Pak. J. Bot., 58(8) August 2026.

ARTICLE | DOI: http://dx.doi.org/10.30848/PJB2026-8(3)

IDENTIFICATION OF QTLS ASSOCIATED WITH PANICLE LENGTH IN SORGHUM BASED ON BSA-SEQ

LIN LIN YANG, YAN QIU WANG, ZE YANG ZHAO, HAN WU, YOU HOU DUAN, FEI ZHANG , FU LAI KE AND KUANG YE ZHANG*

Sorghum Research Institute, Liaoning Academy of Agricultural Sciences, Shenyang 110161, China

*Corresponding author's email: 973219040@qq.com

Abstract

Panicle length represents a key yield-related trait in sorghum (*Sorghum bicolor* L. Moench). Elucidating the genetic basis of panicle length and identifying key regulatory genes are of great significance for high-yield molecular breeding in sorghum. In this study, two sorghum maintainer lines with significantly different panicle lengths, 01-26B and 7009B, were used as parents to construct an F_2 segregating population, and phenotypic investigation and genetic analysis of panicle length were conducted. The results showed that panicle length in the F_2 population exhibited a continuous distribution with obvious transgressive segregation, indicating that this trait is controlled by multiple genes and conforms to the characteristics of quantitative inheritance. Based on individuals with extreme panicle length phenotypes in the F_2 population, long-panicle and short-panicle bulks were constructed, and bulked segregant analysis sequencing (BSA-seq) was performed on the two parents and the bulks. At the 95% confidence level, a total of 14 key SNP loci significantly associated with panicle length were identified, which were distributed on chromosomes 3, 6, 7, 9, and 10. Among them, a nonsynonymous mutation located at 8,814,444 bp on chromosome 7 was identified in the exon region of the gene *LOC8071181*, which encodes an ECERIFERUM 26-like protein involved in the elongation of very long-chain fatty acids. It is speculated that this gene may indirectly regulate panicle rachis development by affecting wax biosynthesis or cell membrane stability. This study provides a theoretical basis and valuable genetic resources for elucidating the genetic mechanism of panicle length and for marker-assisted breeding in sorghum.

Key words: Sorghum, Panicle length, BSA-seq, ECERIFERUM 26-like protein

Introduction

Sorghum (*Sorghum bicolor* L. Moench) is the fifth most important cereal crop worldwide and serves multiple purposes, including food, feed, brewing, and bioenergy production (Zheng *et al*., 2023; Li *et al*., 2024). It possesses remarkable biological advantages such as drought tolerance, salt–alkali tolerance, tolerance to poor soils, and high photosynthetic efficiency compared with many C_3 crops (Zhang *et al*., 2016a; Hussain *et al*., 2024). As a typical stress-tolerant C4 crop, sorghum has considerable agricultural value in the context of marginal land utilization and global climate change. In recent years, with the rapid development of the brewing industry in China and the increasing demand for stress-resilient agriculture, the importance of improving sorghum yield and quality has become increasingly prominent. Therefore, in-depth analysis of the genetic basis of key agronomic traits is of great significance for breeding high-yielding, high-quality, and stable sorghum cultivars.

Panicle length, as a core component of inflorescence architecture, is significantly positively correlated with yield-related traits such as grain number per panicle and thousand-grain weight, and is thus considered a key determinant of crop yield (Li *et al*., 2019; Ding *et al*., 2024; Huang *et al*., 2025; Sun *et al*., 2025). Studies have shown that wheat genotypes with longer panicles often possess a greater number of spikelets, and the increase in spikelet number further enhances grain number per panicle, thereby contributing to yield improvement (Wu *et al*., 2019). In rice, increased panicle length favors the formation of large-panicle types, which is mainly manifested by an increase in primary and secondary branches as well as higher grain density (Agata *et al*., 2020; Dong *et al*., 2010). Wang *et al*., (2019) conducted a path analysis of six panicle-related traits and grain yield per plant across nine maize cultivars and found that panicle length exhibited the strongest correlation with yield and made the greatest direct contribution. Similarly, Zhang *et al*., (2016b) applied grey relational analysis to 14 sorghum hybrids and demonstrated that panicle length showed the highest correlation coefficient with yield, suggesting that long-panicle genotypes have substantial potential in sorghum yield improvement breeding. Therefore, elucidating the genetic basis of panicle length and identifying key regulatory genes is of great practical value for high-yield molecular breeding in sorghum.

Quantitative trait locus (QTL) mapping is an efficient approach for identifying genes underlying complex traits, primarily by analyzing the relationship between genome-wide DNA molecular markers and quantitative phenotypic variation to determine the chromosomal locations of QTLs. Traditional genetic mapping methods usually require genotyping all individuals in a population using a large number of molecular markers, which is time-consuming and labor-intensive (Zhang *et al*., 2021). In contrast, bulked

Received: 11-11-2025 | Revised: 04-03-2026 | Accepted: 09-03-2026 | Online: 01-04-2026

segregant analysis combined with high-throughput sequencing (BSA-seq) offers advantages such as high mapping resolution, rapid analysis, and simplified workflows, and has been widely applied in crop genetic and breeding research. This method involves sequencing DNA from two bulks composed of individuals with extreme phenotypes for a target trait, along with the parental lines, and then calculating allele frequencies of single nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) as well as their differences between the two bulks to rapidly identify genomic regions harboring target genes or QTLs (Zou *et al*., 2016; Ma *et al*., 2024).

This study focused on sorghum maintainer lines exhibiting significant differences in panicle length and the F_2 segregating population derived from them, aiming to identify panicle length–related QTLs through BSA-seq, screen key SNPs, and perform functional annotation and structural analysis of candidate genes. The results are expected to provide a genetic basis for elucidating the molecular mechanisms underlying panicle length formation in sorghum and to offer valuable candidate loci and gene resources for marker-assisted selection and high-yield sorghum breeding.

Materials and Methods

Plant materials: The long-panicle sorghum maintainer line 01-26B was used as the female parent, and the short-panicle sorghum maintainer line 7009B was used as the male parent. Both parental lines were developed by the Liaoning Academy of Agricultural Sciences and possess favorable agronomic traits, including drought tolerance, disease resistance, and stay-green characteristics. Hybridization was performed by artificial emasculation and pollination to obtain F_1 seeds, and the F_1 plants were self-pollinated to generate an F_2 segregating population. The parental lines and the F_2 population consisting of 217 individuals were planted at the experimental field of the Sorghum Research Institute, Liaoning Academy of Agricultural Sciences (41°49′ N, 123°33′ E). Each parental line was planted in two rows, while each F_2 family was planted in a single row, with a row length of 3 m and a row spacing of 0.60 m. Field management followed standard agronomic practices.

Phenotypic evaluation and bulk construction: At the maturity stage, panicle length was measured for all individual plants of the parental lines and the F_2 population. Panicle length was defined as the distance from the base of the panicle to the tip. Based on the panicle length measurements of the F_2 population, 50 individuals with extremely long panicles and 50 individuals with extremely short panicles were selected. Equal amounts of young leaf tissue were collected from these selected individuals as well as from the parental lines. Genomic DNA was extracted from each sample using DNAsecure New-type Plant Genomic DNA Extraction Kit DP320 (Tiangen, Beijing, China) concentration and purity were assessed using a NanoDrop 2000 microspectrophotometer. Genomic DNA from individuals with extreme phenotypes was pooled in equal amounts to construct the long-panicle bulk (PL-L) and the short-panicle bulk (PL-S).

Library construction and high-throughput sequencing: Genomic DNA samples of the two parental lines and the two extreme bulks that passed quality assessment were used for library construction. Libraries were prepared according to the TruSeq DNA Library Construction Kit protocol, including DNA fragmentation, end repair and A-tailing, adapter ligation, and PCR amplification. The constructed libraries were evaluated for concentration and insert size distribution using a Qubit fluorometer and an Agilent 2100 Bioanalyzer. Qualified libraries were sequenced on an Illumina HiSeq 2000 platform (San Diego, CA, USA) using the paired-end 150 bp (PE150) mode to generate raw sequencing data for subsequent BSA analysis.

Sequencing data processing: Raw sequencing reads were subjected to quality control using Fastp software to remove low-quality reads, adapter-contaminated reads, and reads with a high proportion of ambiguous bases (N). Clean reads from the parental lines and extreme bulks were aligned to the sorghum reference genome (BTx623T2T-CAS) using BWA. SAMtools was used for format conversion and duplicate removal (Li *et al*., 2009). SNP and InDel variants were subsequently detected using GATK software (McKenna *et al*., 2010). Functional annotation of the filtered SNPs was performed using ANNOVAR (Wang *et al*., 2010). Using the parental genotype as the reference, allele frequencies at each variant site were calculated for PL-L and PL-S, and the SNP-index was determined. To reduce the effects of sequencing errors and mapping bias, SNPs with SNP-index values lower than 0.3 in both bulks, sequencing depth lower than 7 in both bulks, or missing SNP-index values in either bulk were excluded. After quality filtering, the difference in SNP frequency between PL-L and PL-S was calculated as ΔSNP-index = SNP-index of PL-L − SNP-index of PL-S. A statistical distribution was generated using 1,000 permutation tests, and the 95% confidence interval was used as the threshold for identifying candidate associated regions (Yang *et al*., 2024).

Candidate gene identification and functional annotation: Based on the functional annotation of SNPs within candidate regions, priority was given to variants that potentially result in start/stop codon gain or loss, nonsynonymous substitutions, alternative splicing alterations, and frameshift mutations. Genes harboring these functional variants were defined as candidate genes. To further characterize the molecular features of candidate genes, DNA and amino acid sequences were retrieved from the NCBI database. Physicochemical properties of the candidate proteins, including amino acid composition, molecular weight, theoretical isoelectric point, and instability index, were analyzed using the ExPASy online tool (https://www.expasy.org/). Conserved domains and functional structural features of the candidate proteins were predicted using the SMART online platform (https://smart.embl.de/).

Statistical analysis: Data were organized using Microsoft Excel 2019, and statistical analyses of phenotypic data were performed using SPSS version 24. Descriptive statistics were used to summarize panicle length data. The normality of panicle length distribution in the F_2 population was evaluated using the Shapiro–Wilk test. Differences in panicle length between the two parental lines were assessed using Student's t-test. The significance level was set at $p<0.05$. Graphs and data visualization were generated using GraphPad Prism 7.

Results and analyses

Phenotypic analysis of panicle length in the tested sorghum materials: Phenotypic analysis of panicle length was conducted for the parental lines and their F_2 population. The results showed a significant difference in panicle length between the two parental lines (Fig. 1). The average panicle length of the maintainer line 01-26B (P1) was 28.8 cm, whereas that of 7009B (P2) was 23.8 cm, and the difference between the two parents reached an extremely significant level. In the F_2 population derived from the cross between the two parental lines, the mean panicle length was 21.39 cm, which was lower than that of both parents. Panicle length in the F_2 population ranged from 9 to 34 cm, with a coefficient of variation of 18.56%, indicating substantial phenotypic variation and the presence of transgressive segregation. The Shapiro–Wilk test showed that the distribution of panicle length in the F_2 population deviated from normality (Table 1; Fig. 2).

Sequencing data quality analysis: Sequencing quality statistics were analyzed for the parental lines 01-26B and 7009B and for the extreme bulks PL-L and PL-S, as shown in Table 2. The two parental lines generated 64,771,676 and 56,932,543 reads, with total base numbers of 9.72 Gb and 8.54 Gb, respectively, while the progeny bulks yielded 70,472,498 and 81,140,005 reads, corresponding to total base numbers of 10.57 Gb and 12.17 Gb, respectively. The sequencing error rates of all samples were below 0.026%, the Q20 values ranged from 97.52% to 97.82%, and the Q30 values ranged from 93.47% to 94.12%, indicating high sequencing quality; meanwhile, the GC content of all samples ranged from 43.32% to 44.36% and showed a normal distribution without obvious GC bias, demonstrating the reliability of the library construction and sequencing processes.

Sequencing data alignment analysis: Sequencing reads from all samples were aligned to the sorghum reference genome (Table 3). The results showed that the alignment rates of the four samples ranged from 94.60% to 96.02%, and the sequencing depth ranged from 15.38× to 21.87×, indicating a high degree of concordance between the sequencing data and the reference genome as well as good sequence specificity. Among them, a total of 92,048,512 reads from the parental line 01-26B were aligned to the reference genome, with genome coverage proportions of 91.38%, 87.69%, and 74.50% at 1×, 5×, and 10× depth, respectively. For the parental line 7009B, 81,490,278 reads were aligned, with coverage proportions of 91.36%, 86.68%, and 68.59% at 1×, 5×, and 10× depth, respectively. The alignment rates of PL-L and PL-S were higher than those of the parental lines, reaching 96.01% and 96.02%, respectively, with coverage proportions at 1×, 5×, and 10× depth of 93.61%, 89.34%, and 79.46% for PL-L and 91.88%, 89.39%, and 83.40% for PL-S. Overall, all samples showed uniform alignment, high genome coverage, and low error rates, meeting the requirements for subsequent BSA analysis.

Annotation characteristics of SNP variants: SNP variants identified in the parental lines and the extreme bulks were classified and statistically analyzed. The results showed that, among the four samples, the largest number of SNPs was located in intergenic regions, ranging from 1,521,350 to 1,950,878, accounting for 76.41%–78.97% of the total SNPs in each sample. Within genic regions, the numbers of SNPs located in exons ranged from 63,623 to 75,926, intronic SNPs ranged from 144,999 to 169,181, and SNPs located at splice sites ranged from 171 to 199. In addition, 18,522–21,939 SNPs were located in the 3′-UTR regions and 14,883–17,208 SNPs were located in the 5′-UTR regions. Moreover, the numbers of SNPs located upstream of genes ranged from 96,256 to 113,546, those located downstream of genes ranged from 76,633 to 89,863 (Fig. 3A).

SNPs located in exonic regions were further classified into four types: nonsynonymous mutations, synonymous mutations, premature stop codon mutations, and stop codon loss mutations. The numbers of nonsynonymous mutations in samples 01-26B, 7009B, PL-L, and PL-S were 32,139, 32,254, 37,202, and 38,427, respectively, while the numbers of synonymous mutations were 30,744, 31,174, 35,574, and 36,613, respectively. The numbers of premature stop codon mutations were 499, 428, 521, and 559, respectively, and the numbers of stop codon loss mutations were 113, 109, 123, and 118, respectively (Fig. 3B).

Fig. 1. Phenotypic comparison of parental sorghum lines 01-26B and 7009B.

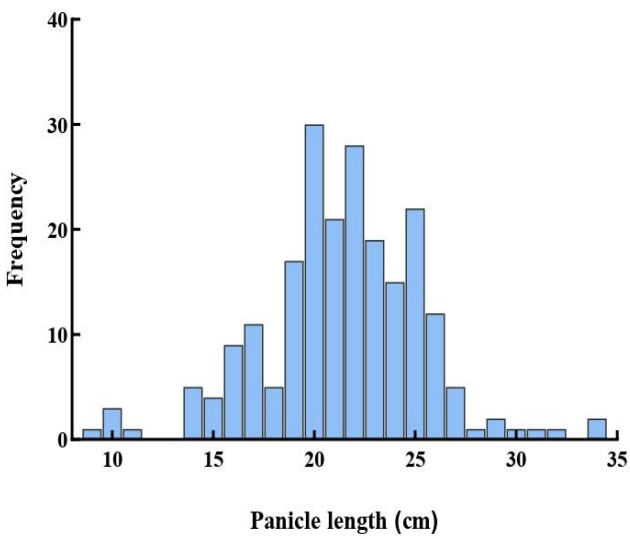

Fig. 2. Frequency distribution histogram of panicle length in F_2 population.

Table 1. Panicle length phenotypic data of the tested sorghum lines.

Trait	Parents			F_2 population					
	P_1	P_2	T-test	Mean ± standard error	Coefficient of variation (%)	Range of variation	Skewness	Kurtosis	Shapiro–wilk test
Panicle length (cm)	28.8	23.8	3.28**	21.39±0.27	18.56	9~34	-0.15	1.2	0.97**

Note: "**" Indicates p<0.01

Table 2. Quality assessment of sequencing data.

Sample ID	Total reads	Total bases	Error%	Q20 %	Q30 %	GC %
01-26B	64771676	9715751400	0.0258	97.5600	93.5800	44.1500
7009B	56932543	8539881450	0.0259	97.5200	93.4700	43.3200
PL-L	70472498	10570874700	0.0252	97.8200	94.1200	43.6200
PL-S	81140005	12171000750	0.0254	97.7100	93.9700	44.3600

Table 3. Mapping statistics of sequencing data.

Sample	Mapped reads	Unmapped reads	Mean depth	Coverage rate (1x) (%)	Coverage rate (5x) (%)	Coverage rate (10x) (%)
01-26B	92,048,512	3,700,735	17.36	91.38	87.69	74.50
7009B	81,490,278	2,869,705	15.38	91.36	86.68	68.59
PL-L	99,633,215	3,979,810	19.17	93.61	89.34	79.46
PL-S	115,658,765	4,606,720	21.87	91.88	89.39	83.40

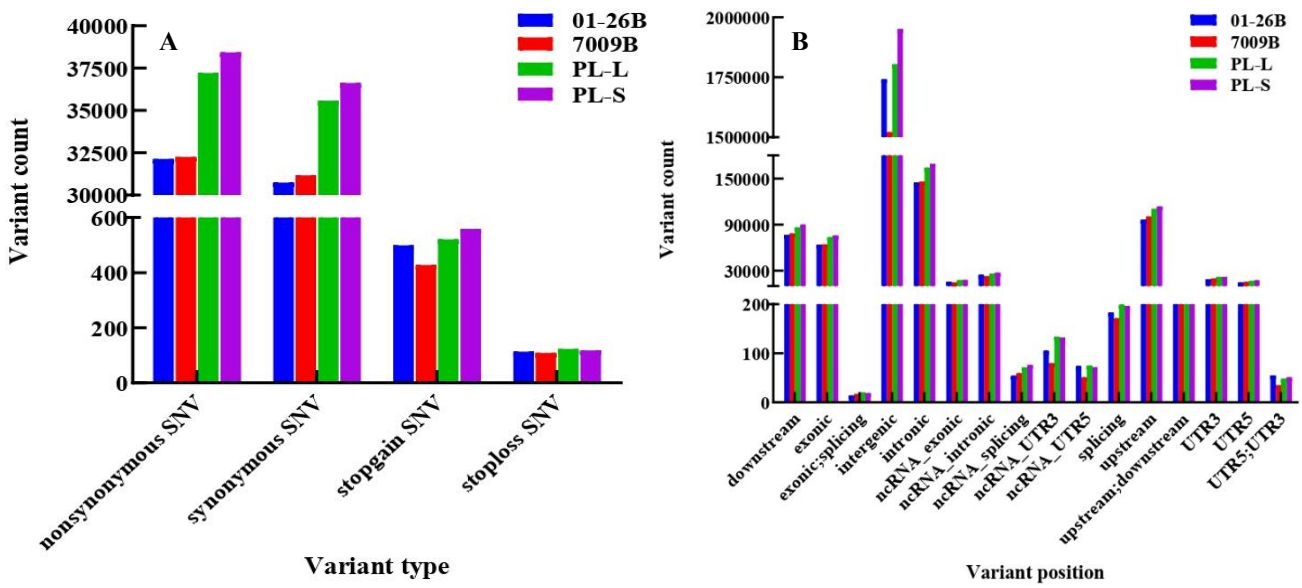

Fig. 3. SNP variant annotation.

Analysis of key SNP loci: Based on the genome-wide ΔSNP-index analysis, a total of 14 SNP loci significantly associated with sorghum panicle length were identified at a 95% confidence level, distributed across chromosomes 3, 6, 7, 9, and 10. On chromosome 7, at position 8,814,444 bp, a nonsynonymous mutation from C to G was detected in the exon of gene *LOC8071181*. On chromosome 6, a C-to-T SNP was found within an intron of gene *LOC8085745* at position 6,706,333 bp. Additionally, two A-to-G SNPs were identified upstream of gene *LOC8072072* on chromosome 6, located at positions 42,572,951 bp and 42,572,954 bp, respectively. The remaining 10 SNPs were located in intergenic regions, including one each on chromosomes 3, 9, and 10, and seven on chromosome 6. Of the seven intergenic SNPs on chromosome 6, five were situated between genes *LOC8072303* and *LOC8072304* (Table 4).

Functional analysis of candidate genes: The nonsynonymous mutation located in the exon is highly likely to alter the protein sequence and function; therefore, *LOC8071181* was selected as a candidate gene for further functional prediction analysis. In NCBI, the locus tag of *LOC8071181* is SORBI_3007G076100, and its protein is annotated as ECERIFERUM 26-like. The basic information of *LOC8071181* is presented in Table 5. The protein consists of 522 amino acids with a molecular weight of 55 kDa. Its instability index is greater than 40, indicating it is prone to degradation in vitro, and the aliphatic index is 88.28, suggesting a high proportion of hydrophobic aliphatic residues, which generally enhance protein thermostability. Structural analysis revealed that the protein contains a Pfam-defined transferase domain, occupying the majority of the protein, and may be involved in the elongation of very long chain fatty acids (VLCFAs) (Fig. 4).

Table 4. Key SNPs associated with sorghum panicle length at the 95% confidence level identified by genome-wide ΔSNP-index analysis.

Chromosome	Mutation site (bp)	Reference base	Variant base	Locus	Associated genes
3	45602338	G	T	Intergenic	*LOC8074956*, *LOC8074958*
6	6706333	C	T	Intronic	*LOC8085745*
6	21743209	A	G	Intergenic	*LOC110436431*, *LOC110436432*
6	42571713	T	C	Intergenic	*LOC8075254*, *LOC8072072*
6	42572951	A	G	Upstream	*LOC8072072*
6	42572954	A	G	Upstream	*LOC8072072*
6	55727394	A	G	Intergenic	*LOC8072303*, *LOC8072304*
6	55727406	C	T	Intergenic	*LOC8072303*, *LOC8072304*
6	55727421	T	C	Intergenic	*LOC8072303*, *LOC8072304*
6	55727424	G	A	Intergenic	*LOC8072303*, *LOC8072304*
6	55727425	G	A	Intergenic	*LOC8072303*, *LOC8072304*
7	8814444	C	G	Exonic	*LOC8071181*
9	24007541	G	A	Intergenic	*LOC110430125*, *LOC8061270*
10	45484861	C	T	Intergenic	*LOC110431331*, *LOC8072635*

Table 5. Basic information of *LOC8071181*.

Gene ID	Number of amino acids (aa)	Molecular weight (KDa)	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity
	522	55	5.99	45.64	88.28	0.06

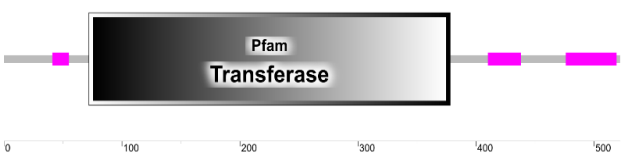

Fig. 4. Structural domain characteristics of *LOC8071181*.

Discussion

Panicle length is an important component of sorghum inflorescence morphology and represents a key agronomic trait that cannot be overlooked in yield formation and variety improvement. Previous studies have shown that panicle length is generally significantly positively correlated with yield components such as the number of grains per panicle and single-panicle grain weight. A longer rachis facilitates the proper branching and distribution of spikelets, thereby providing greater spatial capacity for grain formation (Li *et al*., 2019; Tan *et al*., 2022; Ding *et al*., 2024). Therefore, systematically dissecting the genetic basis of panicle length and identifying key regulatory loci and candidate genes not only enhances theoretical understanding of sorghum inflorescence development but also provides potential targets for high-yield molecular breeding, accelerating the improvement of sorghum varieties.

In recent years, with the development of molecular marker technologies and whole-genome sequencing, genetic studies on sorghum panicle length have made considerable progress. Srinivas *et al*., (2009) constructed a recombinant inbred line (RIL) population of 168 families derived from a cross between sorghum lines 296B and IS18551. Using 152 SSR and morphological markers, they constructed a genetic linkage map and identified four QTL controlling panicle length on chromosomes 2, 6, and 7 using a composite interval mapping approach. Nagaraja *et al*., (2013) used a RIL population of 245 families derived from a cross between M35-1 and B35, constructed a linkage map with 237 molecular markers, and detected panicle length-related QTL on chromosomes 4, 6, and 9 through composite interval mapping. Shehzad and Okuno (2015) studied an F_2 population derived from Red Kafir × Takakibi and identified seven QTL for panicle length, including five major QTL on chromosomes 1, 2, 6, and 7, and two minor QTL on chromosomes 8 and 10. Ding *et al*., (2023) used a RIL population of 205 families derived from BTx623 × Hongyingzi and, based on a high-density linkage map and complete interval mapping, detected 14 QTL associated with panicle length across chromosomes 1, 3, 4, 5, 6, 7, 8, and 9. Collectively, these studies indicate that panicle length-related QTL are repeatedly detected on multiple chromosomes, with chromosomes 6, 7, and 9 consistently identified across studies, suggesting that these regions may harbor important genomic loci regulating panicle length.

Building on these findings, this study used sorghum maintainer lines with significant differences in panicle length as parental materials to construct an F_2 segregating population and systematically analyze the phenotypic variation of panicle length. The results showed that panicle length in the F_2 population displayed a continuous distribution with some degree of transgressive segregation, consistent with the inheritance pattern of typical quantitative traits. Extreme phenotypic individuals were used to construct long-panicle and short-panicle bulks, and BSA-seq was employed for genome-wide mapping of

panicle length-associated loci. At a 95% confidence level, 14 SNP loci significantly associated with panicle length were identified, distributed on chromosomes 3, 6, 7, 9, and 10. Notably, chromosomes 6, 7, and 9 are highly consistent with previous studies, further confirming the importance of these regions in the genetic regulation of panicle length. Of particular interest, a key SNP on chromosome 7 was located in the exon of *LOC8071181*. Functional annotation revealed that this gene encodes an ECERIFERUM 26-like (CER26-like) protein. Further structural analysis showed that the majority of the protein contains a transferase domain, predicted to be involved in the elongation of very long chain fatty acids (VLCFAs).

Previous studies have shown that plant aerial organ surface waxes are primarily composed of VLCFAs and their derivatives, playing critical roles in reproductive development and environmental adaptation. The CER (eceriferum) gene family is widely recognized as a core regulator of VLCFA synthesis and wax biosynthesis (Lai *et al*., 2007; Panahi *et al*., 2025). In rice, OsCER1 participates in VLC alkane synthesis and plays an important role in anther development, plastid differentiation, and pollen formation. Expression analysis indicates that OsCER1 is highly expressed in the tapetal layer and bicellular pollen but shows low expression in vegetative organs, suggesting that CER genes are closely related to inflorescence development (Ni *et al*., 2018). In Arabidopsis and barley, 22 and 85 CER loci have been identified, respectively, revealing organ-specific expression patterns, with some genes mainly affecting inflorescence or internode development (Kunst and Samuels, 2003). In barley, wax-deficient mutants cer.j59 and cer.i16 exhibit significantly shortened panicles(von Wettstein-Knowles, 2020), and studies by Bian (2022) showed that wax-deficient hull mutants display higher non-stomatal transpiration, increased drought sensitivity, and significant reductions in panicle length, panicle width, and thousand-grain weight. These studies further support a functional association between CER genes and inflorescence morphology.

At the molecular level, CER2 and its homologs CER26 and CER26-like are thought to play key regulatory roles in VLCFA elongation. In *Arabidopsis*, CER2 promotes the elongation of C28 fatty acids to C30, whereas CER26 and CER26-like participate in the further elongation of C30 and longer VLCFAs. Mutants of CER26 or CER26-like show significant alterations in ultra-long-chain alkane composition, indicating that these proteins are important for maintaining wax chain-length distribution (Haslam *et al*., 2012; Haslam *et al*., 2015).

Based on previous studies and the results of this study, it is speculated that the CER26-like gene may indirectly affect the growth and development of sorghum panicle rachises by regulating VLCFA biosynthesis and wax deposition. On one hand, VLCFAs are essential components of membrane lipids and sphingolipids, and their abnormal synthesis may affect membrane stability and cell elongation, thereby limiting longitudinal growth of rachis cells. On the other hand, the wax layer plays a critical role in protecting young inflorescence tissues, preventing excessive water loss, and buffering environmental stress; defects in wax synthesis may render panicles more sensitive to stress, consequently affecting normal panicle elongation.

Although this study identified candidate genes significantly associated with sorghum panicle length, the precise regulatory mechanisms remain to be further validated. Future studies could proceed in several directions: (1) expression analysis of rachis tissues at different developmental stages to determine spatiotemporal expression patterns of candidate genes; (2) functional validation through genetic transformation, gene editing, or mutant materials to directly assess their effects on panicle length; and (3) further investigation of significant SNPs located in regulatory regions to explore their potential influence on gene expression. Overall, this study provides new genetic evidence and candidate gene resources for sorghum panicle length, laying a foundation for future studies on the molecular regulation of panicle development and the genetic improvement of related traits.

Conclusions

This study provides new insights into the genetic regulation of panicle length in sorghum through the identification of stable genomic regions and a promising candidate gene. Using an F_2 population derived from parents with contrasting panicle length phenotypes, multiple loci associated with panicle length were detected across several chromosomes, reflecting the polygenic nature of this trait. Among the identified loci, a nonsynonymous variant located in the coding region of *LOC8071181* on chromosome 7 was highlighted as a strong candidate. Functional annotation suggests that this gene encodes an ECERIFERUM 26-like protein potentially involved in very-long-chain fatty acid elongation, a process closely linked to wax biosynthesis, membrane integrity, and organ elongation. These findings indicate that lipid metabolic pathways may play an indirect but important role in regulating panicle rachis development in sorghum.

Overall, the results of this study enrich the understanding of the molecular basis underlying panicle length variation and provide valuable genetic targets for further functional validation. The identified loci and candidate gene offer practical potential for marker-assisted selection and the genetic improvement of panicle architecture and yield-related traits in sorghum.

Funding: This study was supported by the Liaoning Provincial Applied Basic Research Program (2025JH2/101300075); the Collaborative Innovation Project of Liaoning Academy of Agricultural Sciences (2025XTCX0401-2-3); National Grain Sorghum Modern Agricultural Industrial Technology System for Forage Sorghum Breeding Post Scientist Project (CARS-06-14.5-A11); the Agricultural Green and High-Quality Development Project of Liaoning Academy of Agricultural Sciences (2025HQ1307); Liaoning Province's major project on hiding grain in technology (2023JH1/10200001) and the Doctoral Start-up Project of Liaoning Academy of Agricultural Sciences (2025BS1703).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author's Contribution: Linlin Yang, Yanqiu Wang and Kuangye Zhang designed the study. Linlin Yang, Kuangye Zhang and Zeyang Zhao performed the experiments. Han Wu, Youhou Duan and Fulai Ke performed the data analyses. Linlin Yang wrote the manuscript. Fei Zhang and Kuangye Zhang critically reviewed the manuscript. All authors read and approved the submitted version.

References

Agata, A., K. Ando, S. Ota, M. Kojima, Y. Takebayashi, S. Takehara, K. Doi, M. Ueguchi-Tanaka, T. Suzuki, H. Sakakibara, M. Matsuoka, M. Ashikari, Y. Inukai, H. Kitano and T. Hobo. 2020. Diverse panicle architecture results from various combinations of Prl5/GA20ox4 and Pbl6/APO1 alleles. *Comm. Biol.*, 3(1): 302.

Bian, X., L. Yao, E. Si, Y. Meng, B. Li, X. Ma, K. Yang, Y. Lai, X. Shang, C. Li, J. Wang, and H. Wang. 2022. Characterization of glossy spike mutants and identification of candidate genes regulating cuticular wax synthesis in barley (*Hordeum vulgare* L.). *Int. J. Mol. Sci.,* 23(21): 13025.

Ding, H., C. Wang, Y. Cai, K. Yu, H. Zhao, F. Wang, X. Shi, J. Cheng, H. Sun, Y. Wu, R. Qin, C. Liu, C. Zhao, X. Sun and F. Cui. 2024. Characterization of a wheat stable QTL for spike length and its genetic effects on yield-related traits. *BMC Plant Biol*., 24(1): 292.

Ding, Y., C. Wang, J. Xu, X. Gao, B. Cheng, N. Cao and L. Zhang. 2023. QTL identifying for panicle architecture-related traits in sorghum based on high-density genetic map. *J. Plant Genet. Resour.*, 24(04): 1122-1132.

Dong, G.C., J. Ju, X.X. Yu, Y. Zhang, J.N. Zhao, H. Tian, B. Zhang and Y.L. Wang. 2010. Study of difference of yield formation in conventional Indica rice cultivars with different panicle weight. *J. Yangzhou Univ.* (*Agric. Life Sci. Ed.*), 31(01): 49-54.

Haslam, T.M., A. Mañas-Fernández, L. Zhao and L. Kunst. 2012. Arabidopsis ECERIFERUM2 is a component of the fatty acid elongation machinery required for fatty acid extension to exceptional lengths. *Plant Physiol*., 160(3): 1164-1174.

Haslam, T.M., R. Haslam, D. Thoraval, S. Pascal, C. Delude, F. Domergue, A.M. Fernández, F. Beaudoin, J.A. Napier, L. Kunst and J. Joubès. 2015. ECERIFERUM2-LIKE proteins have unique biochemical and physiological functions in very-long-chain fatty acid elongation. *Plant Physiol.*, 167(3): 682-692.

Huang, L.L., J. Ye, R.R. Zhai, Q. Huang, M.M. Wu, G.F. Zhu, F.M. Yu, X.M. Zhang, J.L. Han and S.H. Ye. 2025. Agronomic traits and heterosis in indica-japonica hybrid rice combinations. *Acta Agr. Zhejiang*, 37(10): 2021-2031.

Hussain, I., Z. Hasnain, S. Zafar, M.J. Nazir and S.A. Jatoi, 2024. Salt induced positive and negative effects on germination and other physiological traits of sorghum vulgare seedlings. *Pak. J. Bot.*, 56: 807-814.

Kunst, L. and A.L. Samuels. 2003. Biosynthesis and secretion of plant cuticular wax. *Prog Lipid Res.*, 42(1): 51-80.

Lai, C., L. Kunst and R. Jetter. 2007. Composition of alkyl esters in the cuticular wax on inflorescence stems of *Arabidopsis thaliana* cer mutants. *Plant J.*, 50(2): 189-196.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16): 2078-2079.

Li, M. Y., Z.Q. Guo, Y. Bai, P.F. Zhu, J. Wang, Y.C. Han, F. Zhang and X.H. Lin. 2024. Effects of the nitrogen application rate on dry matter accumulation, nitrogen uptake and utilization, and key enzyme activities of nitrogen metabolism in sorghum. *Pak. J. Bot.*, 56(3): 1099-1108.

Li, R., M. Li, U. Ashraf, S. Liu and J. Zhang. 2019. Exploring the relationships between yield and yield-related traits for rice varieties released in China from 1978 to 2017. *Front. Plant Sci.*, 10: 543.

Ma, L., T.L. Hu, M. Kang, X.K. Fu, P.Y. Chen, F. Wei, H.L Jiang, X.Y. Lü, M. Zhang and Y.L. Yang. 2024. Identification of candidate genes for early-maturity traits by combining BSA-seq and QTL mapping in upland cotton (*Gossypium hirsutum* L.). *J. Integr. Agric.*, 23(10), 3472-3486.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M.A. DePristo. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome. Res.*, 20(9), 1297-1303.

Nagaraja, R.R., R. Madhusudhana, S. Murali Mohan, D.V. Chakravarthi, S.P. Mehtre, N. Seetharama and J.V. Patil. 2013. Mapping QTL for grain yield and other agronomic traits in post-rainy sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.*, 126(8): 1921-1939.

Ni, E., L. Zhou, J. Li, D. Jiang, Z. Wang, S. Zheng, H. Qi, Y. Zhou, C. Wang, S. Xiao, Z. Liu, H. Zhou and C. Zhuang. 2018. OsCER1 plays a pivotal role in very-long-chain alkane biosynthesis and affects plastid development and programmed cell death of tapetum in rice (*Oryza sativa* L.). *Front. Plant Sci*., 9: 1217.

Panahi, B., R. Hamid and F. Jacob. 2025. Genome-wide characterization of the Eceriferum (CER) gene family in barley (*Hordeum vulgare* L.). *Sci. Rep.*, 15(1): 20674.

Shehzad, T. and K. Okuno. 2015. QTL mapping for yield and yield-contributing traits in sorghum (*Sorghum bicolor* (L.) Moench) with genome-based SSR markers. *Euphytica*, 203(1): 17-31.

Srinivas, G., K. Satish, R. Madhusudhana, R.N. Reddy, S.M. Mohan and N. Seetharama. 2009. Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. *Theor. Appl. Genet.*, 118(8): 1439-1454.

Sun, T.N., K. Sui, J.C. Tai, X.G. Wang, Q.X. Zhang, Z. Niu, B.C. Liu and Y.X. Zhang. 2025. Comprehensive evaluation on production performance of 10 summer sown forage-grain dual-purpose maize varieties in Horqin Sandy Land. *Feed Res.*, 20: 140-144.

Tan, Y., C. Kang and S.J. Shou. 2022. Genetic analysis of yield traits in F_2 generation of Xinliang 52×Meiyin-251 hybrid combination. *J. Anhui Agric. Univ*., 50(02): 26-29+41.

von Wettstein-Knowles, P. 2020. Ecophysiology with barley eceriferum (cer) mutants: the effects of humidity and wax crystal structure on yield and vegetative parameters. *Ann. Bot.*, 126(2): 301-313.

Wang, K., M. Li, and H. Hakonarson. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucl. Acids Res*., 38(16): e164.

Wang, N., S.X. Wang, L. Zhang, H.L. Ren, R.H. Dang and L.Y. An. 2019. Correlation and path analysis of yield and yield characteristics of maize. *Helongjiang Agric. Sci.*, 01: 23-27.

Wu, R.G., Y.T. Pei, C. Zhang, F.Y. Quan, Y.R. Jin, P. Liu, D.X. Jia and Z.M. Zhong. 2019. Analysis and evaluation of agronomic character diversity of wheat based on salt stress. *J. Triticeae Crop.*, 39(09): 1029-1037.

Yang, B., L. Yang, L. Kang, L. You, H. Chen, H. Xiao, L. Qian, Y. Rao and Z. Liu. 2024. Integrated analysis of BSA-seq and RNA-seq identified the candidate genes for seed weight in Brassica juncea. *Front. Plant Sci.*, 15: 1458294.

Zhang B., F. Qi, G. Hu, Y. Yang, L. Zhang, J. Meng, Z. Han, X. Zhou, H. Liu, M. Ayaad and Y. Xing. 2021. BSA-seq-based identification of a major additive plant height QTL with an effect equivalent to that of Semi-dwarf 1 in a large rice F_2 population. *Crop J.*, 9(6): 1428-1437.

Zhang, F., Y. Wang, H. Yu, K. Zhu, Z. Zhang, F. Lu and J. Zou. 2016a. Effect of excessive soil moisture stress on sweet sorghum: physiological changes and productivity. *Pak. J. Bot.*, 48(1): 1-9.

Zhang, S., H.J. Sui, Z.Y. Ge, L.X. Wang, H.J. Cheng, Y.X. Pan and F.Z. Yan. 2016b. Grey relational analysis for main agronomic traits and yield of 14 sorghum varieties. *Guizhou Agric. Sci.*, (01): 23-27.

Zheng, H., Y. Dang and N. Sui. 2023. Sorghum: a multipurpose crop. *Agric. Food Chem*, 71(46): 17570-17583.

Zou, C., P. Wang and Y. Xu. 2016. Bulked sample analysis in genetics, genomics and crop improvement. *Plant Biotechnol. J.*, 14(10): 1941-1955.