Genomic exploration and use of genebank collections for breeding: Early insights and experiences from Seeds of Discovery-Maize.

Sarah Hearne, Terry Molnar, Martha Willcox, Jorge Franco, Carolina Sansaloni, Cesar Petroli, Juan Burgueño, Jose Crossa, Alberto Romero, Charles Chen, John Hickey, Gregor Gorjanc, Janez Jenko, Andrzej Kilian, David Marshall, Ed Buckler, Peter Wenzl, Kevin Pixley + many more!



Modernización Sustentable de la Agricultura Tradicional



#### MasAgro Biodiversidad



GANADERÍA, DESARROLLO RURAL, Pesca y alimentación







Potentially valuable genetic variation, the raw material for crop improvement, **remains untapped** on genebank shelves, at a time **when challenges to crop production are unprecedented** 



Genebanks should NOT be museums. Genetic variation is the basic ingredient of all plant breeding Genebanks should be a source of highvalue genetic diversity to **meet tomorrow's challenges** 





#### Systematically identify & mobilize novel, beneficial genetic variation into breeding programs to accelerate and strengthen genetic gains



#### Maize modules













**Genebanks - supermarket** 

Common label - genotypic data





- Genotyped whole MGB
- Composite genotyping GbS (30 individuals per accession in one sample)
- GIS derived information





## Pop type dim2 Collection Composite Group Pools Teosinte Variety dim3 -dim1

## MAIZE genetic diversity survey

- GbS of entire CIMMYT maize genebank (>27,500) completed end 2014
- Initial analysis of 21,000 accessions:
  - 1.2m SNP loci in total
  - Mean 980k loci per accession (~ 20% missing data)
  - 317k loci with coverage ≥5X
- Also genotyping breeding materials (donors) and ex-PVP lines for comparison





ID accessions of interest

dim2 Cluster dim2 Cluster CIMMYT

ID New sources of high value characters and alleles

Combine data resources

- Drought: 9954 landraces come from environments with longterm propensity for drought during flowering (Low AI)
- Genetic analysis using GbS data indicates these landraces cluster into six main groups



## **ID** accessions of interest

 The six groups come from significantly different environments: All dry but some much drier than others- indicating some genotypic patterning- adaptation



LS Means of AI across clusters identified for three months of flowering and the entire 6 month growing period. The effect of cluster on AI for all three flowering months and over the growing period was significant (p<0.0001). Entries with the same letter code do not differ significantly (Tukey multiple comparison test <0.01 following ANOVA).





### SeeD Maize GWAS





#### **Breeder-oriented cores**





## Selection footprints / selection sweep

#### • Look for of selection motifs

· 	
8991F1p39:C>T (0)	
	Palome
· 李潔麗和以和時的有效思想。如何如此用於生物的出生。但是如此是我的人的目光,但是不可能。	
and the second	
· 通过的基础的推动,在某些正式的问题,在这些方面的问题,在这些方面。	
n an	Cápico
医无关的 建物物 化化物物 医子宫神经神经 经济的 化化学学 计分子分子 计分子分子 化合金化合金	
	Arroail
· 그는 것 같은 것 같이 되었다. 이 같은 것 같은	AHOCIII
	1966-2016 CIMMYT







## **Prioritization of traits**



## Genetically complex traits

[heat/drought tolerance]

Genetically simple traits

[some diseases]



## SeeD Maize GWAS





# Accurate, field-based phenotyping is the main bottleneck

Traits	Maize
Abiotic stresses	heat
	drought
	low N
Biotic stresses	tar spot, ear rot, stalk rot, Turcicum, Cercospora
Grain quality and nutritional quality	hardness, starch, oil, protein content, amino acids, phenolics, vitamin A, Zn

 Maize: 800,000 data points from 35 trials across 14 locations









# Signal in a new locus on chromosome 4

- Structural rearrangement-Inv4m locus Previously reported inversion in teosinte and highland maize (Hufford et al, 2013; Pyhäjärvi et al, 2013)
- Introgression with potential selective advantage



CIMMYT



### Inv4m locus 4 has an additive effect on days to anthesis





Drought

#### Alleles not present in elite germplasm





pre6MonthGrowAvg geoCov







### Thank you



GANADERÍA, DESARROLLO RURAL, Pesca y alimentación





research program Maize

