

Measurements of the studied phenotype of 274 *Sorghum* accessions ( $a$ ) in an environment

Normalized phenotype measurements using logistic growth curve model

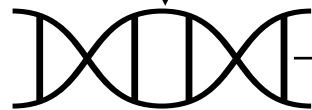
Create *phenotype similarity matrix* of dimension  $274 \times 274$ ; a symmetric matrix with cell values other than diagonal  $\in [0,1]$

	$a_1$	$a_2$	...	$a_{274}$
$a_1$	1			
$a_2$		1		
$\vdots$			.	
$a_{274}$				1

All variants (SNPs) after GATK Variant Filtration, considering 274 *Sorghum* accessions



Apply GWAS with a permissive-filtered threshold



GWAS-filtered ' $N$ ' number of SNPs

Create  $N$  no. *SNP similarity matrices* ( $S_1, S_2, \dots, S_N$ ), each with dimension  $274 \times 274$  and cell values other than diagonal  $\in \{0.0, 0.5, 1.0\}$

	$a_1$	$a_2$	...	$a_{274}$
$a_1$	1			
$a_2$		1		
$\vdots$			.	
$a_{274}$				1

$S_1$

...

	$a_1$	$a_2$	...	$a_{274}$
$a_1$	1			
$a_2$		1		
$\vdots$			.	
$a_{274}$				1

$S_N$

### Sequential SNP Prioritization Algorithm (SSPA)

(Selecting ' $K$ ' no. of *SNP similarity matrices*, the summation of which (i.e., final similarity matrix) results in maximum correlation coefficient with the *phenotype similarity matrix*)



Prioritized Set of SNPs

SNP Annotation using SnpEff package

Analysis of function, carried out on either Phytozome, Gramene, Planteome or the GO online databases

### Generalizability Test in a Different Environment

- Create *phenotype similarity matrix* based on phenotypic measurements of same 274 accessions grown in a different environment.
- Compute correlation coefficient of this *phenotype similarity matrix* with the *final similarity matrix* obtained for prioritized SNPs in the original environment.