Harnessing Global Diversity and Predictive Biology to Accelerate Sorghum Breeding

Invited presentation for the 26th EUCARPIA Maize and Sorghum Conference . Breeding in the era of genome editing and artificial intelligence. June 30th – July 2nd 2025, Bologna Italy

https://www.eucarpiamaizesorghum2025.com/

https://www.eucarpiamaizesorghum2025.com/keynote-speakers/

David Pot[1,2], Vincent Garin[1,2], Clément Bienvenu[1,2], Vincent Segura[1,3], Hugues De Verdal[1,2], Florian Larue[1,2,4], Gregory Beurier[1,2], Concetta Burgarella[1], Aude Gilabert[1,2], Angélique Berger[1,2],Korotimi Thera[5], Michel Vaksmann [1,2], Niaba Teme[5], Jean-François Rami[1,2]

1 AGAP Institut, Université Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France,
2 CIRAD, UMR AGAP Institut, Montpellier,France,
3 Geno-Vigne®, IFV-INRAE-Institut Agro, Montpellier, France
4 LEPSE, Université Montpellier, INRAE, Institut Agro, Montpellier, 34060, France
5 Institut d’Economie Rurale, Bamako, BP 262, Mali

In order to ensure food security, it is essential to adapt varieties and crop management systems to the challenges imposed by climate change. Two key components of the breeder’s equation are i) the genetic variance potentially mobilizable by breeding programs and ii) the accuracy with which the genetic values of candidates to selection can be predicted. To enable West and Central African sorghum breeding programs to utilize beneficial alleles present in locally non-adapted genetic materials, a large multireference backcross nested association mapping population was developed, based on three regional elite parents and 24 donor parents harboring specific beneficial traits. Each of the 3,901 families was characterized in three to four contrasting environments. A new QTL mapping approach was proposed, allowing the estimation of QTL sensitivity to environmental covariates. This framework allowed the potential effects of detected QTL alleles to be projected into unexplored climatic conditions. A subset of this population, comprising two recurrent and 22 donor parents (2,498 families), was then used to explore the relevance of phenomic selection compared to genomic prediction. Our results showed that phenomic selection can be used in sorghum, but its accuracy depends heavily on the considered traits. We also demonstrated that Near-Infrared Reflectance Spectroscopy (NIRS) data acquired in a reference environment can predict the performance of genotypes characterized in other environments. Furthermore, we suggest that the current methods employed by geneticists to take advantage of spectral information is likely suboptimal, as ten randomly selected wavelengths generally suffice to achieve thte same predictive capabilities as the full spectra.

In order to optimize the prediction of genetic values, we demonstrated that Convolutional Neural Networks (CNNs) outperformed traditional genomic prediction models when it came to predicting an integrative trait (fresh aboveground biomass) in a medium-sized association panel. We then explored the integration of genomic prediction with a Crop Growth Model (Ecomeristem), based on a two-step approach: first, estimating the crop model's genetic parameters, and second, using these parameters as inputs for the crop model to predict integrative traits. Our results reinforced the expectations raised by previous studies regarding the potential of integrating genomic prediction with crop growth models.

To unlock the ability to take advantage of the wild sorghum reservoir, we compared nucleotide and transcriptome wild diversity with the cultivated compartment, allowing us to identify potentially relevant alleles to reintroduce into breeding programs.

Maximizing genetic gains per unit of time is critical to meet the challenges posed by climate change. Enhancing each component of the breeder’s equation through coordinated efforts is therefore a strategic priority. In this context, fostering stronger trans-species collaborations between maize and sorghum research communities will be instrumental especially in improving the accuracy of genetic value estimation.

Key words: Multiparental population, BCNAM, GXE, Genomic Prediction, Phenomic selection, Crop growth Model, Crop wild relatives