶ 18,500 accessions with good quality geo-location data
 - ▶ Extracted long term abiotic environment data
 - ▶ Shifts in haplotype distribution and allele frequency with environmental cline across global collection
- Breeding selection footprints
 - ▶ Cycles of recurrent selection populations genotyped
 - ▶ Evaluate response to selection
 - ▶ Evaluate selected genomic regions in derived lines
- Use selection footprints
 - ▶ Race footprints and use footprints

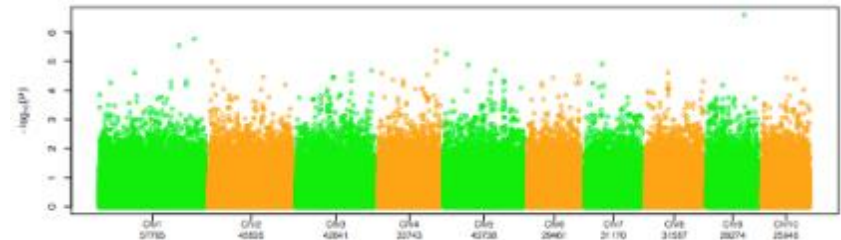
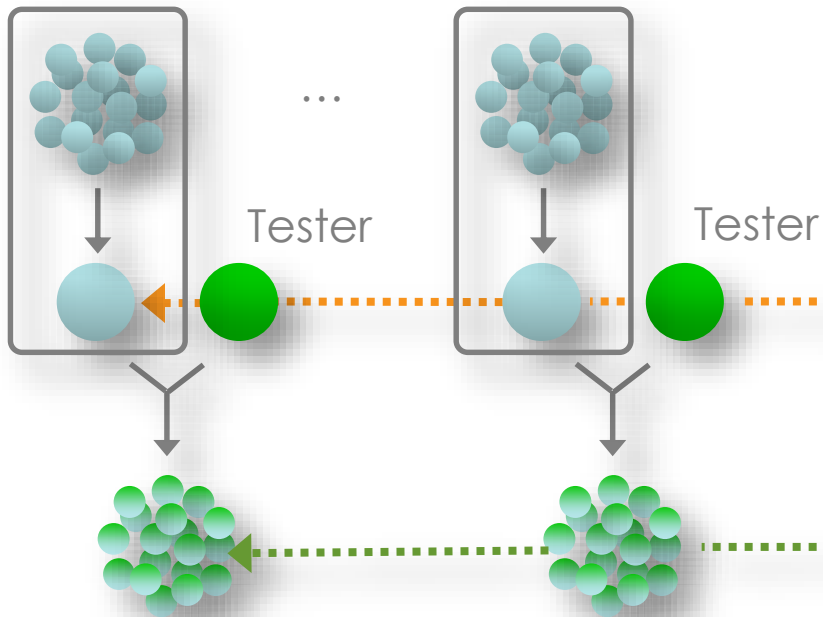


2

Novel, beneficial
alleles and allele
donors

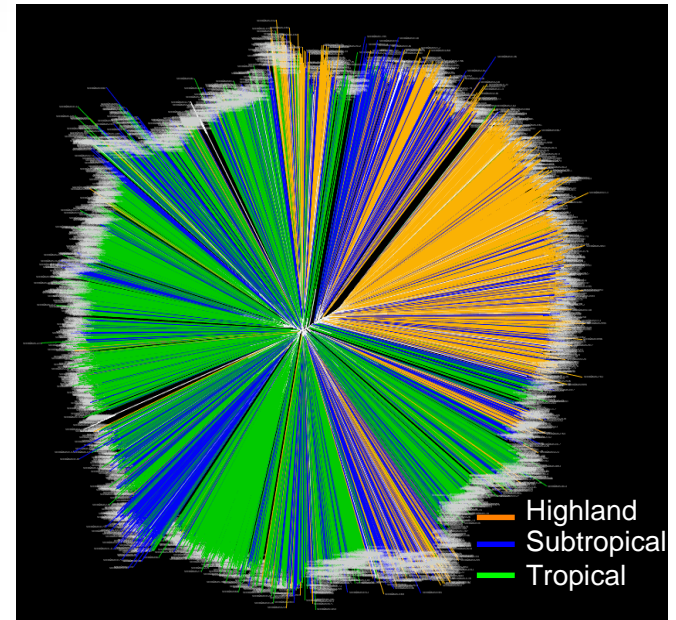
SeeD Maize GWAS

Accession 1 ... Accession 4,000



Three adaptation zones:

- Tropical
- Subtropical
- Highland



36 Latin American countries

Accurate, field-based phenotyping is the main bottleneck

Traits	Maize
Abiotic stresses	heat
	drought
	low N
Biotic stresses	tar spot, ear rot, stalk rot, <i>Turcicum</i> , <i>Cercospora</i>
Grain quality	hardness, starch, oil, amino acids, phenolics

- **Maize**: 700,000 data points from 34 trials across 14 locations



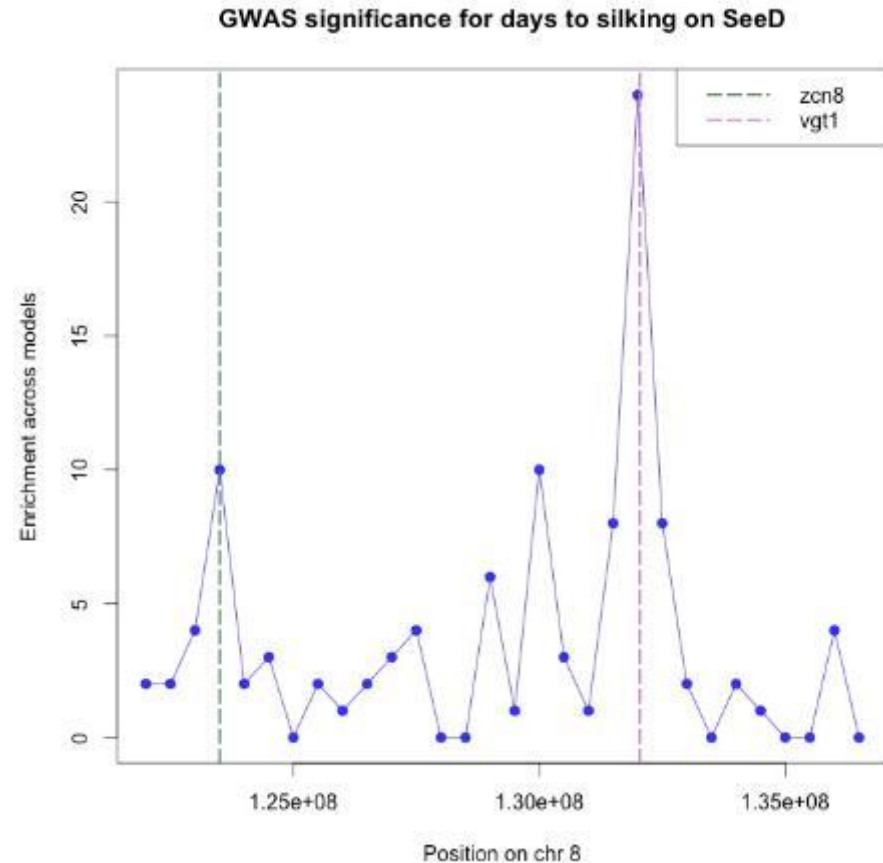
GWAS in SeeD: Association analysis shows overlap in previously reported loci

Association at known loci can provide insight into statistical power

500k SNP, imputation with FILLIN
GWAS were performed per trial using BLUPs for days to silking using naïve GLM and GLM+Q. Non parametric meta-analysis

There are markers with significant association at and close to Vgt1 and ZCN8

In conclusion, we can perform genome wide association in the SeeD panel



Tar Spot Disease Complex

Caused by *Phyllachora maydis* and *Monographella maydis* in association

Yield loss for up to 30% in Mexico in farmers fields (Hock et al. 1995); 46% in unsprayed on-station test plots (Bajet et al. 1994)

Most severe in sub-tropical adaptations (700-1,600 m)



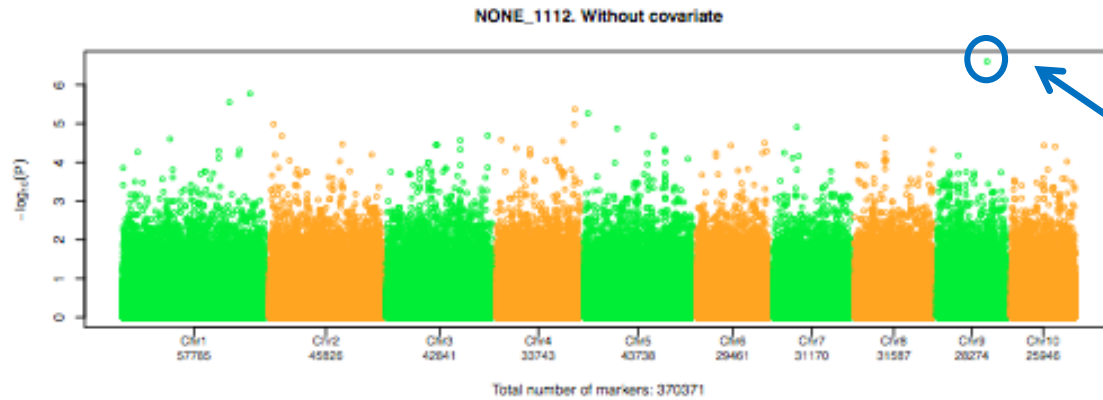
source: CIMMYT

- 1st generation Tar Spot GWAS panel
- 842 gene bank accessions testcrosses evaluated Chiapas in 2011 and 2012
- Average visual mean and averaged row mean
- Filtered GBS2.7 SNP (365,000 SNPs)



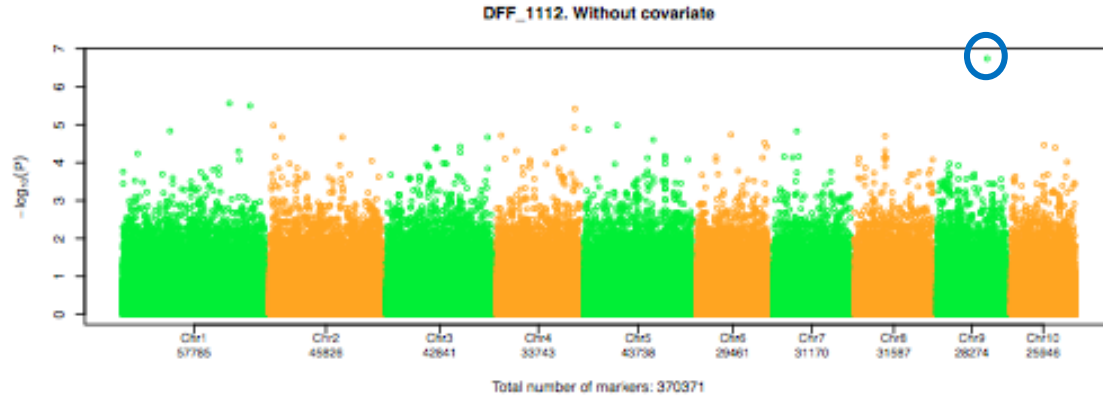
source: www.cabi.org

Tar Spot
Disease
Response
BLUP GWAS
without
covariate

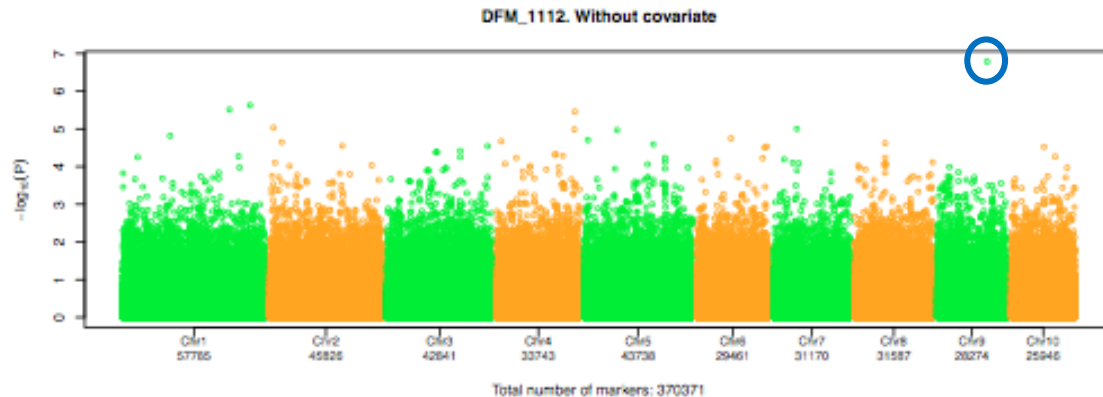


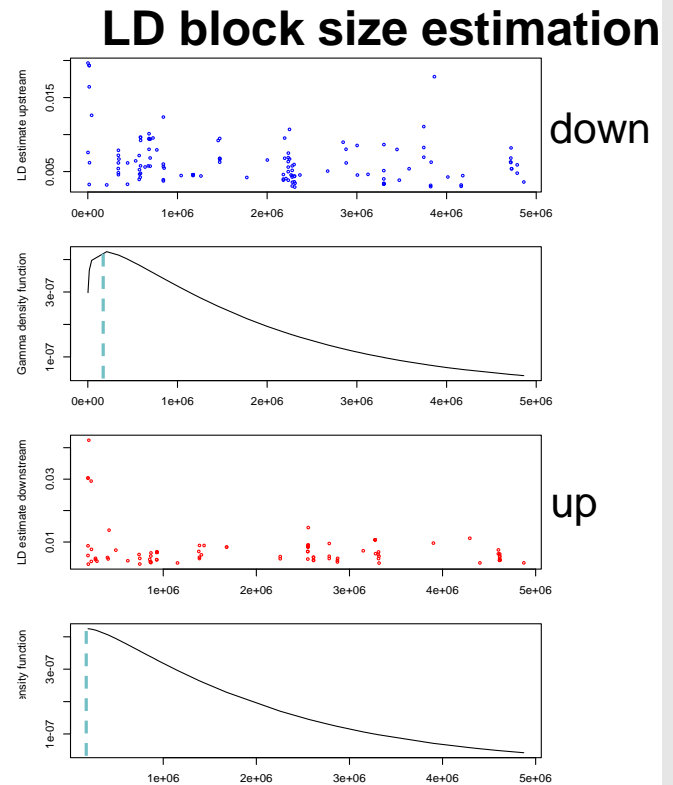
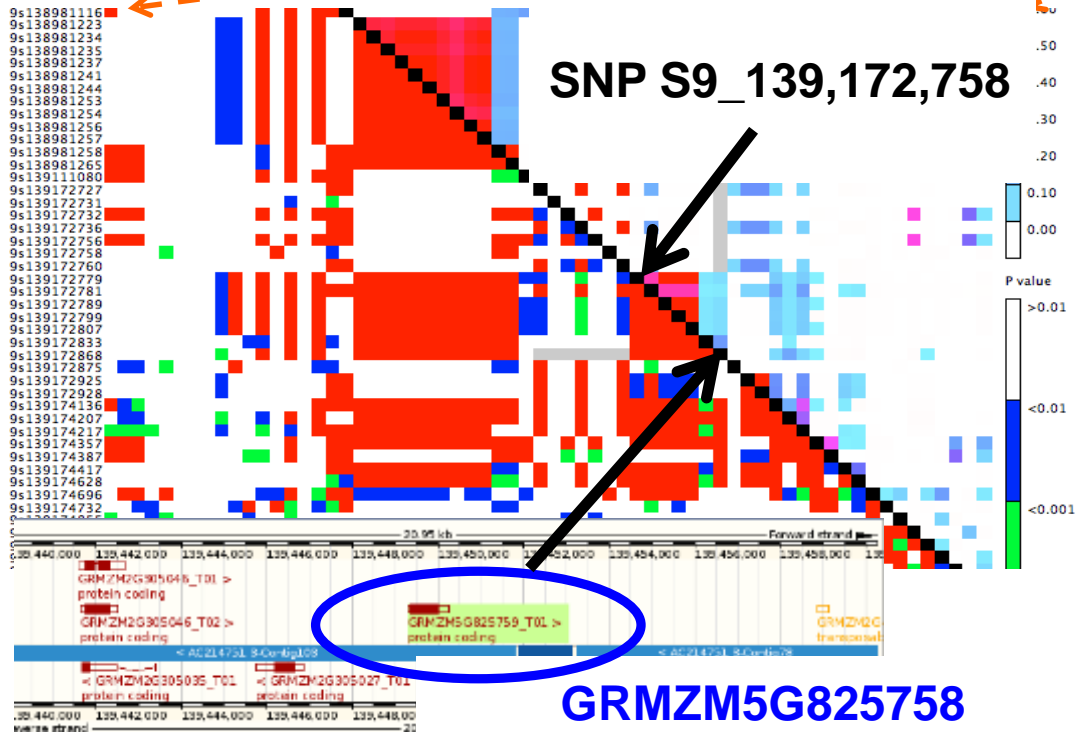
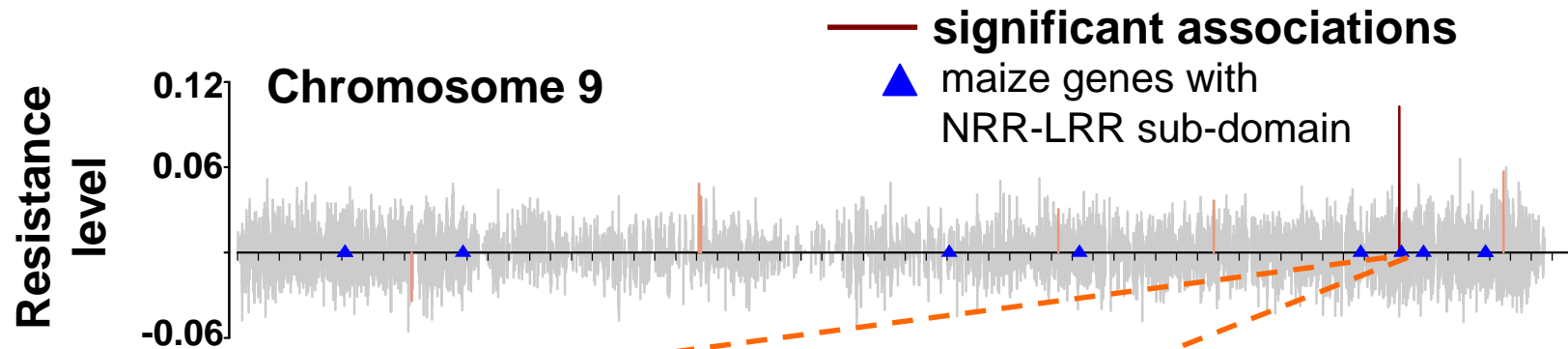
Chr9
pos: 139,172,758
 $p = 1.01e-7$

Tar Spot
Disease
Response
BLUP GWAS
with FFT
covariate



Tar Spot
Disease
Response
BLUP GWAS
with MFT
covariate



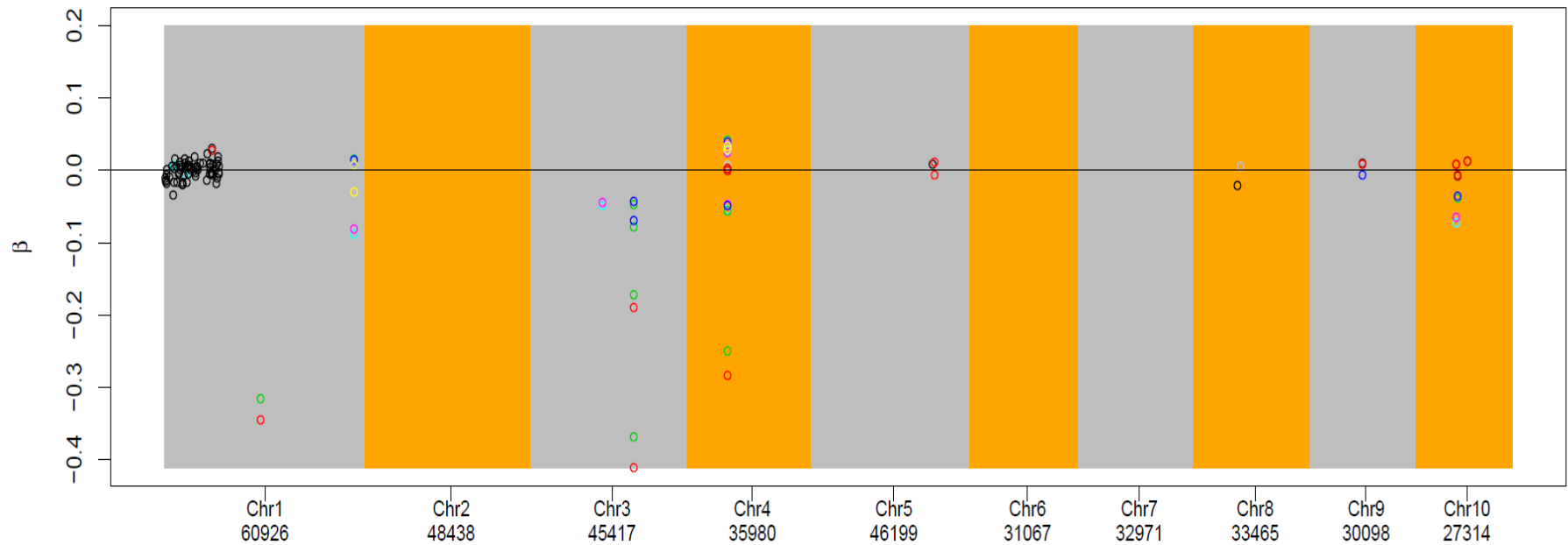


270kb – shorter than mean in NAM

Quality traits

Starch content, oil content, protein content, carotenoid content and profile (14 traits)

Effect plot of 17 traits



Total number of markers: 391875



3

Pre-breeding →
'bridging germplasm'

Maize 'bridging germplasm'

Useful novel
alleles &
haplotypes



Early
generation
lines & pools
enriched for
favorable
alleles

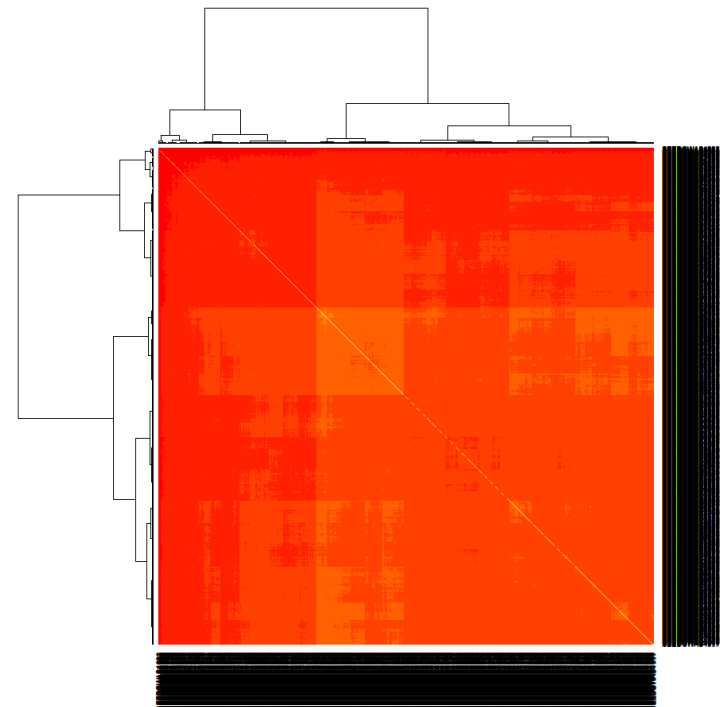


...using **multiple strategies**
defined by trait complexity
and breeder needs
(desired input germplasm,
demand for new sources)

Breeder demand	Trait complexity		
	Monogenic (1-3)	Oligogenic (4-10)	Polygenic (>10)
Urgent	DH from landrace & landrace / line crosses, selfing	DH from landrace & landrace / line crosses, selfing	GS with MABC for BC1S1 development
Medium-term	MABC	MARS & prediction index	GS with MABC for BC1S2 development
Long-term	MABC & GS	MARS, prediction index & GS	GS with MABC for BC1S2 development

Pre-breeding- GWAS/GS

- G-matrix
- Phenotypic ranking of the 4000 testcrosses
- Description of adaptation, colour, grain type “heterotic pattern” not A/B
- Selection of best 20 accessions per adaptation per heterotic group
- What to use for GS test population?
 - Testcrosses
 - DH from accessions
 - Accessions
- Simulations
 - Germplasm input, N_e , marker density and coverage, re-training and test pop strategy



Conclusions simulations

Approach

- ▶ Lower accuracy in initial selection using test-cross materials but higher genetic merit in C4 (beware of reconstructing the tester!!!)
- ▶ More gain up to C4 using landrace DH than segregating materials but add two more seasons for making DH (lower rate of gain)

Genotyping platform

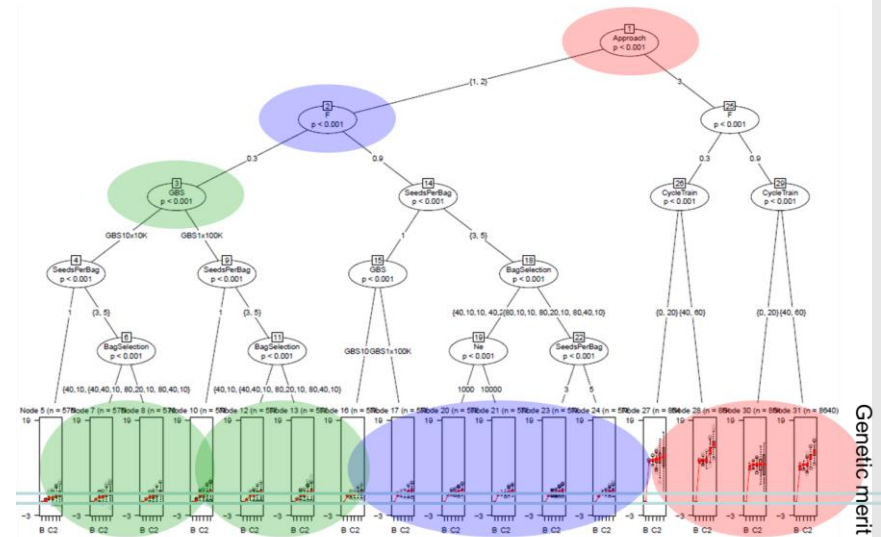
- ▶ larger chip better with high Ne (GBS10x,10K not enough with large Ne=10K)

Retraining

- ▶ Improves accuracy and gain (40 individuals likely enough)

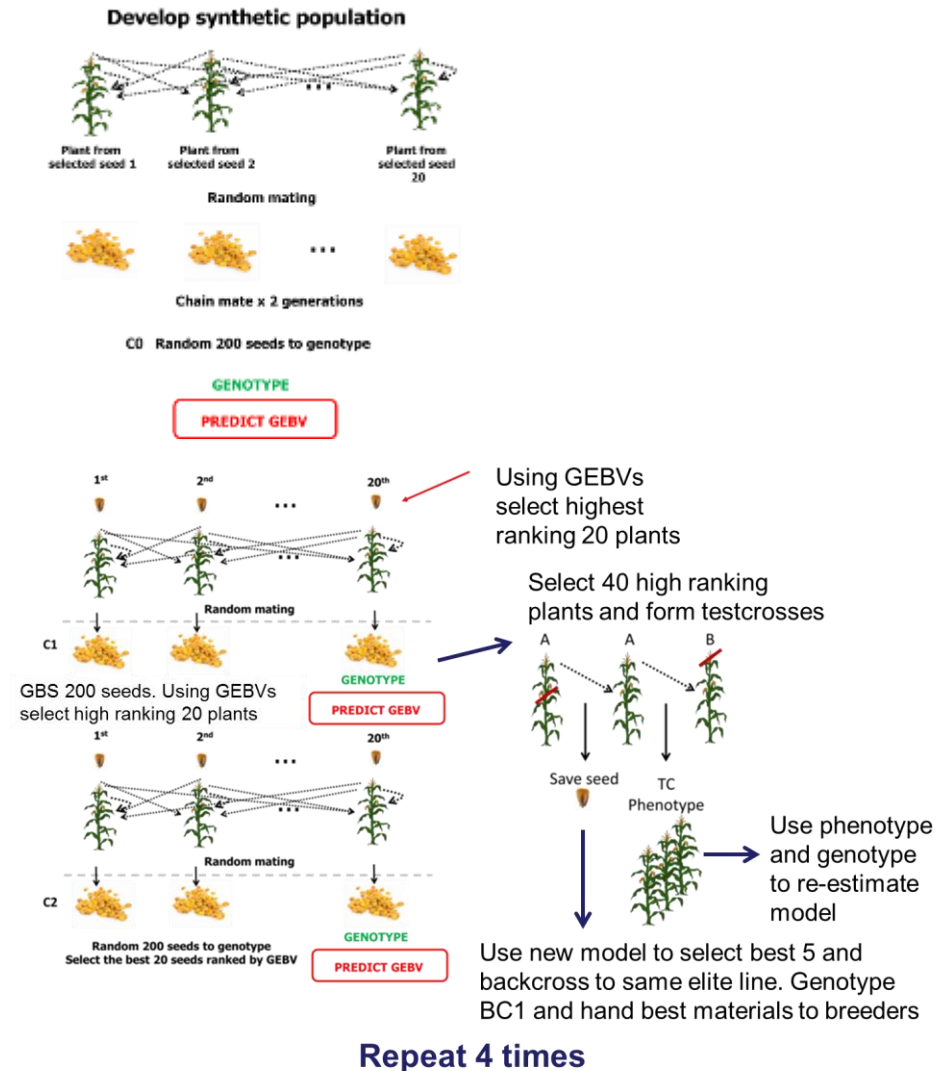
Test more seeds per landrace

- ▶ 3 plants per landrace better for GS applications than looking across more backgrounds



GS populations

- Six populations
- 2 rounds chain mating to increase recombination among genomes
- Between 7 and 14 founders per population
- 100k marker density explored



Bridging the divide- getting the needles from the haystack



Information
Knowledge
Germplasm



Thank you!

<http://seedsofdiscovery.org>; seed@masagro.org

Participants from Mexican
institutions

Participants from
CIMMYT

Participants from
other countries



Jonás Aguirre (UNAM), **Flavio Aragón** (INIFAP), **Odette Avendaño** (LANGE BIO), **Ed Buckler** (Cornell Univ.), **Juan Burgueño**, **Vijay Chaikam**, **Alain Charcosset** (AMAZING), **Gabriela Chávez** (INIFAP), **Jiafa Chen**, **Charles Chen**, **Andrés Christen** (CIMAT), **Angelica Cibrian** (LANGE BIO), **Héctor M. Corral** (AGROVIZION), **Moisés Cortés** (CNRG), **Sergio Cortez** (UPFIM), **Denise Costich**, **Lino de la Cruz** (UdeG), **Armando Espinosa** (INIFAP), **Néstor Espinosa** (INIFAP), **Gilberto Esquivel** (INIFAP), **Luis Eguarte** (UNAM), **Gaspar Estrada** (UAEM), **Juan D. Figueroa** (CINVESTAV), **Pedro Figueroa** (INIFAP), **Jorge Franco** (UDR), **Guillermo Fuentes** (INIFAP), **Amanda Gálvez** (UNAM), **Héctor Gálvez** (SAGA), **Karen García**, **Silverio García** (ITESM), **Noel Gómez** (INIFAP), **Gregor Gorjanc** (Roslin Inst.), **Sarah Hearne**, **Carlos Hernández**, **Juan M. Hernández** (INIFAP), **Víctor Hernández** (INIFAP), **Luis Herrera** (LANGE BIO), **John Hickey** (Roslin Inst.), **Huntington Hobbs**, **Puthick Hok** (DARt), **Javier Ireta** (INIFAP), **Andrzej Kilian** (DARt), **Huihui Li**, **Francisco J. Manjarrez** (INIFAP), **David Marshall** (JHI), **César Martínez**, **Carlos G. Martínez** (UAEM), **Manuel Martínez** (SAGA), **Iain Milne** (JHI), **Terrence Molnar**, **Moisés M. Morales** (UdeG), **Henry Ngugi**, **Alejandro Ortega** (INIFAP), **Iván Ortíz**, **Leodegario Osorio** (INIFAP), **Natalia Palacios**, **José Ron Parra** (UdeG), **Tom Payne**, **Javier Peña**, **Cesar Petrolí** (SAGA), **Kevin Pixley**, **Ernesto Preciado** (INIFAP), **Matthew Reynolds**, **Sebastian Raubach** (JHI), **María Esther Rivas** (BIDASEM), **Carolina Roa**, **Alberto Romero** (Cornell Univ.), **Ariel Ruíz** (INIFAP), **Carolina Saint-Pierre**, **Jesús Sánchez** (UdeG), **Gilberto Salinas**, **Yolanda Salinas** (INIFAP), **Carolina Sansaloni** (SAGA), **Ruairidh Sawers** (LANGE BIO), **Sergio Serna** (ITESM), **Paul Shaw** (JHI), **Rosemary Shrestha**, **Aleyda Sierra** (SAGA), **Pawan Singh**, **Sukhwinder Singh**, **Giovanni Soca**, **Ernesto Solís** (INIFAP), **Kai Sonder**, **Maria Tattaris**, **Maud Tenaillon** (AMAZING), **Fernando de la Torre** (CNRG), **Heriberto Torres** (Pioneer), **Samuel Trachsel**, **Grzegorz Uszynski** (DARt), **Ciro Valdés** (UANL), **Griselda Vásquez** (INIFAP), **Humberto Vallejo** (INIFAP), **Víctor Vidal** (INIFAP), **Eduardo Villaseñor** (INIFAP), **Prashant Vikram**, **Martha Willcox**, **Peter Wenzl**, **Víctor Zamora** (UAAAN)

Contributed at the beginning: Gary Atlin, **Michael Baum** (ICARDA), **David Bonnett**, **Paul Brennan** (CropGen), **Etienne Duveiller**, **Mustapha El-Bouhssini** (ICARDA), **Marc Ellis**, **Ky Matthews**, **Bonnie Furman**, **Marta Lopes**, **George Mahuku**, **Francis Ogbonnaya** (ICARDA), **Ken Street** (ICARDA)