

Improving Candidate Gene Discovery By Combining Multiple Genetic Mapping Datasets

Fausto Rodriguez-Zapata¹, Nirwan Tandukar¹, Jung-Ying Tzeng², **Rubén Rellán-Álvarez***¹
(rrellan@ncsu.edu)

¹Department of Molecular and Structural Biochemistry, ²Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC, 27607.

Project Goals:

- 1) Perform an environmental GWAS in a panel of ~2000 sorghum accessions that have already been genotyped and georeferenced using soil phosphorus availability as the phenotype
- 2) Identify selection signals at the gene and pathway level using natural populations that have adapted to low phosphorus and low temperature
- 3) Characterize the genetic architecture of glycerolipid content during the early stages of sorghum development using the Sorghum Association Panel.
- 4) Test if selection for faster development in low phosphorus availability conditions reduces genome size and more phosphorus efficient varieties.
- 5) Develop algorithms that incorporate all the different types of information we collect (i.e., metabolite levels, GWAS candidate genes, selection signals) to improve our ability to detect signals of minor effects and increase our confidence in selecting candidate genes for further validation.
- 6) Validate candidate gene(s) function under low phosphorus and cold conditions.

Deciphering the number and importance of loci involved in plant local adaptation and identifying candidate genes controlling traits that are relevant for adaptation to environmental stresses are overarching goals of plant scientists across a range of fields. From plant breeders aiming to introduce beneficial genetic diversity from landraces and wild relatives into high-yielding modern varieties to evolutionary biologists striving to identify which loci are shaped by evolutionary forces and molecular biologists and physiologists seeking to understand the role of genetic variation in plant development and stress responses.

However, validating gene function at the molecular and physiological level and confirming the importance of allelic variation for a particular gene are expensive and usually only possible when candidates can be limited to those with large effects on the trait of interest. Reducing the number of candidate genes and making an informed decision on which ones should be validated is particularly challenging in the case of highly polygenic traits.

In this project, we will leverage our *a priori* knowledge of the role of metabolic pathways in plant adaptation to environmental stresses to develop a combinatorial approach to combine data

from targeted metabolic profiling GWAS, environmental GWAS, and population genetics indexes of genetic selection. With these data, we will then design mathematical methods that incorporate pathway-level data and phenotype-genotype associations to develop strategies to reduce the number of candidate genes.

Our preliminary data and the literature show that phosphorus deficiency and low temperatures can have a profound impact on lipid metabolism, particularly on the levels of plant lipids, and that the combination of population and quantitative genetics together with high-precision metabolic phenotyping can be a powerful approach to identify relevant gene/pathway–phenotype associations and ultimately elucidate the functional role of genes/pathways in plant adaptation to environmental stresses.

Funding Statement: This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER) Early Career grant no. DE-SC0021889.