scientific reports

Check for updates

OPEN

Research on the mining of candidate genes for pepper fruit color and development of SNP markers based on SLAF-seq technology

Yaning Meng[1,4], Xinxin Li[1,4], Hongxiao Zhang[1,3], Zhanghong Yu[1], Zhe Zhang[1]✉, Yanqin Fan[1]✉ & Libin Yan[1,2]✉

This study aims to enhance the coloration of pepper fruit by identifying valuable genetic resources through the analysis of single nucleotide polymorphism (SNP). markers and candidate genes associated with fruit pigmentation. Utilizing 197 natural populations of both hot and sweet peppers, we employed specific-locus amplified fragment sequencing (SLAF-seq) to examine 1496 high-quality SNP markers, thereby identifying significant loci contributing to fruit color variation. Our genome-wide association study pinpointed 30 significant SNP sites located on chromosome 6. Further analysis using kompetitive allele-specific PCR(KASP) and phenotypic correlation with fruit color led to the identification of the CA.PGAv.1.6.scaffold919.44 gene, which is implicated in anthocyanin synthesisregulation via the NAC domain, thereby influencing pepper fruit coloration. These findings offer a valuable reference for the advancement of molecular-assisted breeding strategies aimed at improving the fruit color of both sweet and hot peppers.To improve the fruit color of sweet peppers, this study aimed to identify single nucleotide polymorphism (SNP) loci and candidate genes significantly associated with fruit color. A natural population of 197 sweet pepper accessions was used as the material. SLAF-seq was conducted with 1496 high-quality SNP markers to mine excellent variant loci and predict candidate genes. Through Manhattan plot analysis and association analysis with the best linear unbiased prediction (BLUP) values of fruit color, 30 significant loci were detected on chromosome 6. Combining KASP genotyping technology with field phenotypes, the gene CAPGAv.1.6.scaffold919.44 was identified as a candidate gene regulating mature fruit color. It is related to the NAC domain and is hypothesized to alter fruit color by regulating anthocyanin biosynthesis. This study lays the foundation for molecular-assisted breeding of sweet peppers related to fruit color.

Keywords Sweet peppers, Peppers, Fruit color, Specific-locus amplified fragment sequencing, Candidate genes

Peppers (*Capsicum annuum L.*) are the fruits of plants within the Solanaceae family and the Capsicum genus[1]. Among these, sweet pepper is a particularly significant vegetable crop in China,valued for its substantial economic and nutritional contributions[2]. The coloration of the fruit is a critical quality attribute for both peppers and sweet peppers, significantly influencing consumer purchasing decisions[3]. This coloration is primarilydetermined by the relative concentrations of chlorophylls, carotenoids, and flavonoids[4]. This coloration is primarily[5]. The CaGLK2 gene can regulate the color change of pepper in the immature stage by altering the size of the chloroplast compartment[6]. The CaSGR gene can promote chlorophyll degradation in pepper during fruit ripening[7]. Atkins et al. identified that the yellow coloration of pepper fruits is governed by recessive y alleles, whereas the red

[1]Institute of Economic Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang, China. [2]Hebei Province Engineering Research Center for Vegetables, Shijiazhuang, China. [3]College of Horticultural Science & Technology, Hebei Normal University of Science & Technology, Qinhuangdao, China. [4]Yaning Meng and Xinxin Li contributed equally to this work and should be considered co-first authors. ✉email: luanzhi914@163.com; nkyfanyq@126.com; tianjiaoshi@126.com

coloration is determined by the dominant y + allele[8–10]. Furthermore, previous studies have indicated that the CCS gene is instrumental in determining the red and yellow coloratio of pepper fruit, and the yellow color of the fruit is due to the deletion of the CCS gene[11]. However, there are few studies on candidate genes for fruit color traits in sweet pepper and pepper.

The rapid development of high-throughput sequencing technologies provides a powerful tool for conducting biological research. SLAF-seq is one of the most promising technologies[12], and it has several notable features. SLAF-seq has the following advantages: (i) generation of high-density SNP loci numbering in the millions after one sequencing reaction, (ii) capability of detecting novel SNP loci in unknown mutation-harboring loci compared with SNP arrays, (iii) suitability for any species regardless of the presence of a reference genome, and (iv) a higher rate of identified SNP loci that become genuine association markers[13,14]. Recently, SLAF-seq has been applied to various crops, including tomato[15], wheat[16], cucumber[17], and eggplant[18]. Ning et al. utilized SLAF-seq to construct a high-density genetic map of pepper and identified three candidate genes associated with resistance tocucumber mosaic[19].

The researcher constructed a high-density genetic map and mapped a gene CA02g30020 associated with the first flower node trait in pepper using SLAF-seq[20]. Consequently, the application of SLAF-seq technology for the identification of candidate genes related to fruit color traits in sweet pepper and pepper holds significant research potential.In our study, By using a specific locus amplified fragment sequencing (SLAF-seq) sequencing method and combined with Genome-wide association analysis on 197 different sweet pepper and pepper germplasm resources, the candidate genes related to fruit color in pepper were mined so as to provide a theoretical foundation for the further molecular breeding in pepper.

Materials and methods

Plant materials

The 197 different sweet pepper and pepper germplasm resources were Provided by the Solanum Fruit Room, Institute of Economic Crops, and Hebei Academy of Agricultural and Forestry Sciences (Fig. 1) (Table S1). Peppers and sweet peppers are red or yellow when ripe. The red color of peppers and sweet pepper fruit is controlled by the dominant gene, the yellow color is controlled by a recessive gene, and the genetics are stable.

Library construction and sequencing

Frozen young plant leaves, stored at − 80 °C, were ground in liquid nitrogen, and genomic DNA was extracted according to the CTAB method[21]. DNA concentration and quality were determined, with a concentration ≥ 20 ng/μL and a volume ≥ 30 μL, to ensure that the DNA sample met the requirements for sequencing. After the sample DNA was qualified for detection, DNA was fragmented by ultrasound. End-repair, adaptor jointing, and purification of the DNA fragments were performed to construct the library. Sequencing was performed by using an Illumina platform.

Enzymatic digestion and library construction

The published pepper genome (https://www.ncbi.nlm.nih.gov/genome/?Term=Capsicum annuum) was used as the reference genome. The SLAF-predict software (Independently developed by Beijing Baimaike Biotechnology Co., Ltd.) was employed to predict the scheme through the pepper genome, determining the enzyme cutting combination and performing the digestion. The 3′ ends of the obtained enzyme-cut fragments were treated with an A-addition process, andsequencing adapters were ligated onto the A-tailed enzyme-cut fragments. Thequalified library was sequenced using the Illumina Hiseq 4000 sequencing platform after PCR amplification, purification, pooling, and gel recovery of the target fragments.

Data processing and SNP calling

Sequence alignment Sequencing reads were aligned to the pepper reference genome using the BWA0.7.10-r789 software[22] (http://maq.sourceforge.net). A set of SNP sites was obtained by selecting the intersection of the SNP discovered by GATK and Samtools (v1.3.1) (http://samtools.sourceforge.net) to ensure accurate SNP detection[23]. The SNP marker with high quality were obtained by the following standard: integrity > 0.8, minor allele frequency MAF > 0.05.

Fig. 1. The color of pepper and sweet pepper in the turing stage.

Genome-wide association analysis (GWAS) and mining of candidate genes

Total filtered SNPs detected from 197 accessions were used for GWAS with a mixed linear model (EMMAX、FaST-LMM, GEMMA)[24,25]. The EMMAX software version is emmax-intel64 (http://csg.sph.umich.edu//kang/emmax/download/index.html), The FaST-LMM software version is 2.07 (https://github.com/fastlmm/FaST-LMM), The GEMMA software version is 0.98.1 (https://github.com/genetics-statistics/GEMMA) .The GWAS threshold was set to − log10 ($P > 5$), which were defined as significant trait-associated SNPs. Genes that were located within the region 100 kb upstream or downstream of trait-associated SNPs were identified as candidate-associated genes. The Quantile–Quantile plot (Q-Q plot) was drawn by the GGplot2 3.3.0 software[26], and the Manhattan plot was drawn by QQman0.1.9 software[27]. Meanwhile, the candidate-associated genes were annotated by NR, SwissProt[28], GO[29], COG[30], and KEGG[31] databases to explore the function of the genes.

KASP marker development and genotyping

Sequences covering SNP sites and approximately 100 bp in length were selected. The Primer3 v4.1.0 was used to design KASP primers. Each KASP marker contained three primer sequences, namely two forward primers to distinguish SNP alleles and one general reverse primer. The primer sequences were as follows:

F1:5′-GAAGGTGACCAAGTTCATGCTCCTATGCTTTCCAGAGGCC-3′, F2:5′-GAAGGTCGGAGTC AACGGATTCCTATGCTTTCCAGAGGCT-3′ and R:5′-CTTTCCAACGATACCAATTTGGCT-3′. The SNP genotyping was conducted in the Cash Crops Institute of Hebei Academy of Agricultural and Forestry Sciences laboratory. A Thermo Fisher 7500 was used for SNP signal detection.

Results

Statistics of fruit color traits of pepper and sweet peppers

Firstly, an analysis of phenotypic statistics related to fruit color traits was conducted on 197 samples of sweet pepper and pepper at the ripening stage. The results indicated that the predominant colors at maturity were red and yellow. Specifically, the number of red fruits was significantly higher than that of the yellow fruits, comprising 32 yellow peppers, 58 red peppers, 34 yellow sweet peppers, and 73 red sweet peppers (Figure S1).

Genome-wide association analysis

The Q–Q plot (Fig. 2) demonstrated a fundamental overlap between observed and expected values, with only a slight deviation in the P-valuedistribution in the tail region. This suggests a significant correlation between phenotype and genotype. Furthermore, therewas no deviation from the statistical distribution assumptions in the SNP association analysis, indicating that the model is appropriate for this population (Tables S2, S3, and S4). The Manhattan plot analysis revealed that significant SNP sites with a − log10 (P) > 5 were predominantly located on chromosomes 1, 2, and 6, while those with a − log10 (P) > 7 were mainly found on chromosomes 2 and 6. Notably, SNP sites with a − log10 (P) > 8 were primarily concentrated on chromosome 6 (refer to Fig. 3 and Table 1). This distribution suggests that the majority of significant SNP sites are situated on chromosome 6, implicating it as a potential locus for genes influencing fruit color variation.

Mining of candidate genes

The physical location of the significant SNP sites in chromosome 6 from the tip of the chromosome ranged from 232.8 to 236.4 Mb, covering a total length of 3.6 Mb. Linkage disequilibrium (LD) analysis ($r^2 > 0.6$) identified 116 out of 679 SNPs significantly associated with fruit color traits on chromosome 6, which were grouped into 16 candidate genomic regions. Considering the incomplete coverage of the pepper genome by SLAF-Seq, we conducted an analysis of strongly associated loci within the genome's physical position, examining regions extending 100 kb upstream and downstream of these loci to identify potential genes. This investigation led to the identification of six genes, encompassing 30 SNPs, associated with fruit color, as determined through comparisons with the Pfam and SwissProt databases.The six candidate genes were CA.PGAv.1.6.scaffold1432.27, CA.PGAv.1.6.scaffold1726.1, CA.PGAv.1.6.scaffold1755.9, CA.PGAv.1.6.scaffold919.10, CA.PGAv.1.6.scaffold919.38, and CA.PGAv.1.6.scaffold919.44. Their functions were classified as unknown, translation, energy production and conversion, lipid transport and metabolism, and inorganic ion transport and metabolism by COG annotations. According to Pfam and Swissprot databases, the gene functions were classified as bHLH-MYC and R2R3-MYB transcription factors N-terminal and Myb-like DNA-binding, respectively Domain, Lycopene cyclase protein, Plant protein of unknown function, Lecithin: cholesterol acyltransferase and NAC domain-containing protein (Fig. 3).

Identification of candidate genes based on KASP genotyping analysis

A total of 197 natural populations of sweet pepper and pepper plants wereanalyzed to detect 30 mutation sites using KASP, which were correlated with field phenotype. Notably, the mutation site at position 233,136,896 exhibited the highest genotyping accuracy, achieving 96% accuracy of 96% (Table S5, Table S6). The nucleotide bases at these sites transition from C to T and are located upstream of the genes CA.PGAv.1.6.scaffold919.43, CA.PGAv.1.6.scaffold919.44 and CA.PGAv.1.6.scaffold919.45. The SNP loci is especially close to CA.PGAv.1.6.scaffold919.44 gene which contains a NAC40 domain related to fruit color and belongs to the NAC transcription factor family. The CA.PGAv.1.6.scaffold919.44 gene (NAC40 gene) is 3998 bp and contains four exons and three introns. We speculated the CA.PGAv.1.6.scaffold919.44 gene (NAC40 gene) may regulate the fruit color of pepper and sweet pepper (Fig. 4).

Homology of the NAC40 gene with other species

To explore the function of the NAC40 gene, we performed multiple sequence alignment and phylogenetic tree analysis of the NAC40 gene with other species. The result showed that the NAC40 gene is highly conserved and

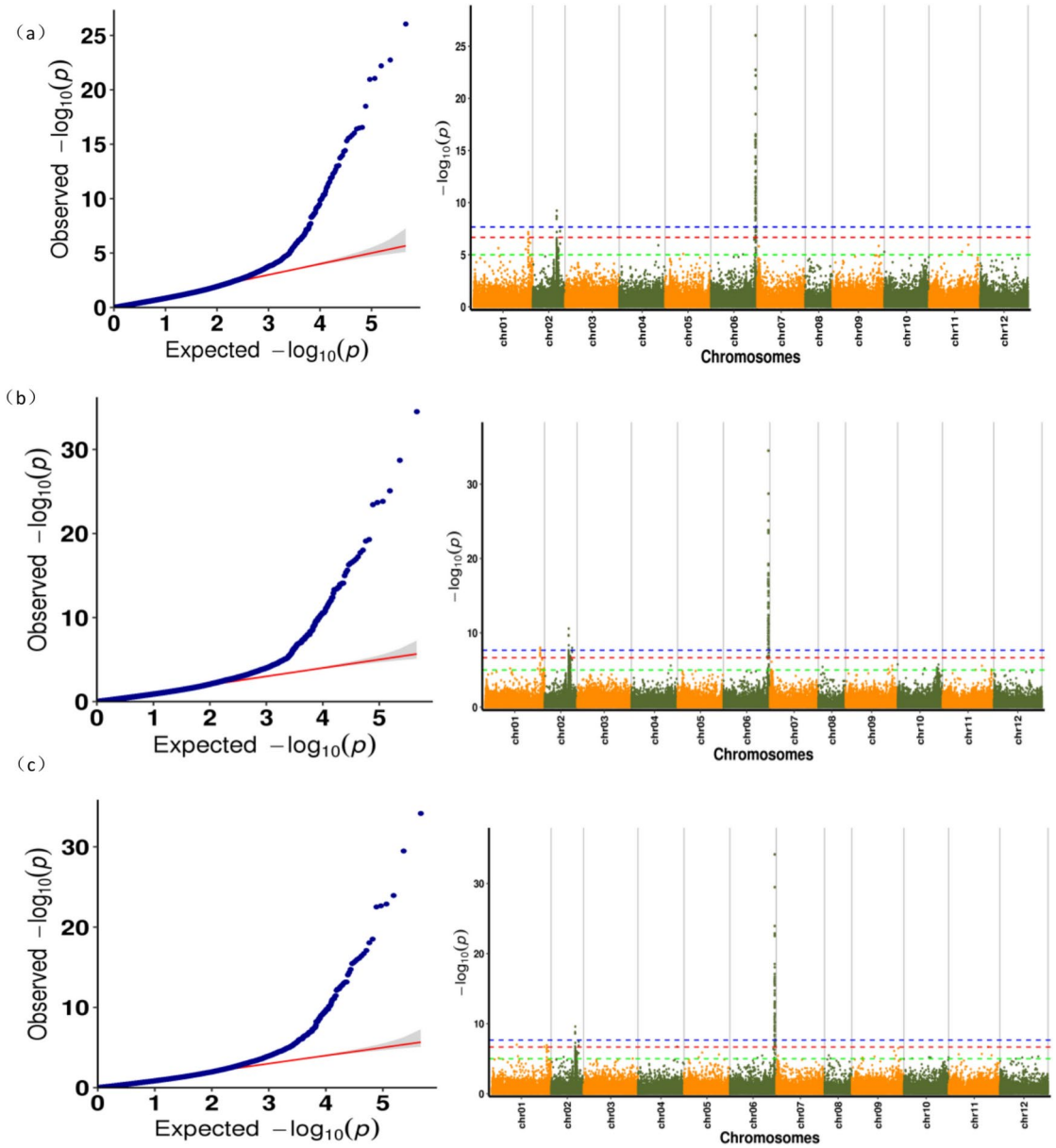

Fig. 2. Quantile–quantile plots and Manhattan plots of genome-wide association study involving pepper color using (**a**) Q–Q plots and Manhattan plots under the EMMAX model, (**b**) Q–Q plots and Manhattan plots under the FaST-LMM model, and (**c**) Q–Q plots and Manhattan plots under the GEMMA model .

belongs to the same branch as tomato, tobacco, and potato, indicating that their homology is very close. The NAC40 gene is likelier to play the same role as a close homolog. We then predicted the function of NAC40 in tomato, tobacco, and potato and found an association with the NAC transcription family (Fig. 5).

Discussion

The SLAF-seq strategy, predicted on high-throughput sequencing technologies, was developed as a simplified genome sequencing method.Its application for de novo SNP discovery, genotyping, and mapping has been successfully demonstrated across various plant species, including sesame[32], cucumber[33], and soybean[34]. Compared to alternative methodologies, SLAF-seq technology offers several advantages, such as the detection of a greater number of markers, enhanced map quality, methodologies, SLAF-seq technology offers several advantages, such as the detection of a greater number of[35]. In this study, we used SLAF-seq technology to sequence 197 sweet pepper and pepper genomes. Our findings revealed a novel SNP associated with fruit color traits, which facilitates subsequent molecular marker-assisted breeding aimed at enhancing the commercial value of sweet pepper and pepper..

Numerous agronomic traitsexhibit diverse phenotypes in sweet pepper and pepper, with fruit coloration being one of the most critical determinants of market value[36,37]. Wu et al. identified a mutation in the 313 base

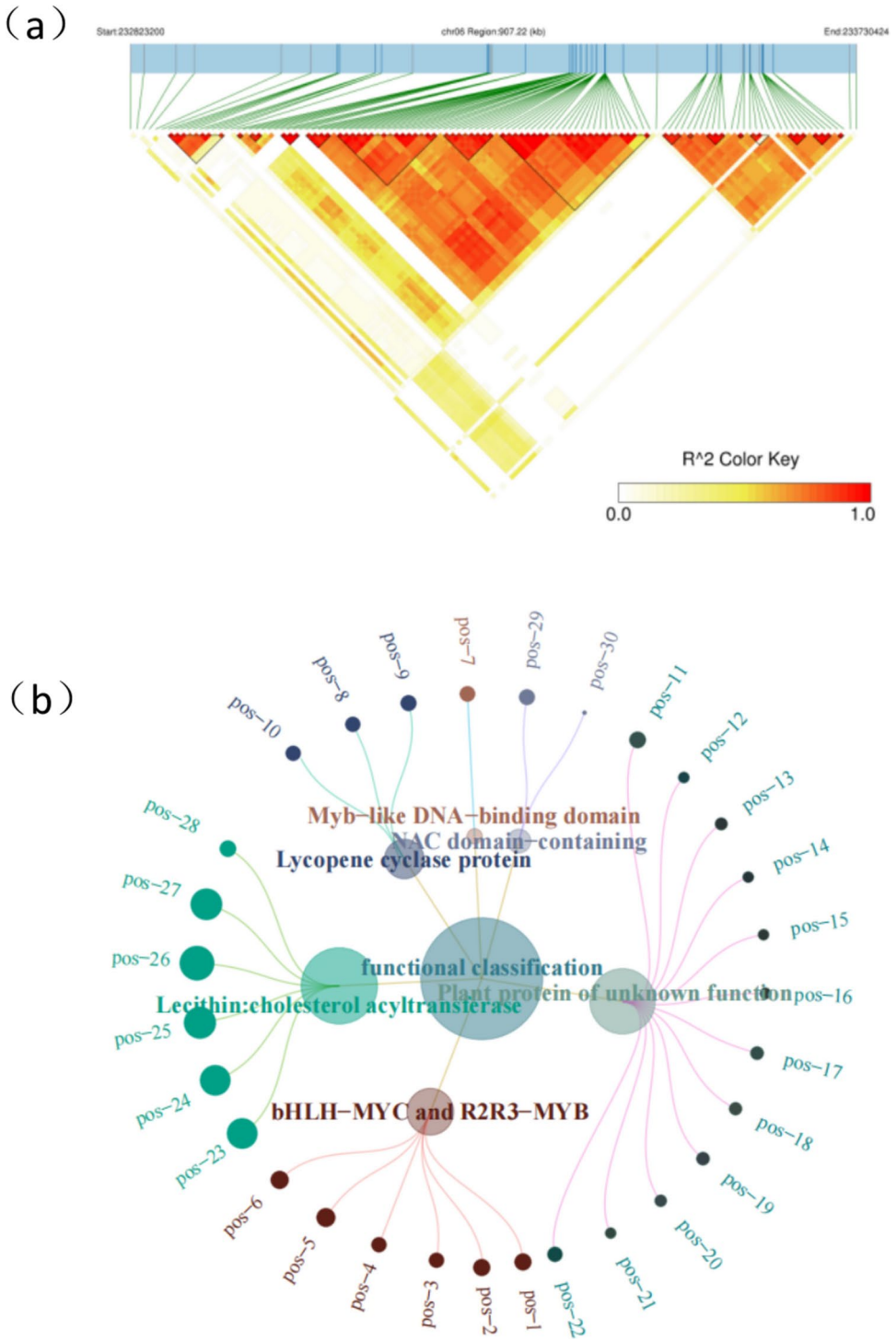

Fig. 3. (**a**) Haploblock showing markers in linkage disequilibrium (LD) ($> r^2 = 0.6$). The square value of R represents the strength of linkage disequilibrium, and the larger the value, the stronger the LD.The redder the color in the picture, the higher the correlation of loci. (**b**) Annotation of functions for the candidate-associated genes.

of the coding sequence of Capana10g001710 located in the 35.07 kb region of chromosome 10, resulting in the conversion of α-helix to β-fold of the encoded PP2C35 protein, which was predicted to change fruit color by regulating the accumulation of chlorophyll content[38]. Shu et al. mapped Bin markers 849, located in chromosome 6, to be closely associated with fruit color genes that regulate ripening[36].Consistently, our study identified that the mutation site responsible for altering the color of sweet pepper and pepper is located on chromosome 6. However, the specific mutation sites differ, potentially due to the involvement of distinct pathways in regulating fruit color changes.In this study, we utilized SLAF-seq to identify the mutation loci responsible for the change

$-\log_{10}(p)$ value	SNP number	chromosome
>5	97	Chr01
	634	Chr02
	6	Chr04
	1	Chr05
	679	Chr06
	2	Chr07
	1	Chr09
	4	Chr10
	2	Chr11
>7	24	Chr01
	109	Chr02
	579	Chr06
>8	490	Chr06

Table 1. SNP loci are significantly associated with the color of pepper.

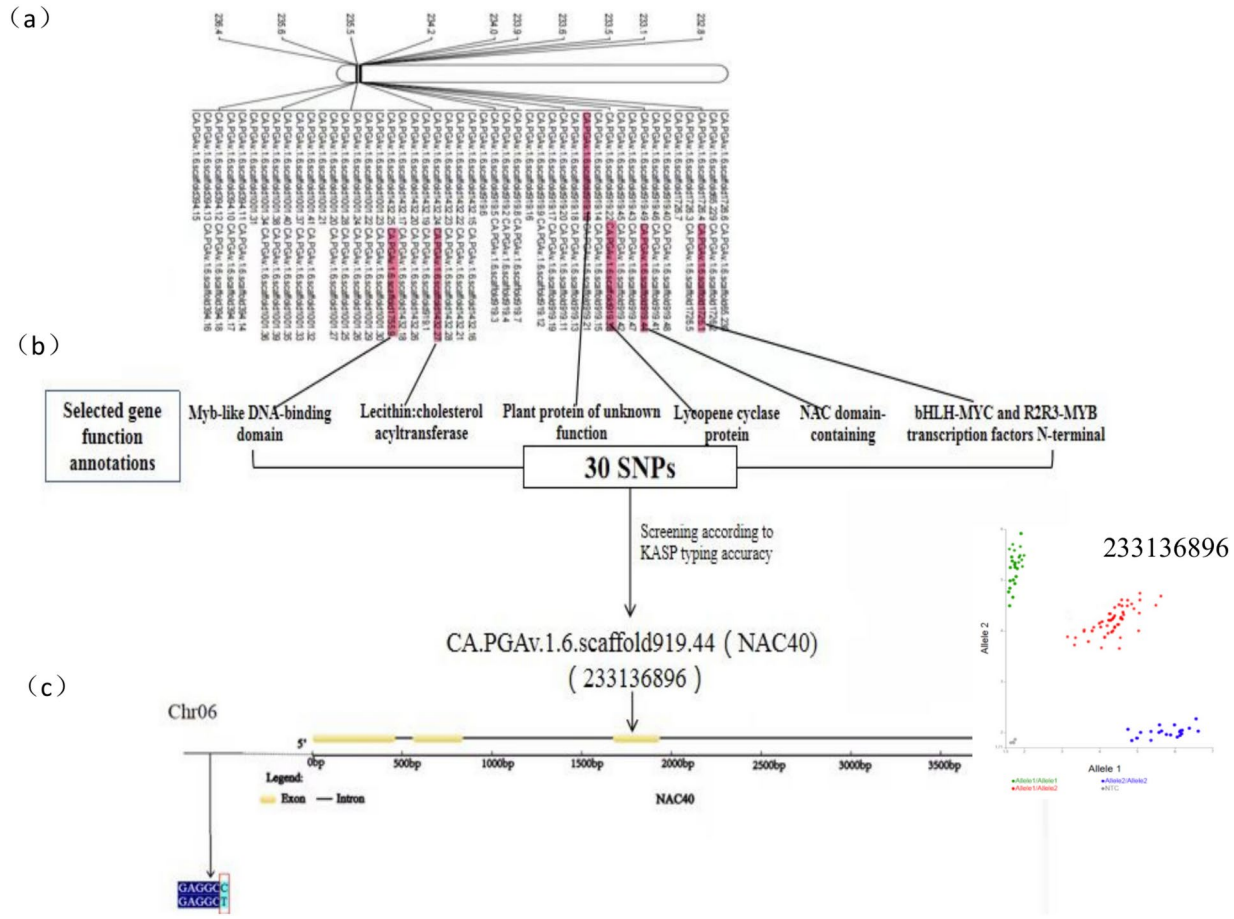

Fig. 4. The information on the candidate-associated genes. (**a**) The sixe candidate-associated genes and their function annotations. (**b**) KASP genotyping results in 233,136,896 mutations. (**c**) PGAv. 1.6 scaffold919.44 gene information graph.

in the physiological mature fruit color of sweet peppers and chili peppers, which were located on chromosome 6. This finding differs from the results of Wu et al. who reported mutations on a different chromosome. The discrepancy may be attributed to the distinct pepper materials used in the studies. Our results are consistent with the findings of Huangying Shu et al. who also identified mutation loci on chromosome 6. However, the specific mutation loci differ between the two studies, which may be due to different pathways regulating the change in fruit color.

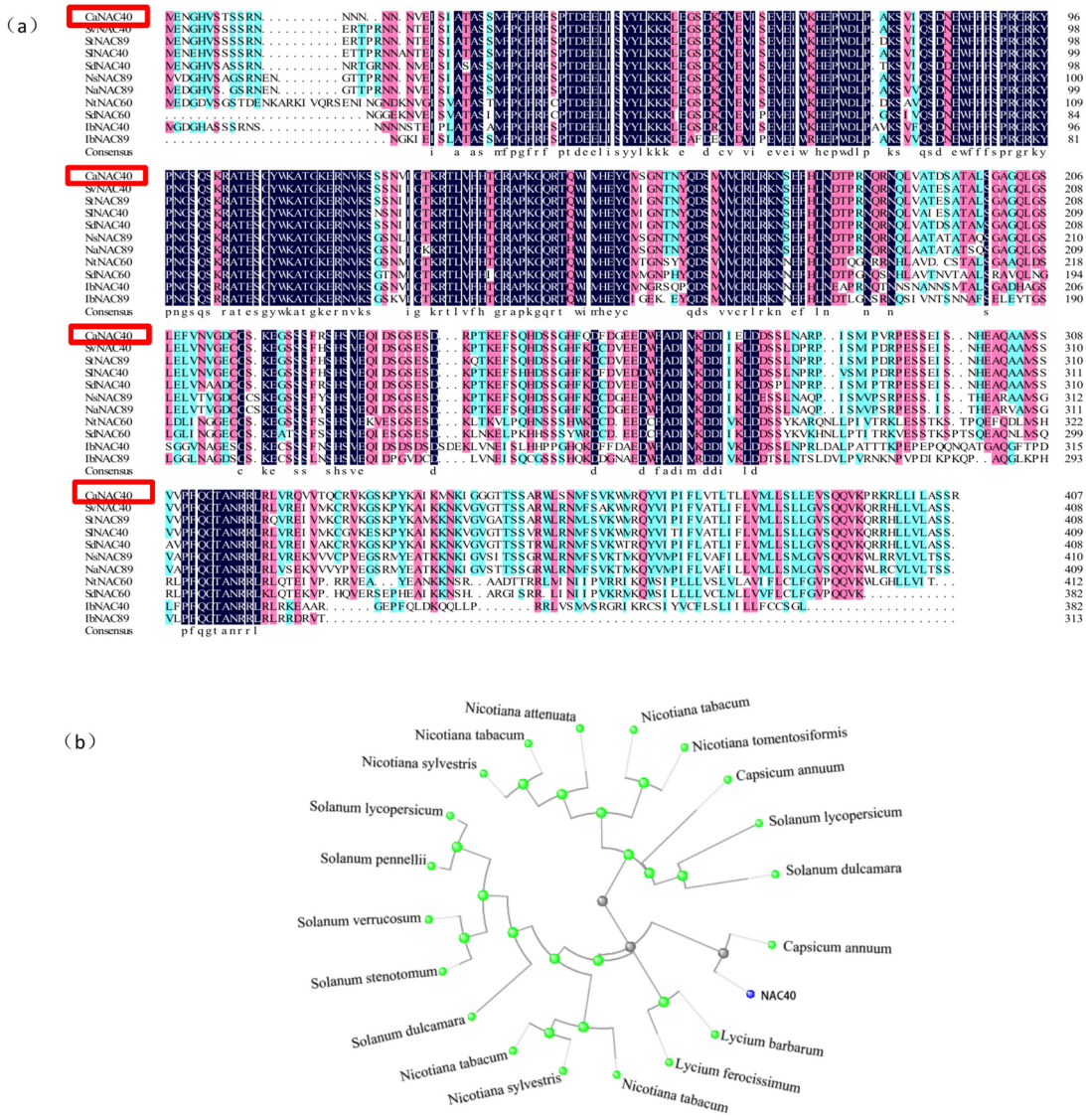

Fig. 5. NAC40 gene homologous relationship in different species (**a**) NAC40 gene homologous analysis in pepper, tomato, tobacco, and potato crops (**b**). NAC40 phylogenetic trees in various species. Red box: The red box represents NAC40 genes. Unknown: represents annotated with gene NAC40.

The phenomenon of point mutation leading to changes in fruit color has been documented in crops such as tomato[39], corn[40], cucumber[41], and eggplant[42], yet there is a scarcity of reports concerning sweet pepper and pepper. Our research demonstrated that the SNP associated with the color of sweet pepper and pepper is situated in the upstream region of the gene. This region contains numerous transcriptional regulatory elementsthat can interact with transcription factors to modulate gene transcription. Consequently, mutations in the upstream regulatory region may influence the binding of transcription factors and the transcriptional activity of genes[43,44]. For instance, the insertion of a Ca-nLTR-A reverse transcription transposon into the CaAN2 upstream regulatory region has been shown to induce the purple phenotype in pepper[45,46]. The Harbinger DNA transposon inserted into the promoter region of Pr-D (R2R3 MYB), which can affect the gene expression related to anthocyanin content in laver[47].The alteration in color observed in sweet pepper and pepper may be attributed to changes in the expression and function of the NAC40 gene, potentially induced by a base mutation in its upstream region. Our findings indicated that the NAC40 gene in pepperexhibits a high degree of homology with the NAC40 gene found in tomato, tobacco, and potato, and is classified within the NAC transcription factor gene family. The NAC transcription factor family is among the most significant gene families in plants, playing crucial roles in plant growth and development, secondary metabolite synthesis, and responses to biotic and abiotic stresses[48]. The NAC and C-terminal domains of the NAC protein are involved in protein binding activity, suggesting that the NAC domain may be critical in determining the protein's function[49,50]. In tomatoes, the NAC transcription factor NOR-like1 influences fruit coloration by directly regulating carotenoid accumulation and chlorophyll metabolism[51]. Similarly, in papaya, CPNAC1 may funcion as a positive regulator of carotenoid biosynthesis through the transcriptional activation of CpPDS2/4[52]. The final step in the carotenoid biosynthesis pathway

in pepper fruits is the conversion of anthocyanins to capsorubin and violaxanthin to capsorubin, which are catalyzed by capsanthin-capsorubin synthase (CCS)[53]. Previous studies have identified an NAC gene involved in the regulation of fruit color through its role in anthocyanin biosynthesis[54].

We hypothesized that mutations in the upstream regulatory regions of the CA.PGAv.1.6.scaffold919.44 gene may lead to alterations in the function of the NAC domain, thereby influencing anthocyanin synthesis and affecting the coloration of pepper and sweet pepper fruits. Subsequent investigations were conducted to elucidate the gene's function and underlying mechanisms. In summary, we identified 30 single nucleotide polymorphisms (SNPs) on Chromosome 6 that are significantly associated with pepper fruit color, corresponding to six associated genes, as determined by genome-wide association studies (GWAS). Using Kompetitive Allele Specific PCR (KASP) genotyping analysis, we precisely mapped genes related to pepper fruit color and preliminarily identified CA.PGAv.1.6.scaffold919.44 as a candidate gene potentially influencing fruit color through modulation of the NAC domain function. Our research offers novel insights into the molecular mechanisms underlying color formation in sweet pepper and pepper.

Data availability

Data is provided within the manuscript or supplementary information files.

Received: 18 September 2024; Accepted: 21 March 2025
Published online: 03 April 2025

References

1. Kim, S. et al. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nat. Genet.* **46**(3), 270–278 (2014).
2. Aguilar-Meléndez, A. et al. Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum; Solanaceae*) from Mexico. *Am. J. Bot.* **96**(6), 1190–1202 (2009).
3. Frank, C. A. et al. Consumer preferences for color, price, and vitamin C content of bell peppers. *HortTechnology* **36**(4), 795–800 (2001).
4. Wang, L. et al. Pigment biosynthesis and molecular genetics of fruit color in pepper. *Plants (Basel)* **12**(11), 2156 (2023).
5. Tomlekova, N. et al. Mutation associated with orange fruit color increases concentrations of β-carotene in a sweet pepper variety (*Capsicum annuum* L.). *Foods* **10**(6), 1225 (2021).
6. Brand, A. et al. CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit. *Theor. Appl. Genet.* **127**(10), 2139–2148 (2014).
7. Barry, C. S. et al. Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. *Plant Physiol.* **147**(1), 179–187 (2008).
8. Atkins, W. & Go Sherrard, A. XXV. The pigments of fruits in relation to some genetic experiments on capsicum annuum. In *The Scientific Proceedings of the Royal Dublin Society*. (The Society, 1915).
9. Deshpande, R. Studies in Indian chillies. (3) The inheritance of some characters in *Capsicum annuum L*. (1933).
10. Hurtado-Hernandez, H. & Smith, P. G. Inheritance of mature fruit color in *Capsicum annuum L. J. Hered.* **76**(3), 211–213 (1985).
11. Huh, J. H. et al. A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (Capsicum spp.). *Theor. Appl. Genet.* **102**, 524–530. https://doi.org/10.1007/s001220051677 (2001).
12. Su, W. et al. Genome-wide assessment of population structure and genetic diversity and development of a core germplasm set for sweet potato based on specific length amplified fragment (SLAF) sequencing. *PLoS ONE* **12**(2), e0172066 (2017).
13. Sun, X. et al. SLAF-seq: An efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE* **8**(3), e58700 (2013).
14. Xie, D. et al. Genome-wide association study identifying candidate genes influencing important agronomic traits of flax (*Linum usitatissimum* L.) using SLAF-seq. *Front. Plant Sci.* **8**, 2232 (2017).
15. Zhao, T. et al. Mapping and candidate gene screening of tomato Cladosporium fulvum-resistant gene Cf-19, based on high-throughput sequencing technology. *BMC Plant Biol.* **16**, 51 (2016).
16. Li, T. et al. High-resolution detection of quantitative trait loci for seven important yield-related traits in wheat (*Triticum aestivum* L.) using a high-density SLAF-seq genetic map. *BMC Genom. Data* **23**(1), 37 (2022).
17. Xu, X. et al. A high-density genetic map of cucumber derived from specific length amplified fragment sequencing (SLAF-seq). *Front. Plant Sci.* **5**, 768 (2014).
18. Wei, Q. et al. Construction of a SNP-based genetic map using SLAF-Seq and QTL analysis of morphological traits in eggplant. *Front. Genet.* **11**, 178 (2020).
19. Li, N. et al. Construction of a high-density genetic map and identification of QTLs for cucumber mosaic virus resistance in pepper (*Capsicum annuum* L.) using specific length amplified fragment sequencing (SLAF-seq). *Breed Sci.* **68**(2), 233–241 (2018).
20. Zhang, X.-F. et al. High-density genetic map construction and QTL mapping of first flower node in pepper (*Capsicum annuum* L.). *BMC Plant Biol.* **19**, 1–13 (2019).
21. Healey, A. et al. Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. *Plant Methods* **10**, 21 (2014).
22. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**(5), 589–595 (2010).
23. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**(16), 2078–2079 (2009).
24. Raman, H. et al. Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant Cell Environ.* **39**(6), 1228–1239 (2016).
25. Kang, H. M. et al. Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**(4), 348–354 (2010).
26. Ginestet, C. ggplot2: Elegant graphics for data analysis. *J. Roy. Stat. Soc.* **174**(1), 245–246 (2011).
27. Turner, S. D. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. Biorxiv, 2014.
28. Yangyang, D. et al. Integrated nr database in protein annotation system and its localization. *Comput. Eng.* **32**(5), 71–72 (2006).
29. Ashburner, M. et al. Gene Ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25 (2000).
30. Tatusov, R. L. et al. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**(1), 33–36 (2000).
31. Kanehisa, M. et al. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**(Database issue), D277-280 (2004).
32. Zhang, Y. et al. Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. *BMC Plant Biol.* **13**, 141 (2013).

33. Zhu, W. Y. et al. A high-density genetic linkage map for cucumber (*Cucumis sativus* L.): Based on specific length amplified fragment (SLAF) sequencing and QTL analysis of fruit traits in cucumber. *Front. Plant Sci.* **7**, 437 (2016).
34. Li, B. et al. Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application to QTL analysis for isoflavone content in Glycine max. *BMC Genomics* **15**(1), 1086 (2014).
35. Tao, J. et al. Construction of a high-density genetic map based on specific-locus amplified fragment sequencing and identification of loci controlling anthocyanin pigmentation in Yunnan red radish. *Hortic. Res.* https://doi.org/10.1093/hr/uhab031 (2022).
36. Shu, H. et al. Fine mapping and identification of candidate genes for fruit color in pepper (Capsicum chinense). *Sci. Hortic.* **310**, 111724 (2023).
37. Zhang, Y. et al. Transcriptome and metabolome analysis of color changes during fruit development of pepper (Capsicum baccatum). *Int. J. Mol. Sci.* **23**(20), 12524 (2022).
38. Wu, L. et al. Mapping of CaPP2C35 involved in the formation of light-green immature pepper (*Capsicum annuum* L.) fruits via GWAS and BSA. *Theor. Appl. Genet.* **135**(2), 591–604 (2022).
39. Liu, X. et al. Association and genetic identification of loci for four fruit traits in tomato using InDel markers. *Front. Plant Sci.* **8**, 1269 (2017).
40. Owens, B. F. et al. Genome-wide association study and pathway-level analysis of kernel color in maize. *G3: Genes, Genomes, Genet.* **9**(6), 1945–1955 (2019).
41. Huang, H. et al. Genome-wide association analysis reveals a novel QTL CsPC1 for pericarp color in cucumber. *BMC Genom.* **23**(1), 383 (2022).
42. Cericola, F. et al. Linkage disequilibrium and genome-wide association analysis for anthocyanin pigmentation and fruit color in eggplant. *BMC Genom.* **15**, 1–15 (2014).
43. Wu, Z. et al. Genome-wide identification and expression profile of Dof transcription factor gene family in pepper (*Capsicum annuum* L.). *Front. Plant Sci.* **7**, 574 (2016).
44. Römer, P. et al. Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper Bs3 alleles. *Plant Physiol.* **150**(4), 1697–1712 (2009).
45. Li, N. et al. Identification of CaPs locus involving in purple stripe formation on unripe fruit, reveals allelic variation and alternative splicing of R2R3-MYB transcription factor in pepper (*Capsicum annuum* L.). *Front. Palnt Sci.* **14**, 1140851 (2023).
46. Ohno, S., Ueno, M. & Doi, M. Differences in the CaMYBA genome between anthocyanin-pigmented cultivars and non-pigmented cultivars in pepper (*Capsicum annuum*). *Hort. J.* **89**(1), 30–36 (2020).
47. Chiu, L.-W. et al. The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol.* **154**(3), 1470–1480 (2010).
48. Liu, G. S. et al. NAC transcription factor family regulation of fruit ripening and quality: A review. *Cells* **11**(3), 525 (2022).
49. Zhu, M. et al. A new tomato NAC (N AM/A TAF1/2/C UC2) transcription factor, SlNAC4, functions as a positive regulator of fruit ripening and carotenoid accumulation. *Plant Cell Physiol.* **55**(1), 119–135 (2014).
50. Kou, X. et al. NAC transcription factors play an important role in ethylene biosynthesis, reception and signaling of tomato fruit ripening. *Mol Genet Genom.* **291**, 1205–1217 (2016).
51. Gao, Y. et al. A NAC transcription factor, NOR-like1, is a new positive regulator of tomato fruit ripening. *Hortic. Res.* https://doi.org/10.1038/s41438-018-0111-5 (2018).
52. Fu, C.-C. et al. The papaya transcription factor CpNAC1 modulates carotenoid biosynthesis through activating phytoene desaturase genes CpPDS2/4 during fruit ripening. *J. Agric. Food Chem.* **64**(27), 5454–5463 (2016).
53. Bouvier, F. et al. Xanthophyll biosynthesis in chromoplasts: Isolation and molecular cloning of an enzyme catalyzing the conversion of 5, 6-epoxycarotenoid into ketocarotenoid. *Plant J.* **6**(1), 45–54 (1994).
54. Wang, J. et al. Overexpression of BoNAC019, a NAC transcription factor from Brassica oleracea, negatively regulates the dehydration response and anthocyanin biosynthesis in Arabidopsis. *Sci. Rep.* **8**(1), 13349 (2018).

Author contributions

Each author is expected to have made substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data; or the creation of new software used in the work; or have drafted the work or substantively revised it and to have approved the submitted version (and any substantially modified version that involves the author's contribution to the study);

Funding

This work was supported by S&T Program of Hebei (No.21326309D-5) , Project of Modern Agricultural Industry Technology System in Hebei Province (No.HBCT2023100209) , HAAFS Science and Technology Innovation Special Project (No.2022KJCXZX-JZS-6) ,HAAFS Science and Technology Innovation Special Project (No.2023KJCXZX-JZS-12).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-95552-3.

Correspondence and requests for materials should be addressed to Z.Z., Y.F. or L.Y.