Article

Not peer-reviewed version

Genome-Wide Analysis to Identify SNP Markers Associated with Plant Height in Sorghum bicolor

Diego Zamudio-Ayala *

Posted Date: 31 March 2025

doi: 10.20944/preprints202503.2343.v1

Keywords: BLINK; farmCPU; GWAS; sorghum

Article

Genome-Wide Analysis to Identify SNP Markers Associated with Plant Height in *Sorghum bicolor*

D. Zamudio

Abelson Hall Room 406C, School of Biological Sciences, Washington State University, Pullman, WA, USA;
diego.zamudioayala@wsu.edu

Abstract: Understanding the genetic basis of plant height (PH) in sorghum is essential for improving crop performance and breeding efficiency. This study analyzed 309 sorghum accessions using genome-wide association studies (GWAS) to identify significant single nucleotide polymorphism (SNP) markers associated with PH. Genotyping was conducted using SeqSNP with 5000 markers, of which 3143 high-quality SNPs were selected for analysis. Principal component analysis (PCA) and kinship analysis revealed two distinct genetic clusters. GWAS was performed using seven models: GLM, MLM, MLMM, CMLM, ECMLM, SUPER, FarmCPU, and BLINK. While GLM and SUPER identified a large number of significant SNPs, their Manhattan and QQ plots indicated high false-positive rates. FarmCPU and BLINK proved to be the most reliable models, detecting 10 and 8 significant SNPs, respectively, with four SNPs shared between both models. The most influential SNPs were located on chromosomes 1 and 8, in genes Sobic.001G017500 and Sobic.008G050800, known for their roles in plant growth and development. These findings demonstrate the advantage of multi-locus GWAS models in reducing false positives and enhancing SNP detection accuracy. The identified markers provide valuable insights for sorghum breeding programs, enabling the selection of desirable traits to optimize plant height and improve overall yield potential.

Keywords: BLINK; farmCPU; GWAS; sorghum

Introduction

Approximately 29% of the total population is moderately or severely food insecure (FAO, 2023). Even though there has been some progress towards reducing chronic hunger in the last four years in regions such as Asia and Latin America and the Caribbean, projected numbers of undernourished indicate that the world is far off from achieving zero hunger by 2030 (FAO et al., 2024). In addition, global food supply and demand are expected to notoriously increase, especially in developing countries, because the world population is expected to grow 1.6 billion people by 2050 (FAO, 2023). To meet these challenges, crop yield needs to increase by 70% by 2050 to satisfy the demands of a growing population (Hussain et al., 2021). Efforts to increase crop yield have included agronomical practices such as intercropping, breeding high-yield crops, better fertilizers and pesticides, and the development of stress-tolerant plants (Hussain et al., 2021). However, the implementation of these techniques is not sufficient to satisfy future food demands (Murchie et al., 2009).

While the majority of studies looking at boosting crop production have focused mainly on C3 species, the genetic factors controlling yield in C4 species are not well understood (Sales et al., 2021). C4 species such as maize, sorghum, and sugarcane are among the top ten crops with the highest annual global production, with sorghum being the fifth most important among cereal crops in terms of both annual metric production and yield (FAO, 2023). Sorghum (*Sorghum bicolor* L.) is known for its dual-purpose crop (staple food for humans and feed crop for livestock) and its greatest drought tolerance among C_4 plants (Hossain et al., 2022), vital to secure future food security. The small relative genome makes sorghum an ideal model crop for crop genomic studies. Moreover, the availability of the sorghum genome sequences in public databases has made it possible to conduct different analyses, including genome-wide-association (GWAS) analysis for the identification of molecular markers associated with key physiological traits.

Genome-wide association studies (GWAS) are known for their efficiency in identifying quantitative traits nucleotides (QTNs) associated with genomic regions of interest (CITATION). However, differentiating true from false QTNs, most often caused by population structure issues, has been challenging (CITATION). Given that, more powerful models have been developed in the last 10 years to overcome GWAS limitations. Among those new models, fixed and random model circulating probability unification (farmCPU) (Liu et al., 2016) and bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) (Huang et al., 2018) are the most powerful models for GWAS. The farmCPU model is a multi-locus linear mixed model (MLMM) that analyzes multiple markers at the same time as cofactors to partially eliminate the confounding effect of markers and kinship coefficients (Liu et al., 2016). As a result, farmCPU has higher statistical power and better computational power compared to other GWAS models. Although the farmCPU model is good for reducing confounding effects caused by multiple markers, it is still inefficient when a large number of markers and genotypes are involved in the analysis. In contrast, BLINK, which uses Bayesian methods, is more time-efficient due to it replaces random effects with a fixed effects model, and therefore, removes the assumption that casual genes are evenly distributed across the genome.

Although GWAS analysis often reveals novel genomic regions associated with important traits that have a high impact on crop productivity and resilience, the genetic variation within and among different plant accessions along with genotype-by-environment interactions (G x E) limits GWAS results robustness and validation, especially when comparing studies conducted under multi-environment field trials. In addition, the development of newer GWAS models with higher power statistical analysis and computational speed needs to be considered when selecting the right model for a given data set. Therefore, this study aimed to conduct GWAS analysis to identify significant SNP makers associated with plant height of 309 independent sorghum accessions. This study hypothesizes that using newer models with higher statistical power and computational speed will help to identify SNP markers associated with plant height in sorghum accessions.

Material and Methods

Plant Material

The number of sorghum accessions was obtained as outlined by Enyew et al. (2025), with the difference that only 309 accessions were used in this study.

Phenotyping Data

Plant height (cm; PH) was measured from each sorghum accession before the flowering stage.

SNP Selection and Genotyping

The 309 accessions were genotyped using SeqSNP. This genotyping method used 5000 SNP makers that were previously identified in the genetic diversity analysis of sorghum accessions by Enyew et al. (2022). The SNP markers are targeted on chromosome 1 to chromosome 10. The majority of the SNP markers (93.7%) came from the sorghum DNP database SorGSD (http://sorgsd.big.ac.cn) and the remaining (6.3%) were identified as outlined by Enyew et al. (2022). Among the 5000 SNP markers, the polymorphic SNP loci and bi-allelic loi were identified as described by Enyew et al. (2025). In addition, from the bi-allelic SNP loci group, further SNP filtering analysis through minimum allele frequency (MAF) was performed to identify high-quality SNP loci (3142 SNP).

Genome-Wide Association (GWAS)

All data were analyzed using R software version 4.2.2. GWAS analysis was performed using the statistical package Genome Association and Prediction Integrated Tool (GAPIT version: 3.0). The GWAS used a total of 3143 SNP markers with phenotype data (plant height, PH) from 309 sorghum accessions. Principal component analysis (PCA) and pairwise genetic association (kinship matrix) were calculated as outlined by Enyew et al. (2022). Five different GWAS models, GLM (general linear model), MLM (mixed linear model), MLMM (multiple loci mixed linear model), CMLM (compressed MLM), SUPER, ECMLM (enriched CMLM), farmCPU (Fixed and random model circulating probability unification), and BLINK (Bayesian-information and linkage-disequilibrium iteratively nested keyway) were used to determine significant SNPs for PH. Significant associations between PH and SNPs were determined at $P < 0.05$ with Bonferroni correction. Manhattan and QQ plots were generated with GAPIT functions in the R software. To assess the statistical power among the GWAS models used in this experiment. FarmCPU and BLINK along with earlier models GLM, MLM, MLMM, CMLM, ECMLM, and SUPER were compared using the GAPIT.Power.compare test (number of replicates = 10) from the GAPIT function (Wang & Zhang, 2021).

Results and Discussion

Principal Components (PCA)

The principal components analysis (PCA) of the 309 sorghum accessions was determined by the package GAPIT. The number of principal components that better explained the population variability was established by the eigenvalues (Figure 1A). Based on these values, 3 principal components (PC) were used to determine the population stratification of the 309 sorghum accessions (Figure 1B). Because PC3 contributed less than 5% of the population variability, only 2 PCs were selected to depict the clustering pattern of the 309 sorghum accessions (Figure 1C). The kinship analysis also showed that sorghum accessions were clustered into 2 groups (Figure 2).

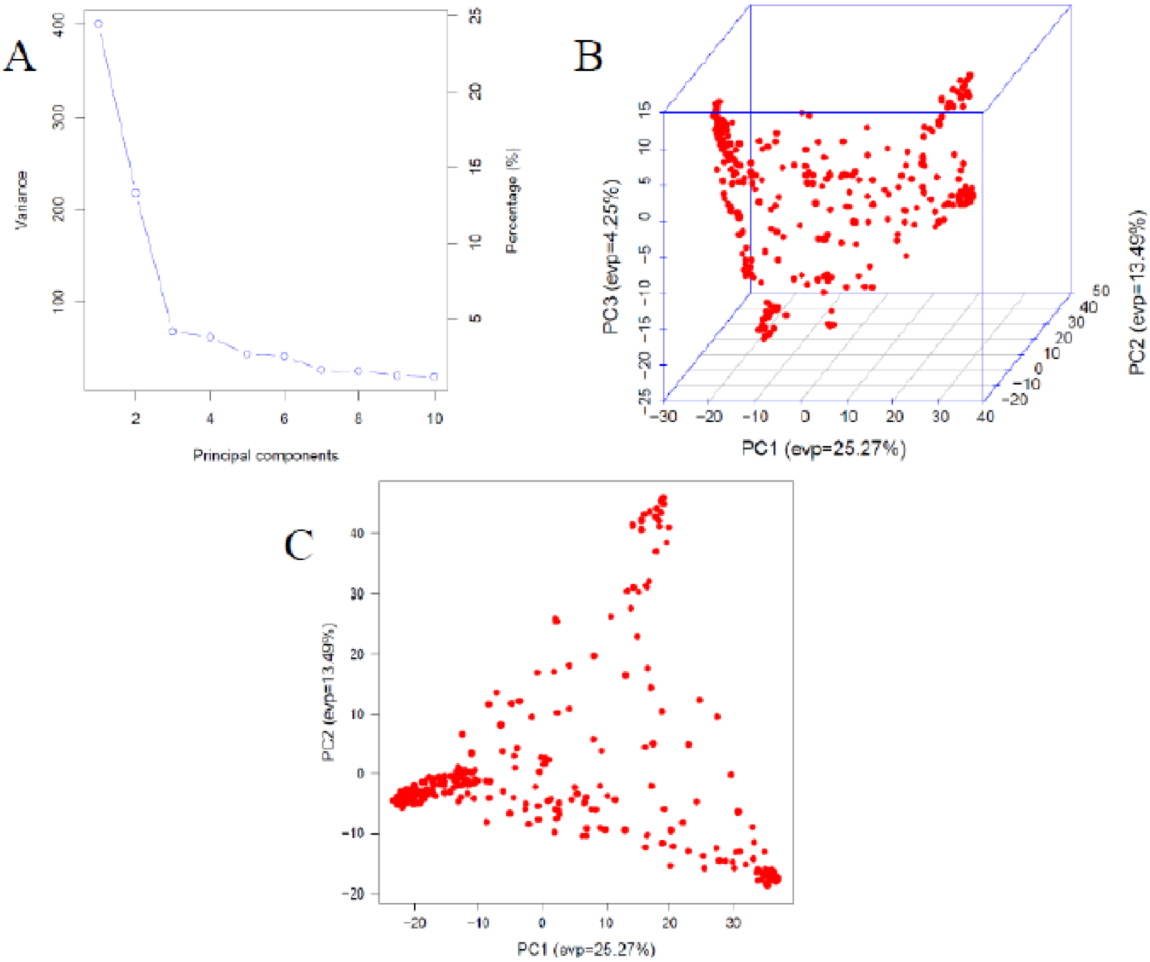

Figure 1. Eigenvalues that show the percentage (%) of variability explained by the most significant components and PCA scatter plots with 3 and 2 components depicting the clustering pattern of the 309 sorghum accessions.

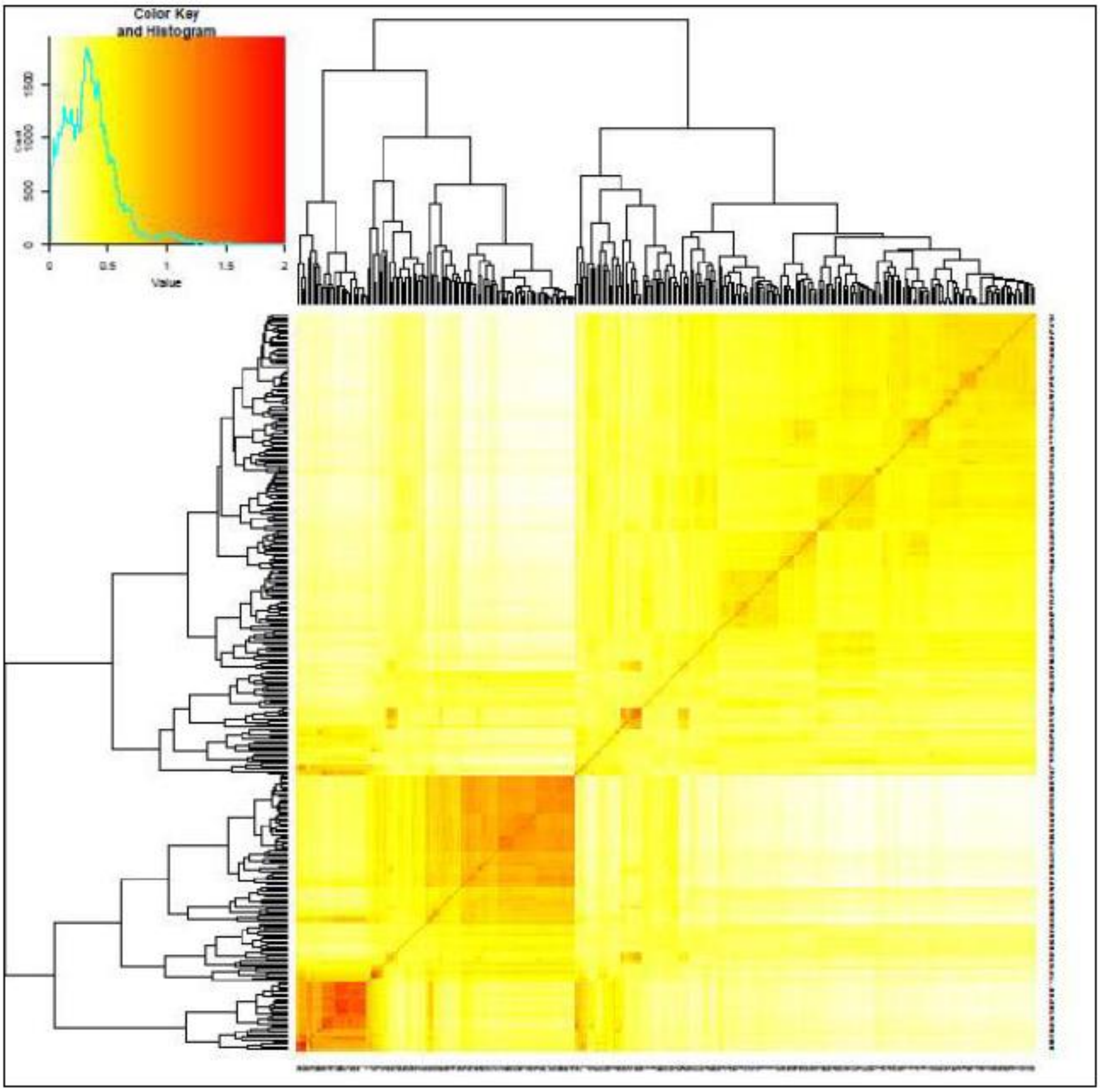

Figure 2. Heat map diagram of kinship matrix depicting that the 309 sorghum accessions are grouped into two clustering groups.

Markers-Trait Associated with Plant Height (PH)

Table 1 shows the significant markers associated with PH. The GLM and SUPER model identified 78 and 124 significant SNP markers, however, based on its QQ plot, the observed vs. predicted values were not correlated, indicating that the majority of significant SNPs identified through GLM were false positives.

FarmCPU and BLINK identified 10 and 8 significant SNP markers, respectively associated with the trait plant height, whereas the other two models MLM and MLMM did not identify any significant SNP markers associated with PH (Figure 3). Among the significant SNP markers for PH, four SNPs were identified by both farmCPU and BLINK. The Manhattan and QQ plots of FarmCPU and BLINK showed that the observed vs. the predicted *P-values* for the 3143 SNPs are matching, therefore, false positive results are less likely (Figure 3). The use of multi-locus GWAS models is better to overcome the limitations associated with single-locus models (Li et al., 2015) because they include kinship values and principal components to avoid confounding effects caused by population structure (Enyew et al., 2022). Moreover, the statistical power analysis showed that only farmCPU and BLINK reduced the likelihood of identifying false positives compared to GLM, MLM, MLMM, CMLM, and SUPER (Figures 4A and 4B). The comparison between multi-locus GWAS models and single-locus models has been also assessed to identify SNP markers associated with plant trains in the vegetative stage of different sorghum accessions (Zou et al., 2024) suggesting that BLINK and farmCPU are models with improved statistical power. Similarly, a study assessing the disease rating for resistance to grain mold in 635 sorghum accessions found that the use of multiple models including those that avoid overfitting and eliminate confounding effects between kinship and testing markers (Nida et al., 2021). Even though these studies have assessed different GWAS models to identify SNP markers in different maturity stages of sorghum accessions, farmCPU and BLINK were still the most powerful models with less likelihood to identify false positives.

Both farmCPU and BLINK identified four markers in chromosomes 3, 4, 6, and 8 (Figure 5). The phenotypic variation explained by these markers ranged between 0.11 (chromosome 3) to 12.6% (chromosome 8). Additionally, four SNP markers were identified by BLINK on chromosomes 1, 5, and 7 with 2.38, 8.19, and 15.3% of phenotypic variation, respectively. Likewise, farmCPU identified 6 additional SNP markers on chromosomes 1, 3, 6, 9, and 10 with 27.4, 8.23, 0.65, 0, and 0.96% of phenotypic variation, respectively. The SNP markers that had the largest effect on PH, in chromosome 1 (PVE = 27.4%) and 8 (PVE=12.6%), were located in the genes Sobic.001G017500 and Sobic.008G050800, respectively. The former encodes for hydroxysteroid dehydrogenase, oxidoreductase, and short-chain dehydrogenase regions, while the latter plays an important role in protein expression. These genes are known for their importance in plant vegetative and reproductive growth (Enyew et al., 2022).

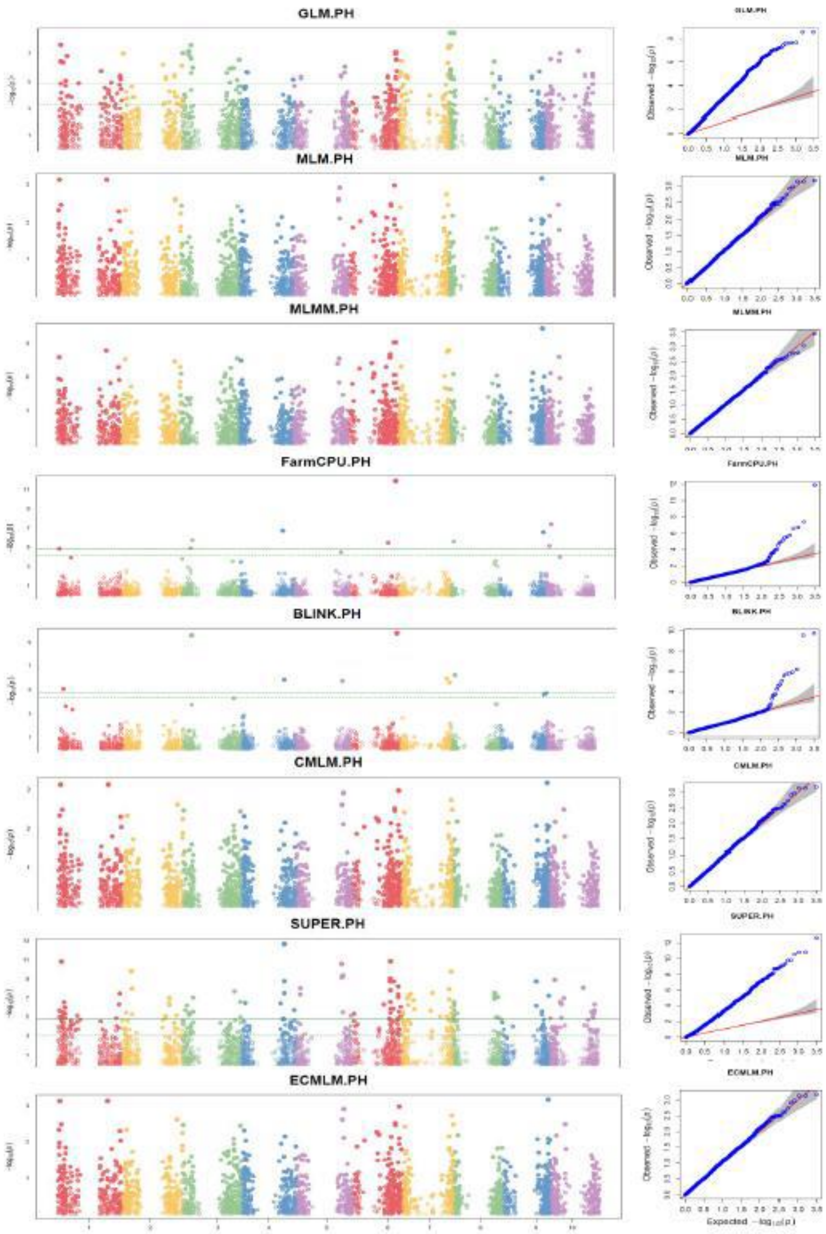

Figure 3. Manhattan and WW plots from GLM (general linear model), MLM (mixed linear model), MLMM (multiple loci mixed linear model), CMLM (compressed MLM), SUPER, ECMLM (enriched CMLM), farmCPU (Fixed and random model circulating probability unification), and BLINK (Bayesian-information and linkage-disequilibrium iteratively nested keyway) GWAS analysis of 309 sorghum accessions using 3143 SNP markers (*P values* threshold were corrected by Bonferroni correction).

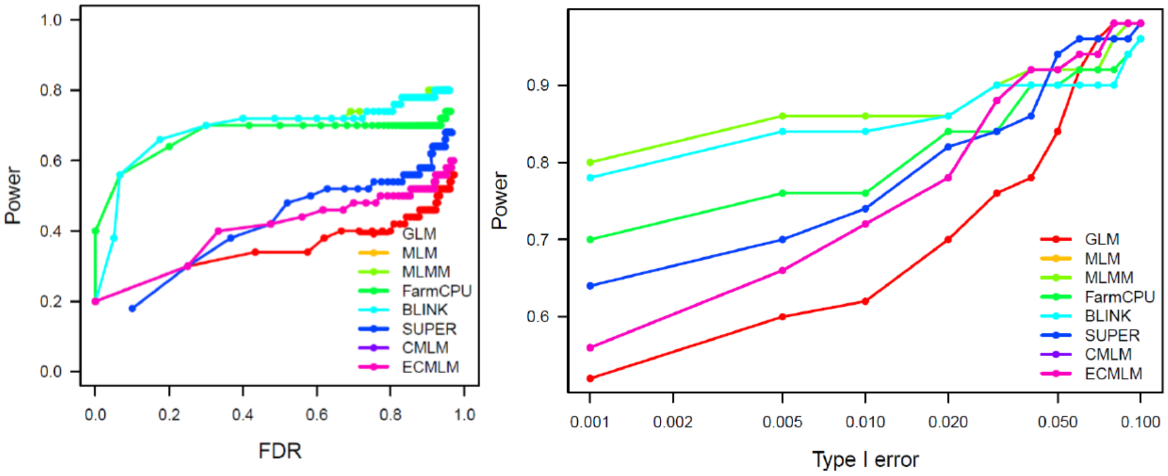

Figure 4. Statistical power comparison between GLM (general linear model), MLM (mixed linear model), MLMM (multiple loci mixed linear model), CMLM (compressed MLM), SUPER, ECMLM (enriched CMLM), farmCPU (Fixed and random model circulating probability unification), and BLINK (Bayesian-information and linkage-disequilibrium iteratively nested keyway) with simulated phenotype data from 309 sorghum accessions.

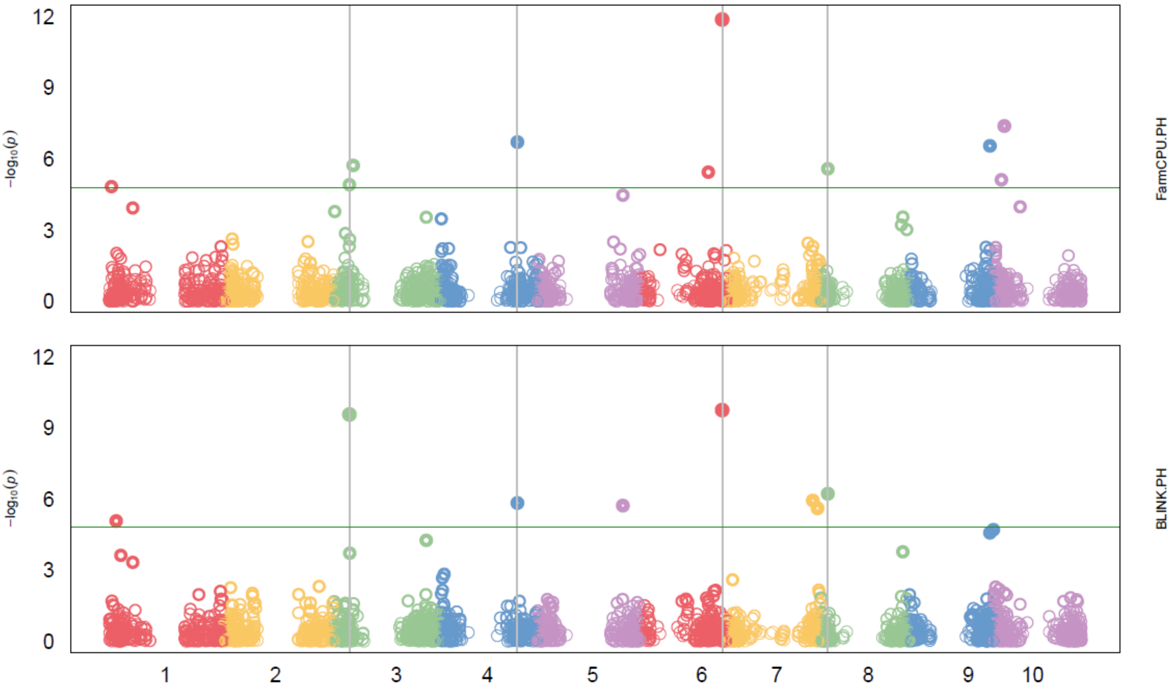

Figure 5. Manhattan plots from farmCPU (Fixed and random model circulating probability unification), and BLINK (Bayesian-information and linkage-disequilibrium iteratively nested keyway) GWAS analysis to identify the number of SNP markers identified by both models. .

Table 1. List of significant SNP markers associated with plant height (PH) of 309 sorghum accessions from GWAS analysis.

Trait	GWAS model	SNP ID	Chromosome number	Position	MAF¹	*P-value*	PVE (%)²

PH[3]	GLM	snp_sb001000003033	1	1803311	0.11650485	2.10E-08	0
		snp_sb001000003043	1	1803552	0.06796117	3.59E-07	1.17
		snp_sb042060173374	1	2053226	0.20711974	1.30E-05	3.06
		snp_sb042060176498	1	4583054	0.12297735	4.58E-07	3.30
		snp_sb042060176733	1	4778713	0.40453074	4.12E-07	0
		snp_sb042060183185	1	8066218	0.26213592	7.89E-06	4.04
		snp_sb040057635750	1	9985873	0.0987055	1.46E-07	7.01
		snp_sb001000040602	1	24769165	0.38673139	1.23E-05	0
		snp_sb042060233820	1	53456591	0.44012945	2.01E-06	0
		snp_sb042060260112	1	66258089	0.25080906	1.09E-05	0.75
		snp_sb042060273650	1	77744466	0.06148867	4.51E-06	0
		snp_sb042060274847	1	78464288	0.32524272	6.09E-06	0
		snp_sb042060279054	2	1028568	0.18770227	9.40E-08	0
		snp_sb042060334327	2	51764695	0.15048544	6.59E-07	3.89
		snp_sb042060339135	2	55665363	0.34951456	5.71E-06	0.51
		snp_sb042060355039	2	66604865	0.12459547	3.85E-06	0
		snp_sb042060363157	2	75994759	0.12621359	7.87E-07	0
		snp_sb042060365193	3	542076	0.26213592	8.95E-08	1.09
		snp_sb042060375446	3	8002244	0.1828479	6.13E-08	3.06
		snp_sb042060375491	3	8012102	0.22168285	1.16E-07	0.42
		snp_sb042060379604	3	10734118	0.06148867	2.10E-07	0
		snp_sb042060379940	3	10836535	0.0776699	2.23E-08	0
		snp_sb042060380150	3	10921954	0.08090615	2.30E-07	0
		snp_sb042060380199	3	10932720	0.07281553	2.71E-06	5.54
		snp_sb042060384040	3	12866372	0.24433657	1.10E-06	0.21
		snp_sb042060421455	3	56322246	0.28964401	2.65E-06	0

snp_sb042060428418	3	60472707	0.0566343	1.07E-05	1.77
snp_sb042060428585	3	60501386	0.05339806	1.29E-06	0
snp_sb042060429232	3	60811994	0.0566343	1.24E-06	0
snp_sb042060434280	3	64830003	0.06148867	1.19E-06	0
snp_sb042060439477	3	69764576	0.2038835	6.96E-06	0
snp_sb042060445645	3	72284237	0.05825243	2.79E-07	0
snp_sb042060463777	4	7424990	0.06148867	2.38E-06	0
snp_sb042060526464	4	8890484	0.08252427	2.51E-06	1.30
snp_sb042060523231	4	66858877	0.25889968	6.67E-06	0.11
snp_sb042060545830	5	8563133	0.05177994	5.31E-06	0
snp_sb042060589862	5	59844243	0.41585761	4.13E-06	4.45
snp_sb042060600784	5	64447085	0.24919094	2.52E-06	0
snp_sb042060601544	5	64669481	0.06634304	9.37E-07	0
snp_sb001000702141	6	47599757	0.35598706	4.80E-06	0
snp_sb001000703068	6	47897247	0.42394822	1.42E-05	0.68
snp_sb001000719123	6	57647043	0.39320388	7.47E-08	0
snp_sb001000719127	6	57647287	0.3802589	1.01E-07	0.43
snp_sb001000719254	6	57704110	0.31067961	5.42E-06	4.23
snp_sb001000720111	6	58537465	0.20226537	2.75E-07	0
snp_sb001000721097	6	59276261	0.1407767	2.04E-06	0
snp_sb042060760081	7	2622240	0.10032362	2.83E-06	0
snp_sb042060760098	7	2623348	0.10355987	2.29E-06	0
snp_sb042060760099	7	2623635	0.0987055	3.48E-06	0
snp_sb042060761490	7	3776173	0.07119741	3.33E-07	7.40
snp_sb042060766358	7	6455441	0.29449838	1.57E-05	0
snp_sb042060772026	7	9428572	0.16504854	4.27E-06	0

snp_sb042060774985	7	12018772	0.43042071	5.26E-06	0
snp_sb042060850373	7	63227425	0.45631068	3.42E-06	0.50
snp_sb042060850458	7	63300346	0.12783172	4.78E-07	5.09
snp_sb042060852727	7	64318593	0.22330097	3.69E-08	0
snp_sb042060852735	7	64319205	0.22491909	2.58E-08	0
snp_sb042060856436	8	606304	0.197411	2.44E-08	0.13
snp_sb042060857050	8	1016470	0.27346278	2.70E-09	0.27
snp_sb042060857649	8	1241472	0.09223301	3.50E-06	0
snp_sb042060857701	8	1243863	0.08090615	1.82E-06	3.31
snp_sb042060858244	8	1459101	0.36893204	1.40E-06	0.16
snp_sb042060858246	8	1462792	0.36731392	5.57E-06	0
snp_sb042060867162	8	5053406	0.08576052	6.16E-06	1.54
snp_sb042060867731	8	5247147	0.23300971	2.68E-09	0
snp_sb042060939820	8	54615437	0.19417476	2.82E-06	2.26
snp_sb042060945149	8	56847745	0.09385113	3.71E-07	0
snp_sb042060958548	9	2005306	0.15372168	5.86E-06	0
snp_sb042061005134	9	42849238	0.08252427	5.11E-06	0
snp_sb042061021148	9	52373447	0.2605178	2.16E-06	0
snp_sb001001049696	10	7303004	0.05987055	2.49E-07	8.78
snp_sb001001075487	10	18388160	0.23300971	3.25E-07	0
snp_sb042061062036	10	19451315	0.19417476	9.11E-08	3.57
snp_sb001001125741	10	42872791	0.40291262	5.98E-08	0
snp_sb001001143490	10	51455048	0.2394822	3.66E-06	1.23
snp_sb001001159813	10	60977250	0.4012945	1.39E-05	0.12
snp_sb001001159934	10	61069500	0.13915858	2.99E-06	0
snp_sb001001160062	10	61144250	0.20064725	3.98E-06	0

SUPER	snp_sb042060173350	1	2036942	0.15534	1.10E-05	-
	snp_sb042060173374	1	2053226	0.20712	2.30E-06	-
	snp_sb042060176733	1	4778713	0.404531	1.60E-11	-
	snp_sb042060177238	1	5073279	0.493528	1.14E-05	-
	snp_sb042060177747	1	5705669	0.461165	9.13E-06	-
	snp_sb042060177833	1	5736900	0.307443	3.18E-06	-
	snp_sb042060183033	1	8034619	0.210356	3.49E-06	-
	snp_sb042060183185	1	8066218	0.262136	2.99E-07	-
	snp_sb001000014533	1	8870319	0.423948	9.51E-07	-
	snp_sb042060184103	1	9022825	0.315534	5.03E-06	-
	snp_sb042060184571	1	9291528	0.457929	1.08E-06	-
	snp_sb001000015314	1	9412963	0.449838	1.40E-05	-
	snp_sb001000040602	1	24769165	0.386731	7.69E-06	-
	snp_sb042060267164	1	73737253	0.349515	1.36E-05	-
	snp_sb042060273269	1	77432019	0.199029	4.72E-07	-
	snp_sb042060274847	1	78464288	0.325243	3.62E-08	-
	snp_sb042060287884	2	5838445	0.459547	3.69E-06	-
	snp_sb042060302008	2	12243572	0.331715	1.62E-10	-
	snp_sb042060304434	2	13177574	0.18123	5.33E-06	-
	snp_sb042060304638	2	13293723	0.427184	1.07E-08	-
	snp_sb042060338476	2	55292382	0.31877	1.32E-05	-
	snp_sb042060339135	2	55665363	0.349515	8.25E-06	-
	snp_sb042060347981	2	59810106	0.480583	3.55E-07	-
	snp_sb042060348862	2	60374738	0.451456	6.24E-07	-
	snp_sb042060349202	2	60415546	0.470874	9.69E-08	-
	snp_sb042060349216	2	60416638	0.438511	1.16E-07	-

snp_sb042060358990	2	69586319	0.072816	6.97E-06	-
snp_sb042060368691	3	3720735	0.364078	2.09E-06	-
snp_sb042060370968	3	5535752	0.375405	6.31E-07	-
snp_sb042060375446	3	8002244	0.182848	3.91E-06	-
snp_sb042060375491	3	8012102	0.221683	6.55E-06	-
snp_sb042060378852	3	10370193	0.122977	3.87E-06	-
snp_sb042060384040	3	12866372	0.244337	3.43E-06	-
snp_sb042060385209	3	13494213	0.380259	2.88E-06	-
snp_sb042060385211	3	13494361	0.496764	9.35E-08	-
snp_sb042060434769	3	65706545	0.472492	2.15E-08	-
snp_sb042060446486	3	72742644	0.343042	2.67E-07	-
snp_sb042060447019	3	73010484	0.187702	1.08E-05	-
snp_sb042060523179	4	4755657	0.158576	3.40E-07	-
snp_sb042060527001	4	49335014	0.302589	3.70E-07	-
snp_sb042060525591	4	52459901	0.478964	1.44E-05	-
snp_sb042060524523	4	54044421	0.446602	2.22E-13	-
snp_sb001000469722	4	54108806	0.18932	3.37E-08	-
snp_sb042060524643	4	54165068	0.383495	1.76E-09	-
snp_sb042060526825	4	66343002	0.496764	2.07E-06	-
snp_sb042060539991	5	5752947	0.453074	4.32E-08	-
snp_sb042060541082	5	6039745	0.346278	9.84E-09	-
snp_sb042060587348	5	57995949	0.443366	2.81E-11	-
snp_sb042060587478	5	58078499	0.365696	6.82E-10	-
snp_sb042060589677	5	59776476	0.482201	3.34E-06	-
snp_sb042060589862	5	59844243	0.415858	4.54E-10	-
snp_sb042060613793	5	69766819	0.299353	7.77E-06	-

snp_sb001000591841	6	1893544	0.067961	4.97E-06	-
snp_sb001000685308	6	39650537	0.433657	4.84E-07	-
snp_sb001000699642	6	46593261	0.067961	1.52E-05	-
snp_sb001000702141	6	47599757	0.355987	9.67E-10	-
snp_sb001000702204	6	47614312	0.415858	1.83E-09	-
snp_sb001000702655	6	47747134	0.389968	2.63E-07	-
snp_sb001000703068	6	47897247	0.423948	1.61E-08	-
snp_sb001000703373	6	48007508	0.126214	3.89E-06	-
snp_sb001000703397	6	48019015	0.386731	1.50E-11	-
snp_sb001000703435	6	48022337	0.326861	1.39E-06	-
snp_sb001000710036	6	51604058	0.391586	1.49E-06	-
snp_sb001000710785	6	51877933	0.404531	1.99E-09	-
snp_sb001000719123	6	57647043	0.393204	4.25E-08	-
snp_sb001000719127	6	57647287	0.380259	8.49E-08	-
snp_sb001000719254	6	57704110	0.31068	6.61E-09	-
snp_sb001000719606	6	57947690	0.393204	1.40E-05	-
snp_sb001000720111	6	58537465	0.202265	2.82E-06	-
snp_sb001000722757	6	60369078	0.050162	9.03E-06	-
snp_sb001000723739	6	60908842	0.300971	8.88E-07	-
snp_sb042060760081	7	2622240	0.100324	3.49E-06	-
snp_sb042060760098	7	2623348	0.10356	5.02E-06	-
snp_sb042060760099	7	2623635	0.098706	4.78E-06	-
snp_sb042060770742	7	8646573	0.360841	8.44E-08	-
snp_sb042060770748	7	8647644	0.415858	1.17E-07	-
snp_sb042060771257	7	8852228	0.346278	5.59E-07	-
snp_sb042060774985	7	12018772	0.430421	3.83E-08	-

snp_sb042060777297	7	15926543	0.456311	1.70E-07	-
snp_sb042060778451	7	17139586	0.381877	1.57E-06	-
snp_sb042060810749	7	37587848	0.286408	2.36E-06	-
snp_sb042060814271	7	39762751	0.441748	3.17E-08	-
snp_sb042060829083	7	49513148	0.300971	9.34E-06	-
snp_sb042060840008	7	57778908	0.425566	9.31E-06	-
snp_sb042060844978	7	60737379	0.331715	1.46E-07	-
snp_sb042060844980	7	60737490	0.33657	1.26E-08	-
snp_sb042060850373	7	63227425	0.456311	1.77E-10	-
snp_sb042060850888	7	63387510	0.482201	4.16E-06	-
snp_sb042060851559	7	63687599	0.242718	9.21E-07	-
snp_sb042060852735	7	64319205	0.224919	1.08E-05	-
snp_sb042060856436	8	606304	0.197411	1.09E-05	-
snp_sb042060857701	8	1243863	0.080906	6.59E-06	-
snp_sb042060857985	8	1398622	0.228155	2.77E-07	-
snp_sb042060858064	8	1417507	0.436893	6.98E-07	-
snp_sb042060858244	8	1459101	0.368932	3.51E-07	-
snp_sb042060858246	8	1462792	0.367314	2.15E-06	-
snp_sb042060867731	8	5247147	0.23301	4.64E-06	-
snp_sb042060935546	8	52669272	0.377023	1.27E-07	-
snp_sb042060935548	8	52669338	0.36246	6.20E-08	-
snp_sb042060935700	8	52708389	0.349515	3.18E-08	-
snp_sb042060939820	8	54615437	0.194175	1.13E-05	-
snp_sb042060944584	8	56617256	0.440129	8.10E-08	-
snp_sb042060945149	8	56847745	0.093851	1.14E-05	-
snp_sb042061005177	9	42889324	0.360841	1.89E-09	-

	snp_sb042061005325	9	42941172	0.430421	6.97E-07	-
	snp_sb042061016556	9	50101247	0.378641	2.71E-06	-
	snp_sb042061016557	9	50101376	0.372168	5.49E-06	-
	snp_sb042061032708	9	58880811	0.428803	6.21E-08	-
	snp_sb001001039940	10	1485692	0.320388	1.38E-05	-
	snp_sb001001039971	10	1488770	0.338188	1.46E-05	-
	snp_sb001001049162	10	7201807	0.359223	1.23E-09	-
	snp_sb001001066189	10	14137762	0.44822	2.53E-07	-
	snp_sb001001067710	10	15183661	0.47411	4.61E-07	-
	snp_sb001001068022	10	15305891	0.378641	1.23E-05	-
	snp_sb001001075487	10	18388160	0.23301	1.03E-07	-
	snp_sb001001125741	10	42872791	0.402913	8.92E-09	-
	snp_sb042061092070	10	53108112	0.050162	3.14E-06	-
	snp_sb001001150146	10	54629603	0.401295	1.02E-05	-
	snp_sb042061095323	10	55477852	0.338188	2.49E-06	-
	snp_sb001001151401	10	55478741	0.305825	1.33E-05	-
	snp_sb001001151403	10	55479226	0.309061	5.54E-06	-
	snp_sb001001153208	10	56339534	0.092233	4.37E-06	-
	snp_sb001001154375	10	56953779	0.490291	4.74E-07	-
MLM	-[4]	-	-	-	-	-
MLMM	-[5]	-	-	-	-	-
CMLM	-[6]	-	-	-	-	-
ECMLM	-[7]	-	-	-	-	-
farmCPU	snp_sb001000002532	1	1495213	0.113269	1.45E-05	27.4
	snp_sb042060379940	3	10836535	0.07767	1.20E-05	0.11
	snp_sb042060385209	3	13494213	0.380259	1.85E-06	8.23
	snp_sb001000469722	4	54108806	0.18932	1.91E-07	0
	snp_sb001000703373	6	48007508	0.126214	3.56E-06	0.65

	snp_sb001000719254	6	57704110	0.31068	1.29E-12	5.50
	snp_sb042060867162	8	5053406	0.085761	2.56E-06	0
	snp_sb042061028714	9	56601353	0.192557	2.76E-07	0
	snp_sb001001045903	10	5270775	0.11165	7.46E-06	0
	snp_sb001001049696	10	7303004	0.059871	4.02E-08	0.96
BLINK	snp_sb042060176733	1	4778713	0.404531	8.40E-06	2.38
	snp_sb042060379940	3	10836535	0.07767	2.70E-10	10.4
	snp_sb001000469722	4	54108806	0.18932	1.47E-06	7.34
	snp_sb042060589862	5	59844243	0.415858	1.89E-06	8.19
	snp_sb001000719254	6	57704110	0.31068	1.77E-10	5.38
	snp_sb042060843522	7	59997937	0.119741	1.15E-06	15.3
	snp_sb042060850458	7	63300346	0.127832	2.46E-06	5.29
	snp_sb042060867162	8	5053406	0.085761	5.97E-07	12.6

[1] MAF, minor allele frequency (>0.05).
[2] PVE, phenotypic variation (%).
[3] PH, plant height (cm).
[4,5,6,7] Non-significant SNP markers were identified by MLM, MLMM, CMLM, and ECMLM.
[a] SNPs in bold were identified by two GWAS models (farmCPU and BLINK).

Conclusions

This study highlights the effectiveness of BLINK and farmCPU to identify significant SNP markers associated with plant height in 309 sorghum accessions, indicating that these models can be implemented for further GWAS analysis with other phenotypic traits. Both FarmCPU and BLINK identified four common SNP markers on chromosomes 3, 4, 6, and 8, with additional unique markers detected by each model. Notably, the SNPs with the largest effects on plant height were located in genes Sobic.001G017500 and Sobic.008G050800, which are involved in key metabolic and growth-related processes. Overall, these findings reinforce the importance of selecting robust GWAS models for trait association studies and provide valuable genetic insights that can aid in sorghum breeding programs aimed at improving plant height and overall crop performance.

References

1. Enyew, M., Feyissa, T., Carlsson, A. S., Tesfaye, K., Hammenhag, C., Seyoum, A., & Geleta, M. (2022). Genome-wide analyses using multi-locus models revealed marker-trait associations for major agronomic traits in Sorghum bicolor. *Frontiers in Plant Science*, *13*, 999692-999692. https://doi.org/10.3389/fpls.2022.999692
2. Enyew, M., Geleta, M., Tesfaye, K., Seyoum, A., Feyissa, T., Alemu, A., Hammenhag, C., & Carlsson, A. S. (2025). Genome-wide association study and genomic prediction of root system architecture traits in Sorghum (Sorghum bicolor (L.) Moench) at the seedling stage. *BMC plant biology*, *25*(1), 69-69. https://doi.org/10.1186/s12870-025-06077-w

3. FAO. (2023). *FAOSTAT statistical database. Rome, Italy*. Accessed on August 5. https://www.fao.org/faostat/en/#data/FS.
4. FAO, IFAD, UNICEF, WFP, & WHO. (2024). The State of Food Security and Nutrition in the World 2024 – Financing to end hunger, food insecurity and malnutrition in all its forms. https://doi.org/10.4060/cd1254en
5. Hossain, M. S., Islam, M. N., Rahman, M. M., Mostofa, M. G., & Khan, M. A. R. (2022). Sorghum: A prospective crop for climatic vulnerability, food and nutritional security. *Journal of Agriculture and Food Research*, *8*, 100300. https://doi.org/https://doi.org/10.1016/j.jafr.2022.100300
6. Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2018). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience*, *8*(2). https://doi.org/10.1093/gigascience/giy154
7. Hussain, S., Ulhassan, Z., Brestic, M., Zivcak, M., Weijun, Z., Allakhverdiev, S. I., Yang, X., Safdar, M. E., Yang, W., & Liu, W. (2021). Photosynthesis research under climate change. *Photosynthesis Research*, *150*(1-3), 5-19. https://doi.org/10.1007/s11120-021-00861-z
8. Li, X., Li, X., Fridman, E., Tesso, T. T., & Yu, J. (2015). Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. *Proceedings of the National Academy of Sciences - PNAS*, *112*(38), 11823-11828. https://doi.org/10.1073/pnas.1509229112
9. Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies: e1005767. *PLoS genetics*, *12*(2). https://doi.org/10.1371/journal.pgen.1005767
10. Murchie, E. H., Pinto, M., & Horton, P. (2009). Agriculture and the new challenges for photosynthesis research. *The New phytologist*, *181*(3), 532-552. https://doi.org/10.1111/j.1469-8137.2008.02705.x
11. Nida, H., Girma, G., Mekonen, M., Tirfessa, A., Seyoum, A., Bejiga, T., Birhanu, C., Dessalegn, K., Senbetay, T., Ayana, G., Tesso, T., Ejeta, G., & Mengiste, T. (2021). Genome-wide association analysis reveals seed protein loci as determinants of variations in grain mold resistance in sorghum. *Theoretical and Applied Genetics*, *134*(4), 1167-1184. https://doi.org/10.1007/s00122-020-03762-2
12. Sales, C. R. G., Wang, Y., Evers, J. B., & Kromdijk, J. (2021). Improving C4 photosynthesis to increase productivity under optimal and suboptimal conditions. *Journal of Experimental Botany*, *72*(17), 5942-5960. https://doi.org/10.1093/jxb/erab327
13. Wang, J., & Zhang, Z. (2021). GAPIT Version 3: Boosting Power and Accuracy for Genomic Association and Prediction. *Genomics, proteomics & bioinformatics*, *19*(4), 629-640. https://doi.org/10.1016/j.gpb.2021.08.005
14. Zou, G., Ding, Y., Xu, J., Feng, Z., Cao, N., Chen, H., Liu, H., Zheng, X., Liu, X., & Zhang, L. (2024). Genome-wide dissection of genes shaping inflorescence morphology in 242 Chinese south–north sorghum accessions. *Scientific Reports*, *14*(1), 25828. https://doi.org/10.1038/s41598-024-76568-7