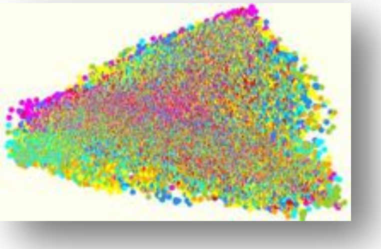
Extending the Frontiers of Genomics-Assisted Breeding for Grain Yield, Stress-Resilience and Quality of Bread Wheat Using the Reference Sequence

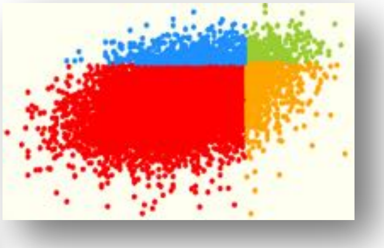
Philomin Juliana, Ravi Singh, Jesse Poland, Julio Huerta-Espino, Sandesh Shrestha, José Crossa, Leonardo Crespo-Herrera, Fernando Toledo, Velu Govindan, Suchismita Mondal, Uttam Kumar, Sridhar Bhavani, Pawan K. Singh, Mandeep S. Randhawa, Xinyao He, Carlos Guzman, Susanne Dreisigacker, Matthew Rouse, Yue Jin, Paulino Pérez-Rodríguez, Osval Montesinos-López, Daljit Singh, Mokhlesur Rahman and Felix Marza

CIMMYT

Application of the IWGSC's bread wheat reference sequence (Refseq v.1.0) for genomics-assisted breeding in the Global Wheat Program at CIMMYT

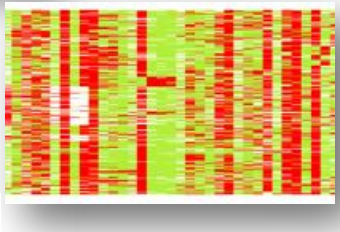


Genomic resources

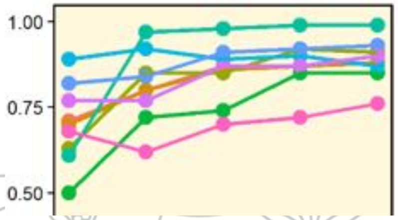


Genomic selection

Genome-wide association mapping

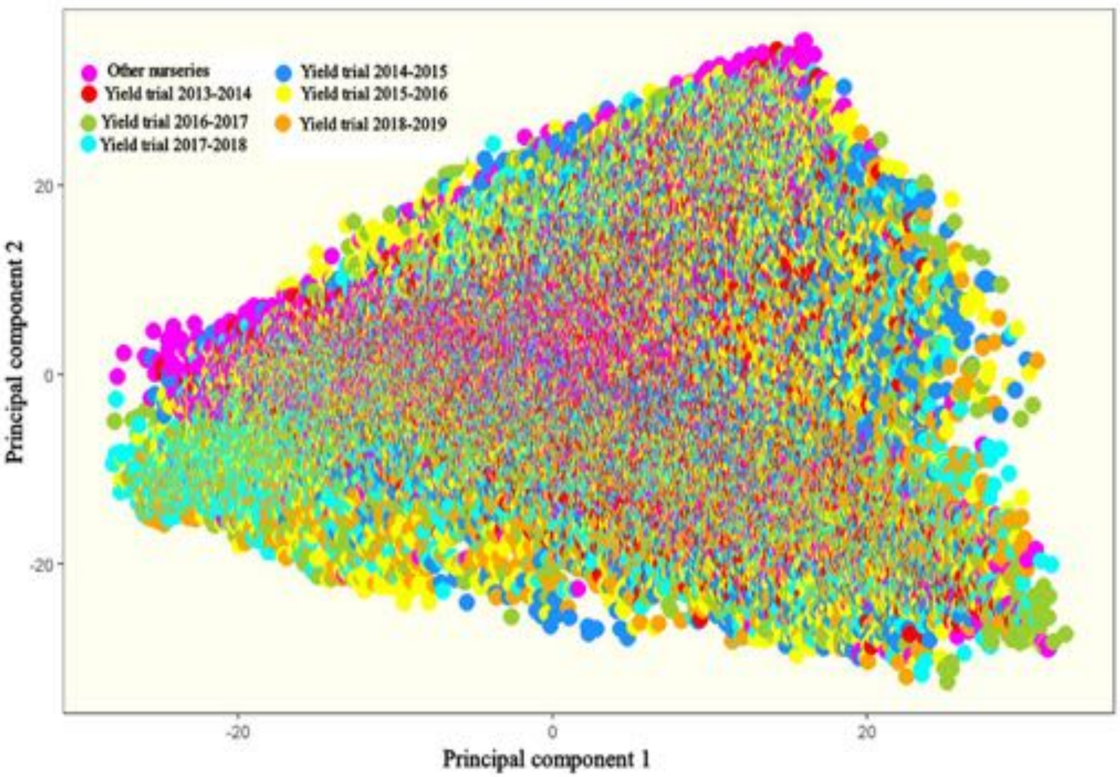


Genomic fingerprinting



Molecular tracking alleles

Genotyping and phenotyping resources in the Global Wheat Program



☐ Through the USAID Feed the Future and Delivering Genetic Gain in Wheat projects, a large panel of >62,000 CIMMYT wheat breeding lines in Stage 1 yield trial nurseries from 2013-2019 (>6,000 lines in each panel) was genotyped using >77,000 genotyping-by-sequencing (GBS) markers aligned to the RefSeq v1.0, generating over 4 billion marker data-points.

☐ About 3,485 lines from this panel that comprise the Stage 2 yield trials have phenotypic data on 24 traits including grain yield, disease resistance and quality related traits.


Yield trials, Ciudad Obregon, Mexico


Stem rust, Njoro, Kenya

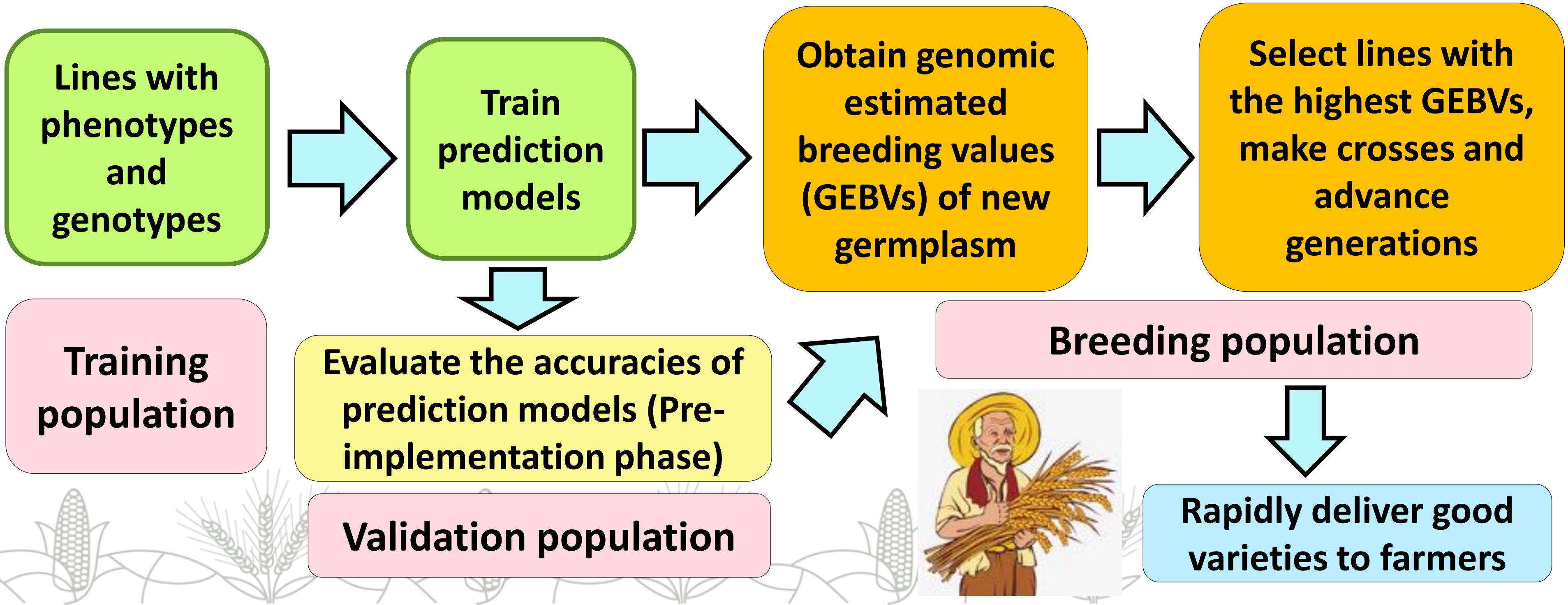

Septoria blotch, Toluca, Mexico


Quality lab, El Batan, Mexico

Evaluation of genomic selection in the Global Wheat Program

- **Phenotyping and selections can be laborious, expensive and time consuming.**
- **The current breeding cycle time for spring bread wheat in the GWP is ~5-6 years.**
- **Can we shorten the cycle time and quicken the process of delivering varieties to farmers using genomic selection?**


Toluca, Mexico


Lines with phenotypes and genotypes

Train prediction models

Obtain genomic estimated breeding values (GEBVs) of new germplasm

Select lines with the highest GEBVs, make crosses and advance generations

Training population

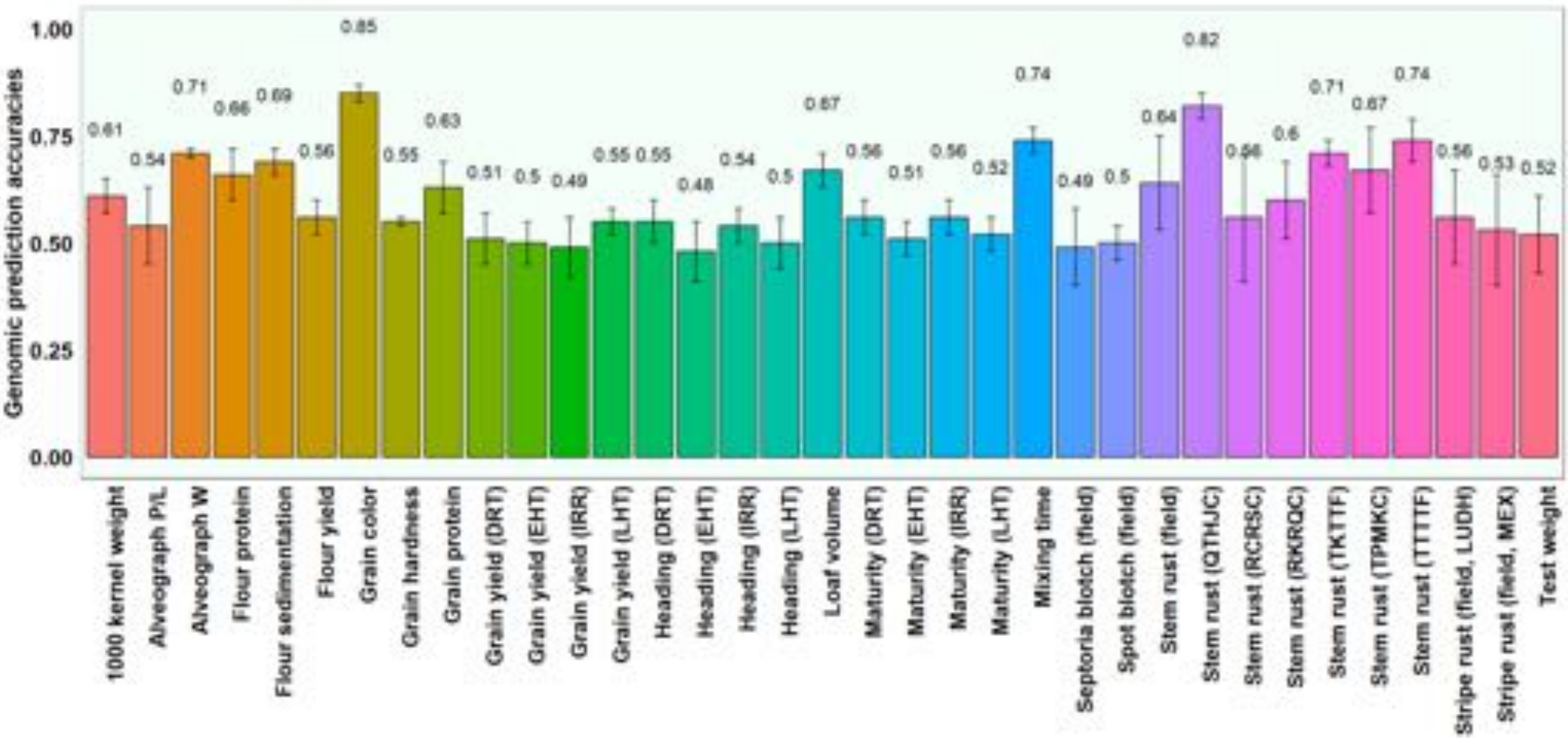
Evaluate the accuracies of prediction models (Pre-implementation phase)

Validation population

Breeding population

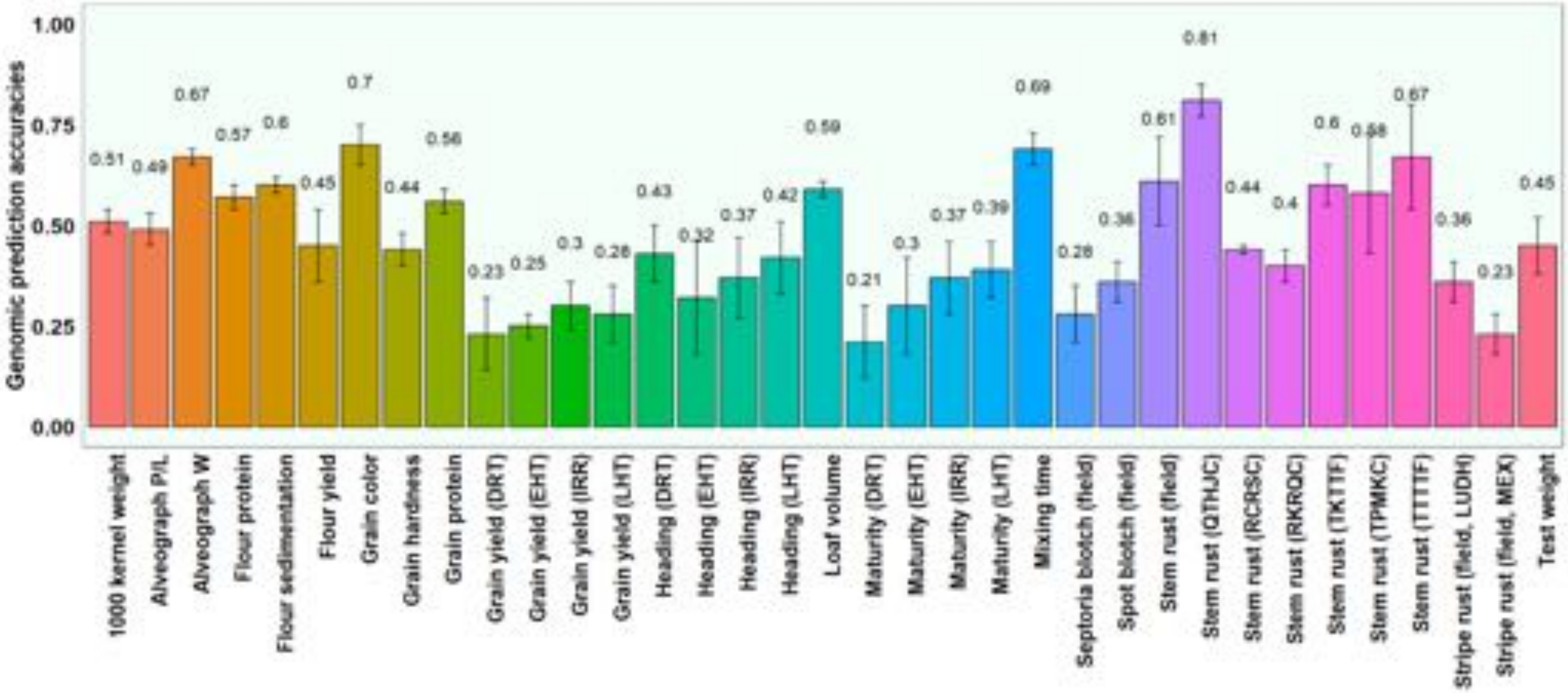
Rapidly deliver good varieties to farmers

Genomic prediction accuracies within panels and environments (five-fold cross validations)



Mean genomic prediction accuracies of four different panels with 1,092 lines each, evaluated during the 2014-2017 seasons. In each panel, about 218 lines were predicted using 874 lines from the same panel. DRT, drought stressed; EHT, early-sown heat stressed; IRR, irrigated; LHT, late-sown heat stressed; LUDH, Ludhiana; MEX, Mexico.

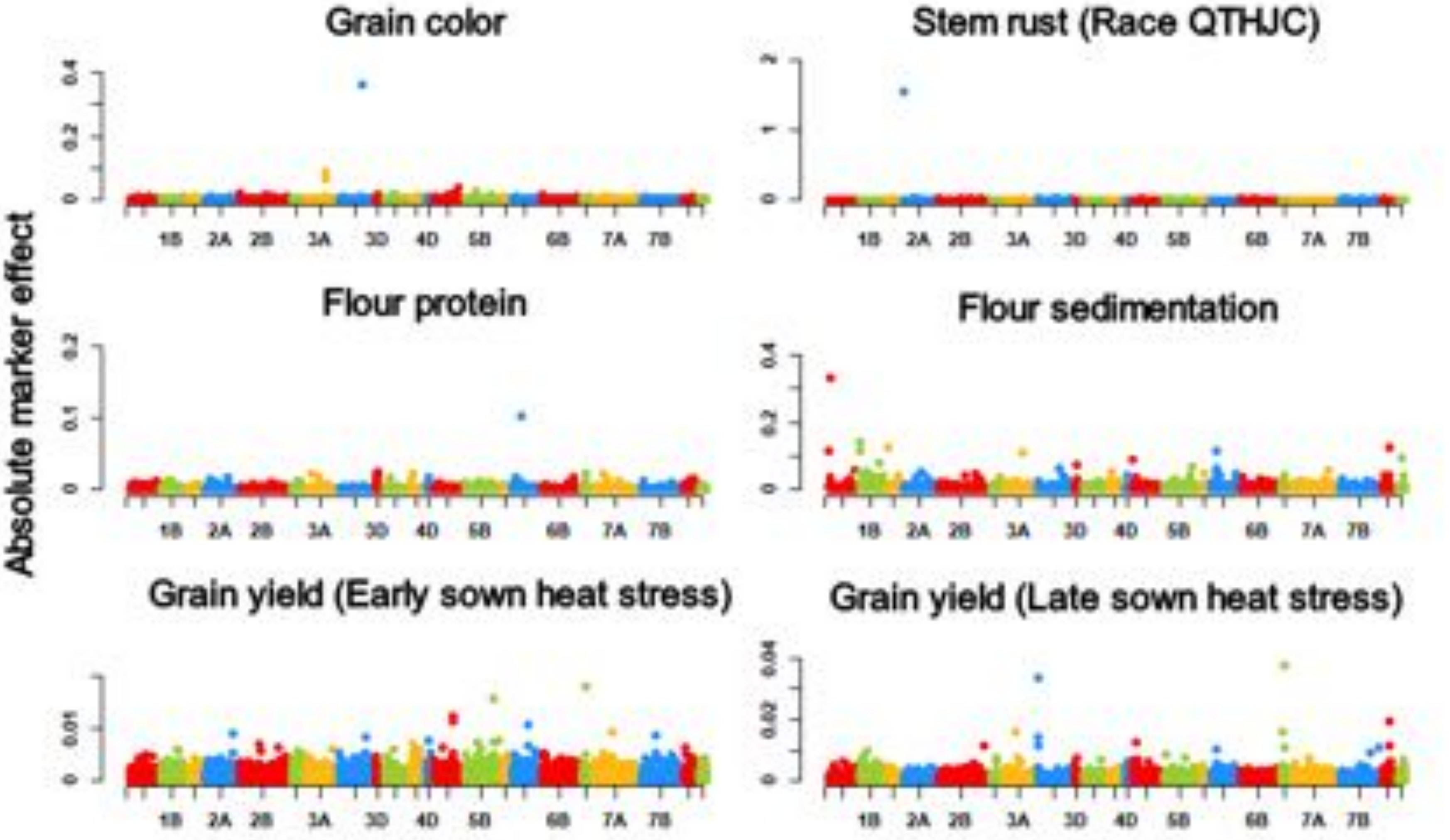
Forward genomic predictions across nurseries/environments using historic training sets



Mean genomic prediction accuracies of four different panels with 1,092 lines each evaluated during the 2014-2017 seasons in Obregon (the 1,092 lines were predicted using three other panels of 3,276 lines). DRT, drought stressed; EHT, early-sown heat stressed; IRR, irrigated; LHT, late-sown heat stressed; LUDH, Ludhiana; MEX, Mexico.

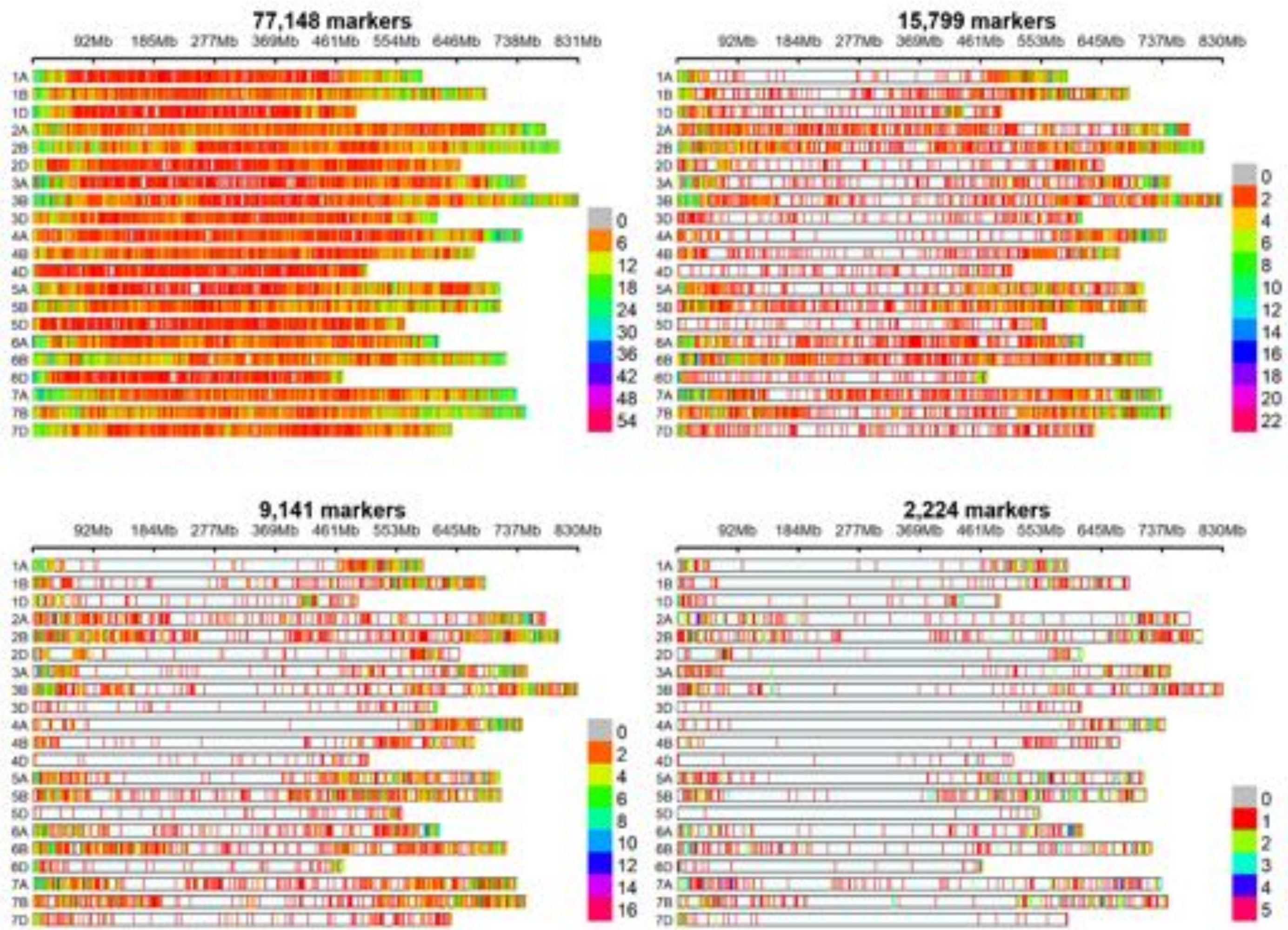
End-use quality and disease resistance are predicted well from historic training sets at CIMMYT, but grain yield predictions for unknown future environments are challenging due to the low heritability of yield across years and environments.

How did the RefSeq help with genomic predictions?



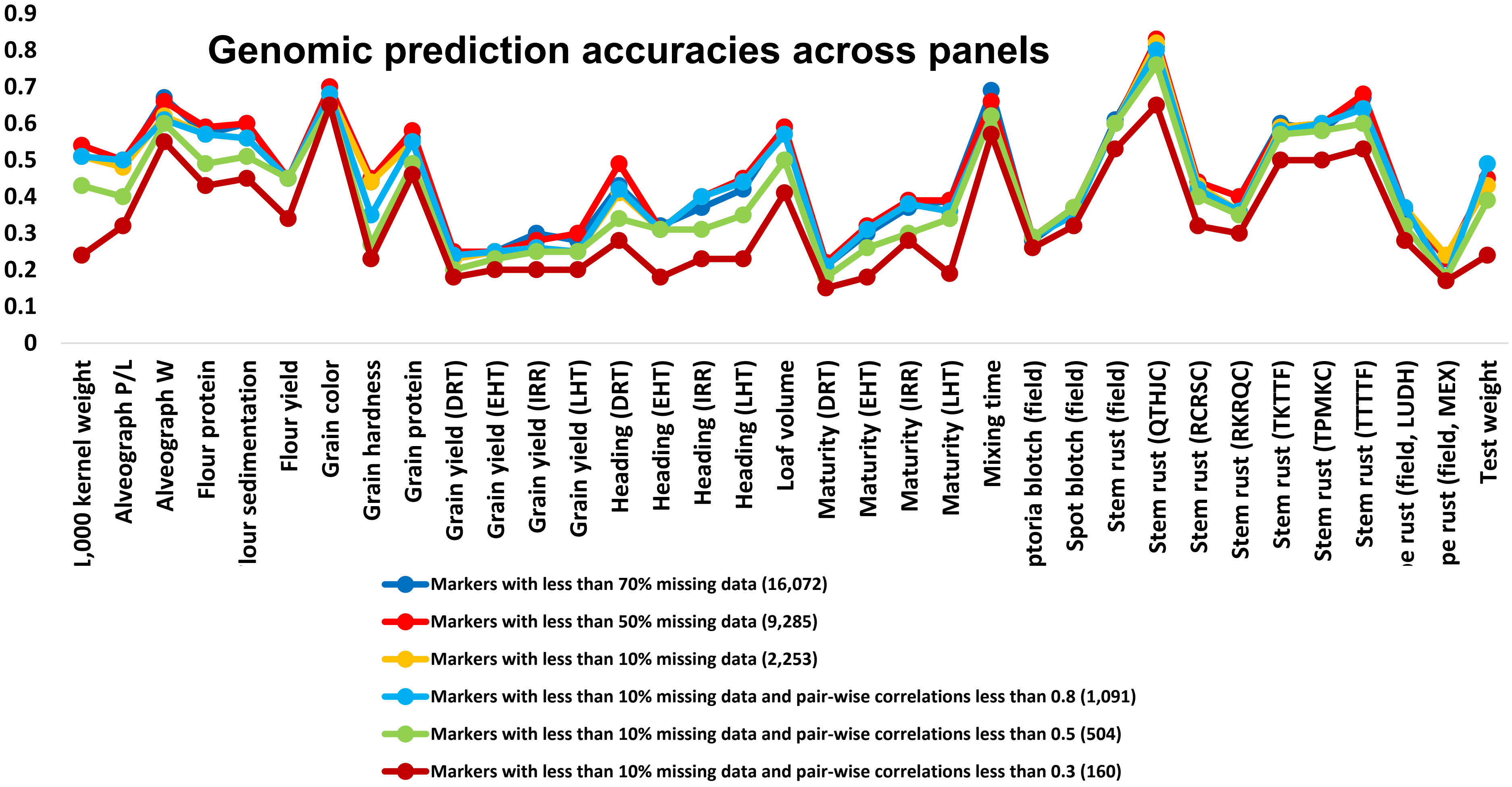
The RefSeq helped us better understand genomic prediction accuracies and knowledge of the loci that have large effects on traits gave us more confidence on trait predictabilities.

What is the effect of different marker densities on genomic prediction accuracies?



- The 77,148 markers denote the unfiltered set and subsets of 15,799, 9,141 and 2,224 markers denote filtered sets with 70%, 50% and 10% of missing data, respectively (the unaligned markers are not included). The color key with marker densities indicates the number of markers within a window size of 1 Mb.
- The genomic coverage associated with the marker subsets shows a decreasing trend toward the proximal centromeric regions with stringent filtering for missing data.

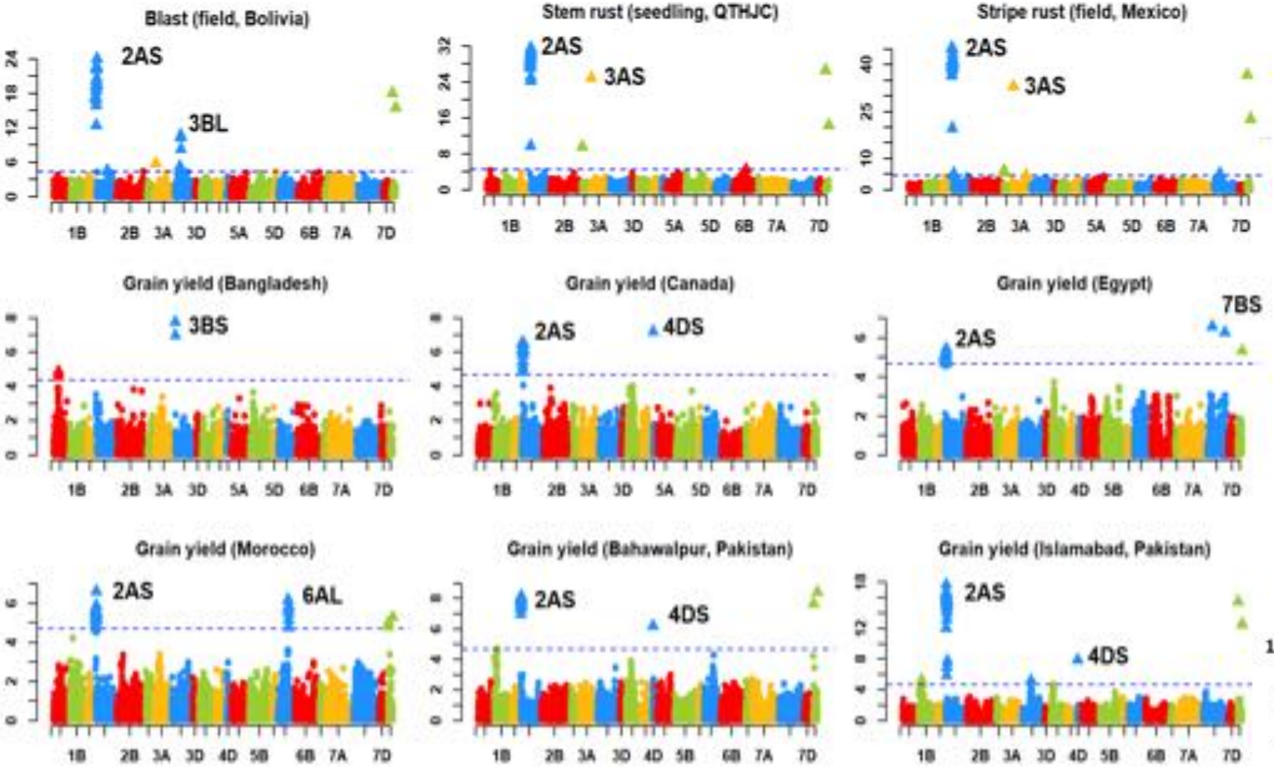
What is the effect of different marker densities on genomic prediction accuracies?


Genomic prediction accuracies across panels

Except for the smallest marker subset with 10% missing data and pairwise correlations less than 0.3, which led to an average decrease of 0.13 ± 0.06 in prediction accuracy across panels, all other marker subsets resulted in similar accuracies or had only a marginal difference in accuracies in both cross-validations and prediction across panels, across all traits.

DRT, drought stressed; EHT, early-sown heat stressed; IRR, irrigated; LHT, late-sown heat stressed; LUDH, Ludhiana; MEX, Mexico.

Large-scale genome-wide association mapping studies for identifying genomic regions associated with grain yield and key traits



Blast (field, Bolivia)

Stem rust (seedling, QTHJC)

Stripe rust (field, Mexico)

Grain yield (Bangladesh)

Grain yield (Canada)

Grain yield (Egypt)

Grain yield (Morocco)

Grain yield (Bahawalpur, Pakistan)

Grain yield (Islamabad, Pakistan)

- ☐ The 2NS translocation from *Aegilops ventricosa* was found to be associated with grain yield in several locations globally, stripe rust resistance (Mexico), seedling resistance to some stem rust races, lodging resistance and wheat blast resistance.

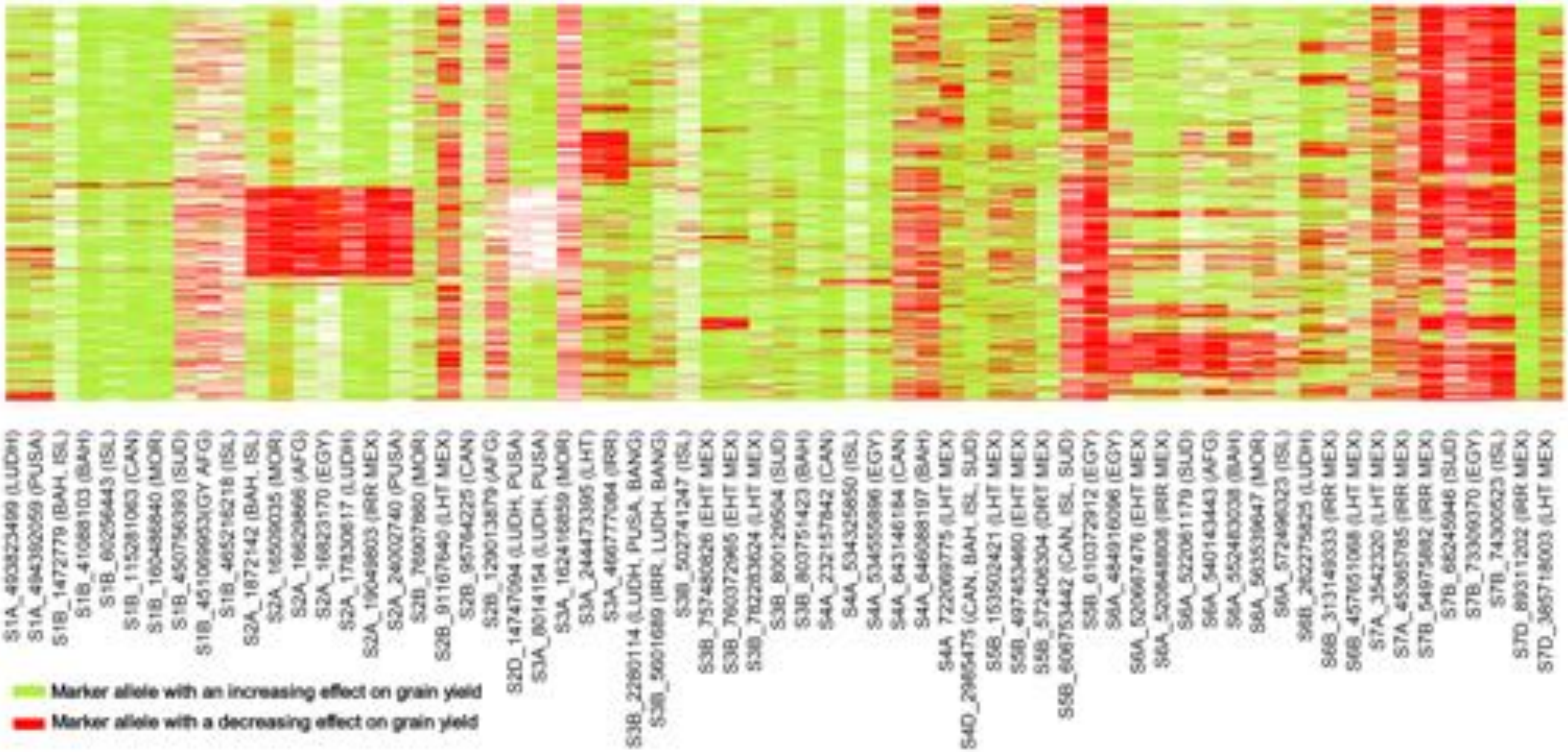
Trends in the 2NS translocation frequencies in CIMMYT's yield trial germplasm (>9,000 lines each year)

About 96% of >9,000 Stage 1 yield trial breeding lines in the 2018-2019 cycle had the 2NS translocation!

A reference genotype-phenotype map with key trait-linked markers aligned to the Reference Genome

Genotyping-by-sequencing markers associated with all the traits evaluated by the breeding program at CIMMYT have been aligned to the RefSeq v1.0.

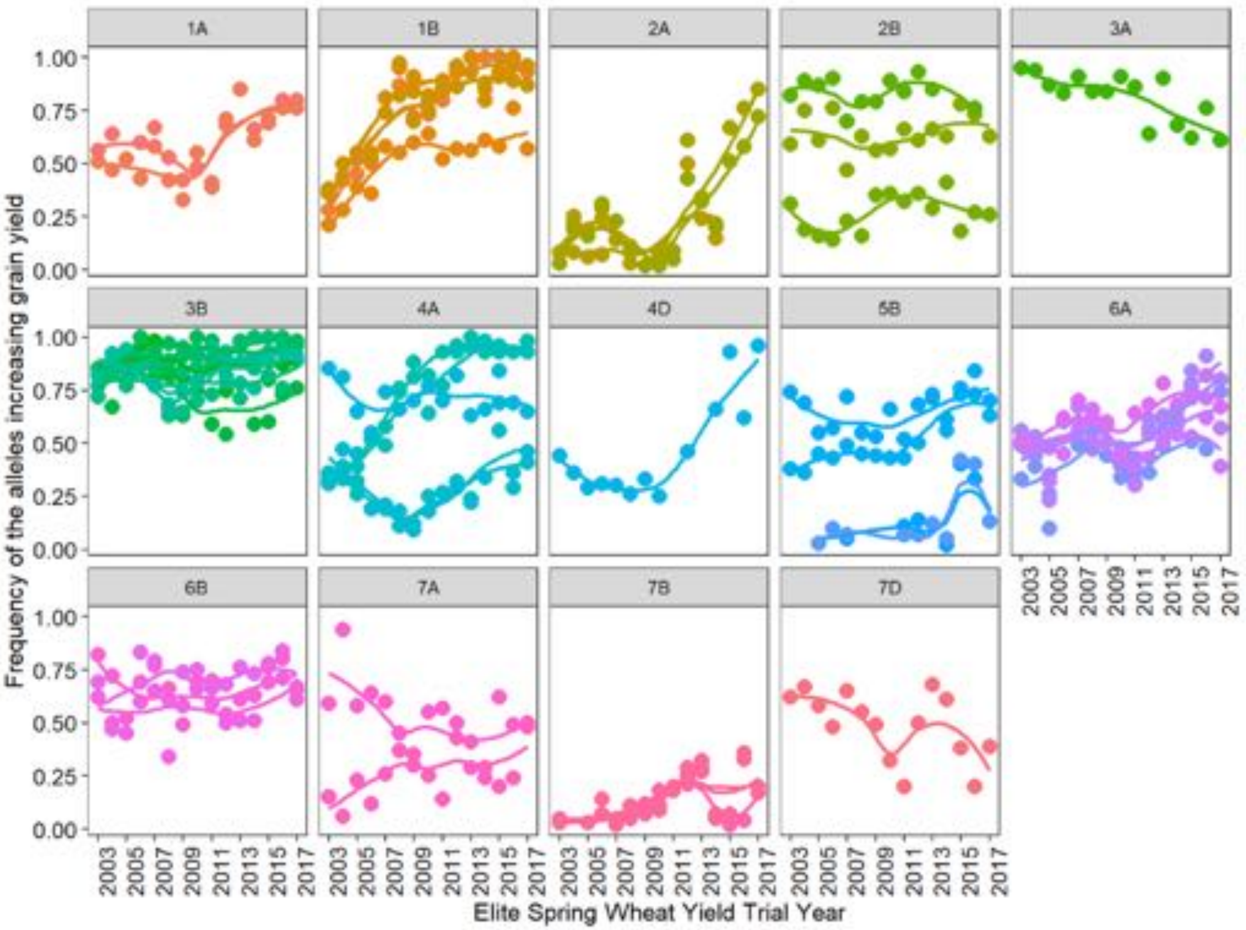
Genomic-fingerprints of CIMMYT’s global wheat breeding germplasm (44,624 lines) for key trait-linked markers



Genomic fingerprints for markers associated with grain yield evaluated in the irrigated (IRR) environments of Bahawalpur, Pakistan (BAH), Dehdadi Farms, Afghanistan (AFG), Jamalpur, Bangladesh (BANG), Swift Current, Canada (CAN), Ety El Barud, Egypt (EGY), Obregon, Mexico (IRR MEX), Islamabad, Pakistan (ISL), Ludhiana, India (LUDH), Marchouch, Morocco (MOR), Pusa, India (PUSA), Wad Medani, Sudan (SUD), and in the drought-stressed (DRT MEX), early-heat stressed (EHT MEX), and late-heat stressed (LHT MEX) environments of Obregon, Mexico.

The genomic fingerprints of 44,624 wheat lines has been generated for utilization in strategic and informed crossing decisions.

Molecular tracking of favorable allele frequencies and the effect of selection in 15 years of wheat breeding at CIMMYT



Trends in the grain yield favorable allele frequencies in the globally distributed Elite Spring Wheat Yield Trials

- ❑ The frequency dynamics of the favorable alleles for grain yield in the CIMMYT's Elite Spring Wheat Yield Trials (ESWYTs) due to selection for 15 years (2003–2017) was assessed using 47 markers associated with grain yield.
- ❑ The favorable alleles for several grain yield linked markers have reached near-fixation indicating the effectiveness of phenotypic selection at CIMMYT, and also emphasize the importance of integrating genomic data in achieving accelerated favorable allele enrichment.

Summary and acknowledgments

- ❑ The RefSeq has shifted the paradigm of genomics-assisted breeding in wheat, with numerous benefits to CIMMYT and its partners globally.
- ❑ The current challenges facing wheat production are immense. But, we can go right after those challenges and combat them with persevering audacity, motivated by a genuine passion to do something for the poorest of the world's poor!

IWGSC International Wheat Genome Sequencing Consortium

BILL & MELINDA GATES foundation

DFID Department for International Development

USAID FROM THE AMERICAN PEOPLE

KANSAS STATE UNIVERSITY

USDA

Colegio de Postgraduados

UNIVERSIDAD DE COLIMA Facultad de Telemática

INIAF Instituto Nacional de Innovación Agropecuaria y Forestal

Jesse Poland
Sandesh Shrestha
Daljit Singh
Mokhles Rahman

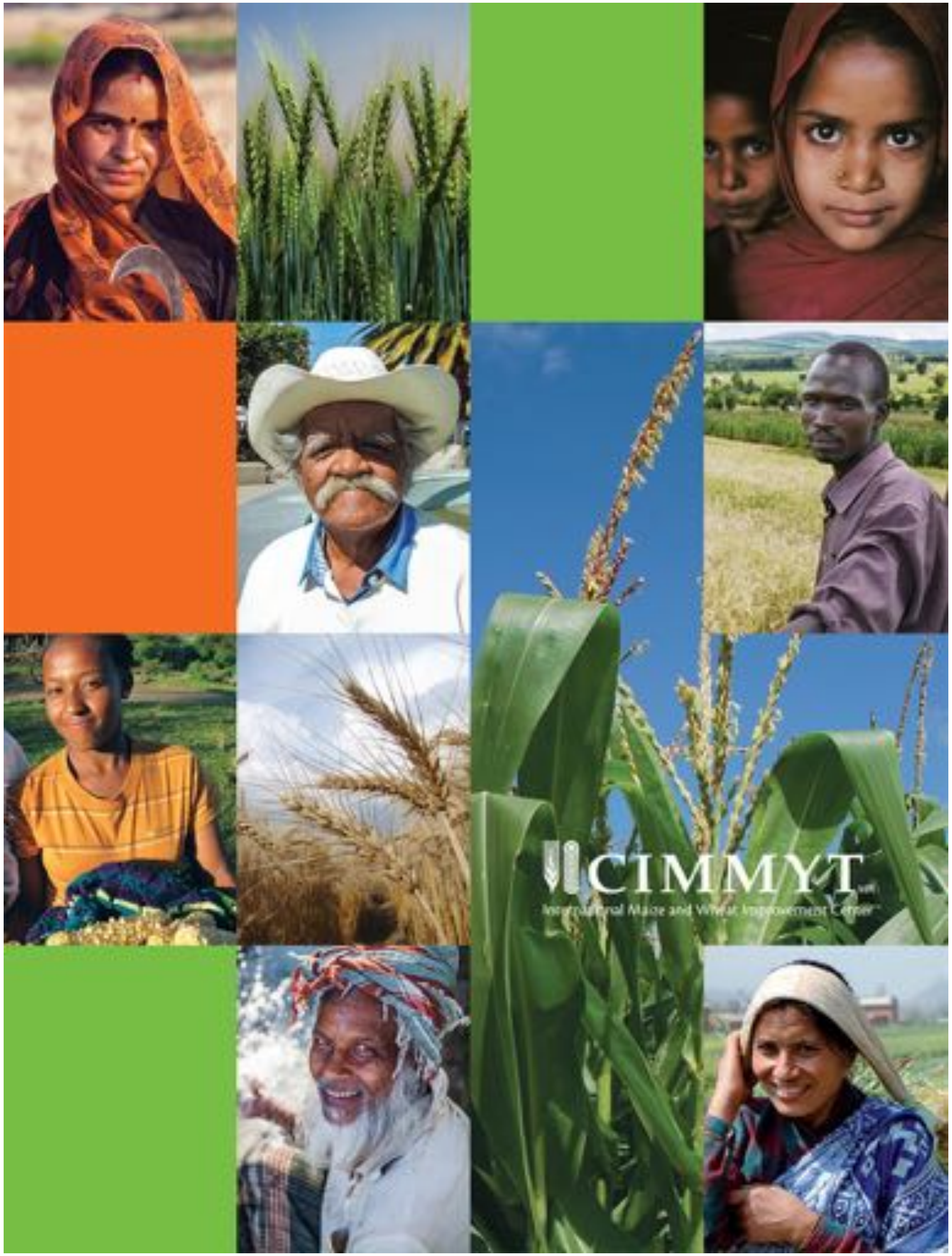
Matthew Rouse
Yue Jin

Paulino Pérez-Rodríguez

Osval A. Montesinos-López

Felix Marza

CIMMYT



Thank you for your interest!