BMC Genomics

RESEARCH

Open Access

SLAF-seq technology-based genome-wide association and population structure analyses of hot pepper and sweet pepper

Yaning Meng[1†], Hongxiao Zhang[1,3†], Zhe Zhang[1], Xinxin Li[1], Zhanghong Yu[1], Yanqin Fan[1*] and Libin Yan[1,2*]

Abstract

Background Utilizing Single Nucleotide Polymorphism (SNP) marker technology, a phylogenetic and agronomic trait network analysis was conducted on the collected hot pepper and sweet pepper germplasm resources, providing a theoretical basis for parental selection and new varieties.

Results Specific-locus amplified fragment sequencing (SLAF-seq) technology was employed for a genome-wide association study (GWAS) on 197 hot pepper and sweet pepper germplasm resources, generating 1404.88 Mb clean reads data with an average Q30 of 91.5% and mean GC content of 37.96%. Through sequencing data analysis, a total of 639,815 SLAF tags were obtained with an average sequencing depth of 12.16x, among which 86,381 were polymorphic SLAF tags, leading to the development of 18,145,155 SNP markers. The identified SNP markers were used for cluster analysis of the genetic structure and phylogenetic relationships of hot pepper and sweet pepper germplasm resources, dividing the 197 hot pepper and sweet pepper germplasm resources into 9 clusters. Additionally, a genome-wide association analysis was conducted on 25 agronomic traits of the 197 hot pepper and sweet pepper materials, yielding a substantial number of significantly associated SNP loci with agronomic traits. A correlation network analysis diagram was drawn among the various agronomic traits, preliminarily determining the relationships between the 25 agronomic characteristics of hot pepper and sweet pepper and the positions of 15 agronomic traits ($p < 1.707 \times 10^{-8}$) on the chromosomes were annotated, forming multi-trait aggregation regions.

Conclusions Our research reveals the genetic diversity, phylogenetic relationships, and population structure of 197 hot pepper and sweet pepper germplasm resources, providing a basis for germplasm identification, resource utilization, and breeding.

Keywords Hot pepper and sweet pepper, Agronomic traits, Genome-wide association study (GWAS), Specific-locus amplified fragment sequencing (SLAF-seq)

[†]Yaning Meng and Hongxiao Zhang are co-first authors.

*Correspondence:
Yanqin Fan
nkyfanyq@126.com
Libin Yan
tianjiaoshi@126.com
Full list of author information is available at the end of the article

Introduction

The hot pepper and sweet pepper *(Capsicum annuum* L.*)*, belonging to the family Solanaceae and genus Capsicum, is an annual or perennial crop and one of the most important vegetable crops worldwide. As one of the primary condiment vegetables, peppers are favored for their unique flavor and rich nutritional value. Identifying and describing field agronomic traits are the most fundamental methods and pathways for pepper germplasm resources. They are the primary concern for researchers in cultivation and breeding [1, 2]. The genetic diversity, phylogenetic relationships, and population structure of pepper germplasm resources can reveal their genetic characteristics, providing references for studying the origin and evolution of different pepper types and offering a basis for germplasm identification, resource utilization, and breeding.

Traditional methods for studying pepper germplasm resources' genetic diversity and evolutionary origins are susceptible to various factors, such as environmental conditions and sampling parts. However, applying molecular biology techniques can overcome these influences, interpreting the genetic diversity of germplasm at the gene level [3, 4]. Genome-wide association studies (GWAS) are extensively used to identify relationships between molecular markers or candidate genes and phenotypes of interest. This technique uses high-density SNP markers, correlates molecular polymorphisms with a phenotype, and compared with traditional linkage analysis, this method has the advantages of more accurate location and mapping, simultaneous assessments of multiple alleles at a locus, and no requirements for linkage group construction [5, 6]. The method has been successfully applied to identify candidate genes controlling complex agronomic traits in various plant species, including tomatoes, cucumbers, chickpeas, cabbage, and alfalfa [7–11]. There are few reports on the application of GWAS in hot and sweet peppers, and most focus on one or a few agronomic traits. Nimmakayala et al. utilized GWAS association analysis to identify 16 SNP loci related to single fruit weight (Wpf) and 36 SNP loci related to fruit pedicel length (Fpl) [12, 13]. Ahn et al. also located candidate genes related to powdery mildew resistance on chromosome 4 through GWAS association analysis [14].

Specific-locus amplified fragment sequencing (SLAF-Seq) technology is a simplified genome sequencing technique that uses restriction enzymes to fragment the genome, conducting high-throughput sequencing on specific regions to obtain a large number of genetic polymorphic tag sequences, thereby revealing the entire genomic sequence characteristics. SLAF-Seq is primarily used to construct high-density genetic maps in natural and biparental populations, predict genome structure, and Quantitative Trait Loci (QTL) mapping of important agronomic traits [15, 16]. Genome-wide association analysis based on SLAF-seq sequencing technology for crop field agronomic traits provides a foundation for the study population's systematic evolution, germplasm resource evaluation, and molecular breeding. Currently, this technology has been widely applied in the development of genetic markers, construction of genetic maps, and target gene localization in various plants such as tomato [17], rapeseed [18], ramie [19], soybean [20], and cotton [21].

This study used 197 hot peppers and sweet peppers germplasm resources as experimental materials. SLAF-Seq technology was used to conduct Genome-Wide Association and PopulationStructure Analyses of hot and sweet peppers. This study explores the genetic background and phylogenetic relationships of the sweet (hot) pepper germplasm resources preserved by our research group, laying the foundation for classifying and identifying the 197 hot peppers and sweet peppers germplasm resources. It also provides a scientific reference for collecting, preserving, developing, and utilizing hot pepper and sweet pepper germplasm resources and selecting hybrid combinations.

Materials and methods

Materials

In our study, we utilized 197 hot peppers and sweet peppers accessions provided by the Institute of Cash Crops, Hebei Academy of Agriculture and Forestry Sciences, to comprehensively explore correlations among the 25 agronomic traits and subsequent GWAS analysis. The 25 agronomic traits were surveyed during different growth stages and were categorized into four plant tissues: stem (axillary bud formation, stem villi, the height of the main stem, first flower node, plant height), leaves (cotyledon color, fluff on leaves, blade length, blade width, petiole length), flower (inflorescence morphology, flower color, style color, stigma growth, male sterility), and fruit (flesh thickness, single fruit weight, mature fruit color, fruit-bearing form, fruit length, transverse diameter of fruit, fruit shape, fruit brightness, ventral number). Correlations among the 25 agronomic traits were analyzed using the R software based on phenotypic data (Table S1).

DNA extraction and detection

A DNA extraction kit from Chengdu Fuji Biotech Co., Ltd.was used to extract genomic DNA from the tender leaves of the experimental materials. The concentration was detected using a micro-ultraviolet spectrophotometer, and the extracted DNA solution concentration was 100 ng/μl, stored at −20 °C for future use.

Enzymatic digestion and library construction

In this study, we used the reference genome of C. annuum cv. CM334, with a size of 3.48 Gb (http://peppergenome.snu.ac.kr/.). Genomic DNA of 197 hot peppers and sweet peppers was incubated at 37 °C with 0.6 units MseI, T4 DNA ligase (NEB), ATP (NEB), and *MseI* adapters. Restriction–ligation reactions were heat-inactivated at 65 °C and then digested with *HaeIII* at 37 °C. The PCR was performed using diluted restriction–ligation samples, dNTP, *Taq* DNA polymerase (NEB), and MseI primer containing barcode 1. The PCR productions were purified using E.Z.N.A.H® Cycle Pure Kit (Omega, London, UK) and pooled. The pooled sample was incubated at 37 °C with *MseI*, T4 DNA ligase, ATP, and the Solexa adapter. After incubation, the sample was purified using a Quick Spin column (Qiagen, Hiden, Germany) and electrophoresed through a 2% agarose gel. After gel purification, DNA fragments with indices and adaptors (SLAFs) of 550–600 bp were excised and diluted for paired-end sequencing. SLAF-seq was performed using the Illumina High-seq2500 sequencing platform.

Sequencing and data quality control

The dual-index was applied to identify the raw reads obtained by sequencing to find the reads of the samples. The adaptor-filtered reads were evaluated for the sequenced quality score and data size, in which the sequenced quality score was an important indicator to evaluate the error rate of the high-quality single base. Real-time monitoring was performed for each cycle during sequencing, the ratio of high quality reads with quality scores greater than Q30 (means a quality score of 30, indicating a 1% chance of an error, and thus 99% confidence) in the raw reads and guanine-cytosine (GC) content were calculated for quality Control [22]. Afterwards, high-quality reads were mapped onto the reference genome of pepper using BWA software [23], and then the paired-end mapped reads located at the same position with more than 95% identity were grouped into one SLAF locus.

SLAF tag and SNP marker development

Cluster analysis was conducted based on sequence similarity on the 197 hot peppers and sweet peppers germplasm resources. The SLAF tags were mapped to the reference genome using the BWA software [23].

Cluster analysis was conducted based on sequence similarity on the 197 hot peppers and sweet peppers germplasm resources. Raw reads were sorted to each material according to duplex barcode sequences. After the barcodes and the terminal 5-bp positions were trimmed, high-quality reads from the same samples were mapped onto the reference genome sequence using the BWA software. All SLAF pair-end reads with clear index information were clustered together, based on sequence similarity which was detected using one-to-one alignment by BLAT. Sequences located at the same position with over 95% identity were grouped into one SLAF locus. Alleles among the materials were defined in each SLAF by the minor allele frequency (MAF) evaluation. True genotypes had markedly higher MAF values than genotypes containing sequence errors. SLAFs filtration and genotype definition were performed according to the method described by Sun et al. [24], and all SLAFs had been filtered and quality assessed many times. SLAFs containing more than four tags were filtered out as repetitive SLAFs. Because pepper is diploid species, one locus can only contain at most four SLAF tags. Only SLAFs with two to four alleles were identified as polymorphic and considered potential markers. All polymorphism SLAFs loci were genotyped with consistency in the parental and offspring SNP loci. The SLAFs with > 3 SNPs were filtered out. It is considered to be a high frequency variation region of sequencing, if SLAFs have more than 3 SNPs. SLAF markers with average sequence depths of less than 10-fold in parents were discarded.

The number of SLAF tags and polymorphic SLAF tags on different chromosomes was counted. Two methods, GATK [25] and SAMtools [26] were used to develop SNP markers, and the intersection of SNP markers obtained by both methods was considered the final reliable set of SNP markers. All SNPs were filtered based on completeness > 0.8 and minor allele frequency (MAF) > 0.05. The SnpEff [27] software was used for variant site annotation (SNP, Small InDel) and prediction of variant effects. Based on the position of the variant sites on the reference genome and the gene location information on the reference genome, the region where the variant occurred in the genome and the impact of the variation could be determined.

Population structure analysis

The population structure of 197 accessions was calculated [28] using by using the admixture software. The analysis utilized the SNPs of 197 accessions to infer the genetic background of an accession belonging to a cluster under a given number of subgroups (K). For all accessions, the number of genetic clusters was predefined (K = 1–15) to explore the population structure. The relatedness coefficient was calculated with SPAGeDi (version 1.4C) software [29]. Linkage disequilibrium between SNPs was estimated using R^2 using HAPLOVIEW software [30].

Phylogenetic tree analysis

Based on the selected polymorphic SNPs, a phylogenetic tree of the hot pepper, sweet pepper, and peppers germplasm was constructed using the Neighbor-joining algorithm in MEGA X [31, 32], with Bootstrap set to 1000. The branch lengths of the phylogenetic tree reflected the genetic distances among samples, with shorter lengths indicating closer phylogenetic relationships.

Genome-wide association analysis of 25 agronomic traits

Total filtered SNPs detected from 197 accessions were used for GWAS. GWAS for all traits (based on LMM, LM, FaST-LMM and EMMAX models) was conducted using GEMMA (https://github.com/genetics-statistics/GEMMA), FaST-LMM (https://github.com/fastlmm/FaST-LMM), and EMMAX (http://csg.sph.umich.edu//kang/emmax/download/index.html.) software with default settings. The genes within 100 kb up or downstream of these significant SNPs were found by Jbrowser in SGN (Sol Genomics Network https://solgenomics.net/) and reported as potential candidate genes.

Results and analysis

Library construction and sequencing quality assessment

Using the SLAF-predict software and referring to the pepper genome for electronic enzyme digestion prediction, the RsaI + HaeIII enzyme combination was selected for digestion. Sequences with lengths between 550–600 bp were defined as SLAF tags. A total of 1404.88 Mb of reads were obtained in the experiment, with an average sequencing Q30 of 91.5% for all samples; the average GC content was 37.96%, which was generally low but met the sequencing requirements (Table S2).

SLAF tag and polymorphic SNP marker screening

Six hundred thirty-nine thousand eight hundred fifteen SLAF tags were obtained from sequencing 197 hot peppers and sweet peppers germplasm resources, with an average sequencing depth of 12.16x. Among these, there were 86,381 polymorphic SLAF tags, and a total of 18,145,155 SNPs were obtained (Table S3-S4). All SNPs were filtered based on minor allele frequency MAF > 5% and a missing rate < 20% to select highly consistent population SNPs. The number of polymorphic SLAF tags and SNP markers on different hot pepper and sweet pepper chromosomes was counted, and a chromosome distribution map was drawn based on the data (Fig. 1). The developed polymorphic SLAF tags and SNP sites were evenly distributed across the 12 chromosomes of the pepper genome, indicating expected sequencing results and readiness for further analysis.

Population genetic structure analysis

Based on the developed sweet pepper SNP sites, the population structure of 197 hot and sweet peppers germplasm resources was analyzed using the Admixture software. Figure 2 and Table 1 show that the genetic structure among the 197 hot peppers and sweet peppers germplasm resources was relatively straightforward and could be divided into nine categories. The first group included 22 pepper materials, such as L108, L109, and L113; the second group included 39 sweet pepper materials, such as L130, L143, and L155; the third group included ten sweet pepper materials, such as L10, L12, L2; the fourth group included 14 pepper materials such as L1, L107, L140; the fifth group included 17 sweet pepper materials such as L11, L129, L135; the sixth group included 19 sweet pepper materials such as L100, L101, L102; the

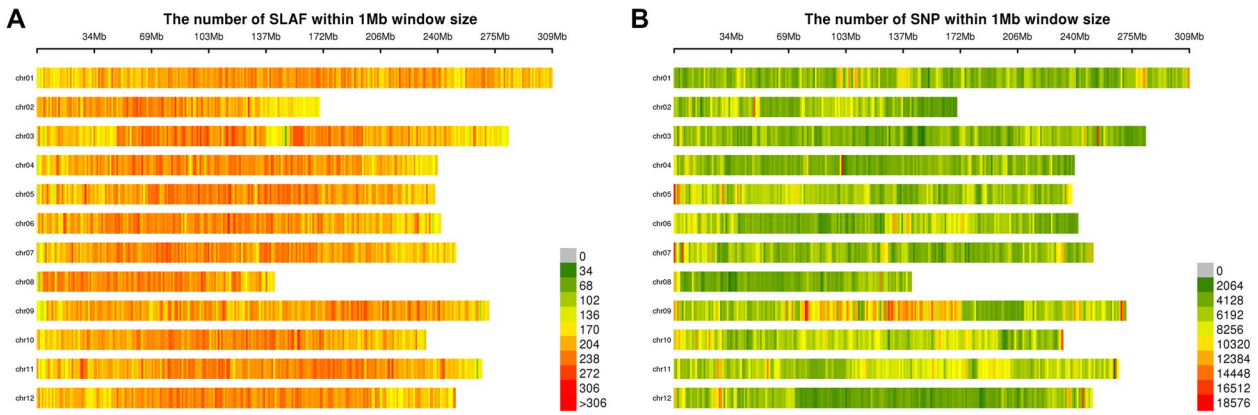

Fig. 1 Distribution of SLAF tags (**A**) and SNP (**B**) on the 12 chromosomes of the pepper reference genome. The horizontal axis represents the length of the chromosome, with each bar representing a chromosome. The genome has been divided into segments of 1 Mb in size. The darker the color within each window, the greater the number of SLAF tags and SNP markers; conversely, the lighter the color, the fewer the SLAF tags and SNP markers

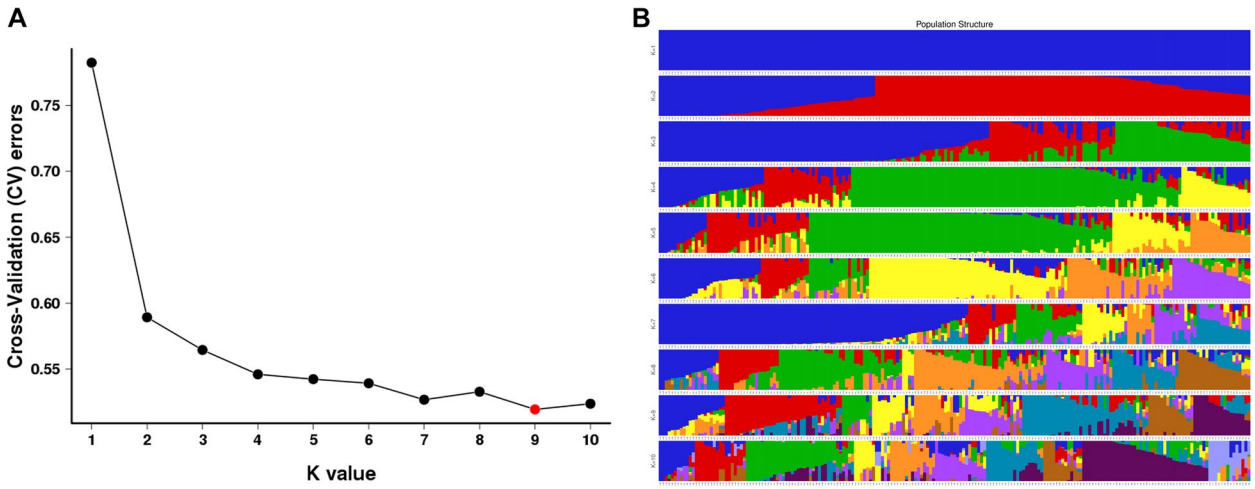

Fig. 2 Population structure analysis of 197 sweet pepper germplasm resources. **A** Cross-validation error rates corresponding to different K values; **B** Analysis of population structure by using the identified SNPs

Table 1 197 Sweet pepper and pepper germplasm resources population structure classification

Group	Count	Pepper number
Q1	22	L108、L109、L113、L123、L126、L128、L131、L134、L136、L137、L138、L139、L142、L149、L150、L153、L34、L50、L51、L52、L68、T132
Q2	39	L130、L143、L155、L94、T109、T130、T131、T133、T135、T137、T139、T14、T140、T141、T26、T29、T31、T36、T37、T51、T54、T58、T62、T7、T72、T94、T98、TZ1、TZ10、TZ14、TZ15、TZ16、TZ3、TZ4、TZ5、TZ6、TZ7、TZ8、TZ9
Q3	10	L10、L12、L2、L3、L42、L43、L49、L6、L9、T18
Q4	14	L1、L107、L140、L145、L146、L148、L24、L45、L76、L79、L80、L81、L86、L93
Q5	17	L11、L129、L135、L144、L15、L151、L17、L19、L22、L28、L30、L31、L33、L4、L96、LZ8、T91
Q6	19	L100、L101、L102、L21、L23、L53、L62、L69、L72、L74、L77、L85、L87、L90、L92、L97、L99、LZ10、T142
Q7	42	L13、L47、L98、T1、T10、T11、T128、T13、T144、T145、T146、T15、T2、T21、T22、T24、T3、T32、T38、T39、T4、T40、T42、T44、T45、T47、T48、T5、T50、T53、T59、T61、T64、T70、T71、T8、T9、T90、T95、T97、TZ2
Q8	15	L16、L20、L38、L39、L40、L55、L7、L78、L8、LZ1、LZ3、LZ4、LZ5、LZ6、LZ9
Q9	19	T110、T115、T12、T121、T122、T124、T16、T17、T19、T20、T23、T25、T27、T28、T43、T49、T55、T56、T74

seventh group included 42 sweet pepper materials such as L13, L47, L98; the eighth group included 15 pepper materials such as L16, L20, L38; and the ninth group included 19 sweet pepper materials such as T110, T115, T12. Some sweet pepper germplasm resources were not classified, indicating that some peppers and sweet pepper resources in the 197 resources share common parental blood relationships (Table 1).

Genetic relationship cluster analysis

Based on the developed SNPs (Table S5) identified from the 197 hot peppers and sweet peppers materials, the phylogenetic relationships was constructed in MEGA 5.0 [24] using the neighbor-joining algorithm [33]. A genetic relationship clustering diagram (Fig. 3) was constructed based on the distance matrix. The results showed that the resources of 197 hot peppers and sweet peppers could be divided into 9 clusters, consistent with the population structure analysis results. The phylogenetic tree clearly showed the closeness of the phylogenetic relationships among the 197 hot peppers and sweet peppers resources, providing a reference for the selection of parents in breeding new sweet pepper varieties (Fig. 3).

Global exploration of correlations among the 25 agronomic traits

The correlations among the 25 traits were explored using correlation network analysis (Fig. 4). Among the stem-related traits, HMSOP was positively correlated with PH (0.62), and ABF was positively correlated with FFN (0.84). Among the leaf-related traits, BL was positively correlated with BW (0.75). Among the fruit-related traits, FS was positively correlated with FL (0.73), SFW was positively correlated with TDF (0.85), FS was negatively

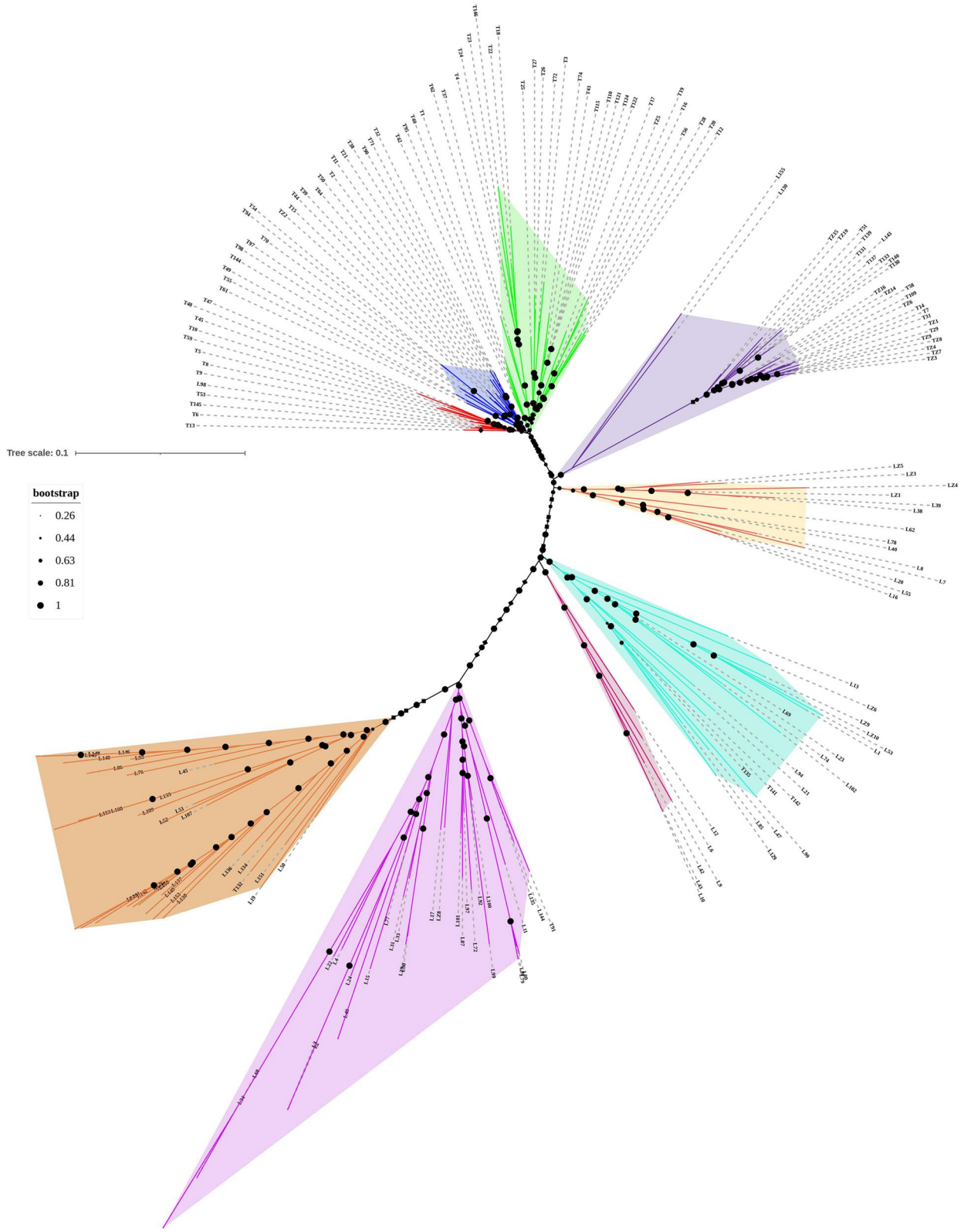

Fig. 3 Phylogenetic tree of 197 sweet pepper and hot pepper germplasm resources

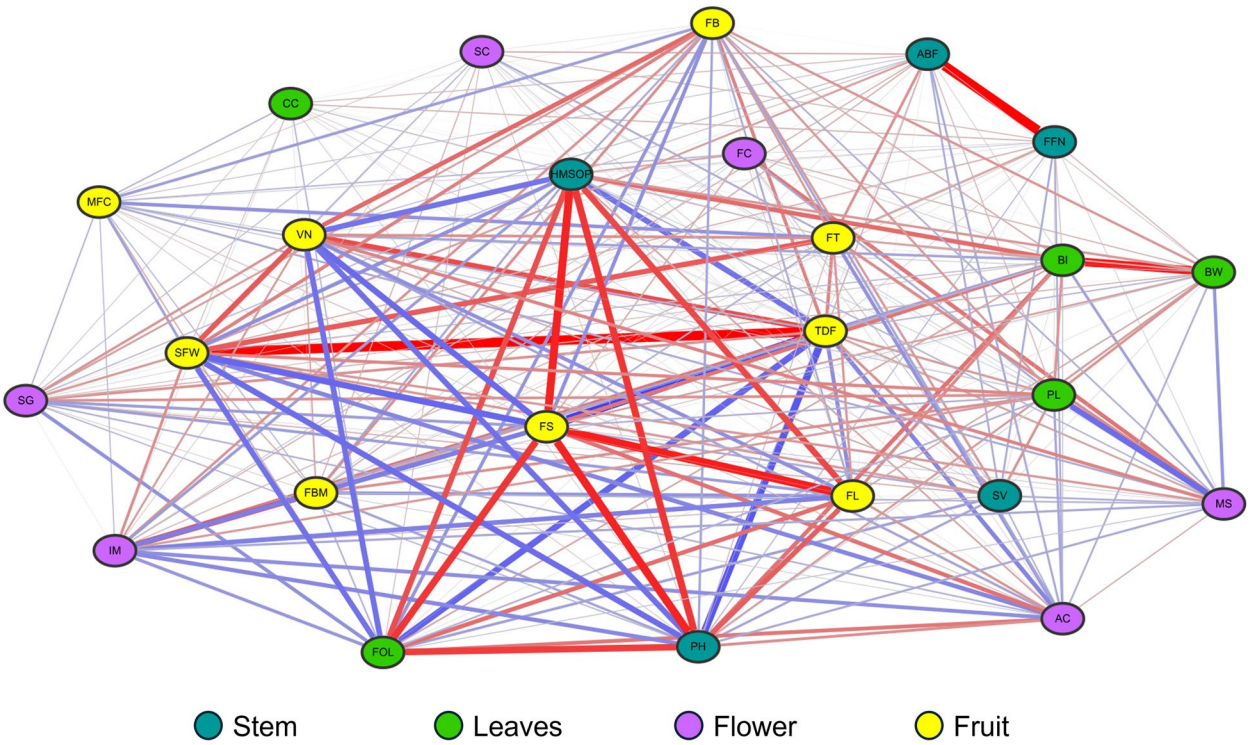

Fig. 4 Phenotype correlation network of agronomic traits in pepper. The red and blue lines represent positive and negative correlations, respectively. Line width is proportional to the strength of the correlation. Different color nodes represent stem-related, leaf-related, flower-related, and fruit-related traits, respectively

correlated with VN (−0.54), FS was negatively correlated with SFW (−0.56), and FS was negatively correlated with TDF (−0.79). The correlation coefficients between each pair of traits are shown in Table S6.

Large-scale GWAS for 25 agronomic traits

The highly consistent filtered SNPs in the 197 hot peppers and sweet peppers accessions were used for GWAS analysis of 25 agronomic traits. To reduce the influence of population structure and increase the reliability of GWAS results, four statistical models were used: efficient mixed-model association expedited (EMMAX), factored spectrally transformed linear mixed model (FaST-LMM), general linear model (GLM), and linear mixed model (LMM). The performance of the four models was compared based on Q-Q plots (Figure S1), and the most appropriate statistical model for each trait was selected for subsequent GWAS analysis. As shown in Table S7, EMMAX was the optimal model for 13 traits; Fast-LMM was the optimal model for eight traits; LMM was the optimal model for four traits. These results showed that the optimal models for different traits can vary, implying that it was necessary to select one optimal model for each trait for GWAS analysis.

In total, 726 SNPs were identified with $p < 1.707 \times 10^{-8}$ as the highly significant threshold from the optimal models, and 846 candidate genes in the 100 kb region upstream or downstream near those SNPs were detected for 15 traits (Table S7). Traits such as Flower color, Male sterile, Style color, Anther color, and Cotyledon color formed aggregation areas where multiple agronomic traits could be controlled, laying the foundation for future research on the pleiotropy of genes (Fig. 5).

Discussion

Peppers have rich germplasm resources and a long cultivation history in China. With the continuous development of China's pepper industry and the increase in consumable products, the demand for new varieties of hot peppers and sweet peppers with different levels of spiciness and uses is growing. Therefore, the collection and evaluation of sweet and hot pepper germplasm resources and the breeding of superior germplasm have become vital issues that need to be addressed for the sustainable development of hot and sweet peppers.

The study of gene localization for pepper agronomic traits is mainly based on traditional QTL mapping methods, which only study single or multiple traits in the same study [34–37]. GWAS can examine many agronomic

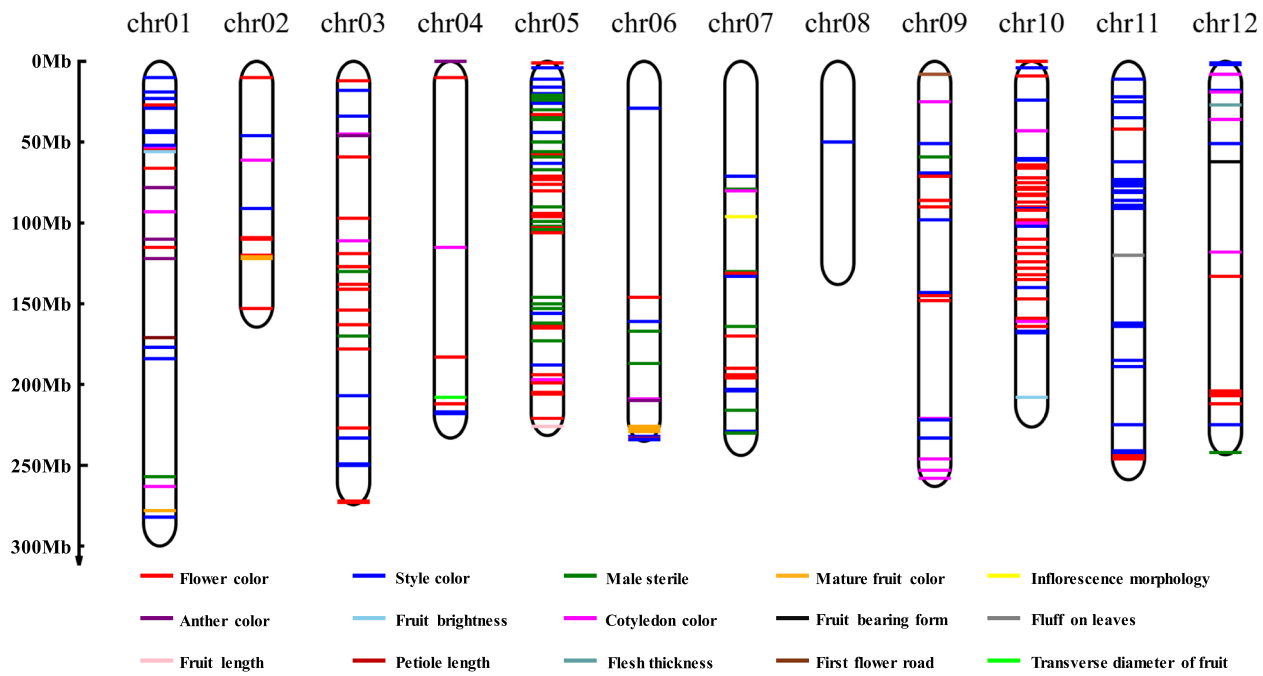

Fig. 5 The position of 15 agronomic traits on each chromosome

traits simultaneously, significantly improving the efficiency of gene localization for pepper agronomic traits. With the development of DNA sequencing technology, the SLAF-seq technology has been widely used in plants due to its advantages of rapid, low-cost, and high-throughput development of SNP markers [24, 38, 39]. Domestic and international researchers have successfully applied this technology in crops such as onion, wheat, tea, soybean, flax, forsythia and sunflower [40–46]. Lee et al. the GWAS of cross-validated results using 17 fruits species as materials. They identified 16 candidate genes associated domestication traits related to fruit morphology, which are linked to molecular functions such as cell division and proliferation [47]. Ghulam Mustafa Wassan et al. identified 142 significant SNP markers closely related to salt tolerance by conducting a genome-wide association study on 228 rapeseed germplasm resources to identify genetic variations linked to differential responses to salt stress, among which 78 SNPs were located on the C genome and 64 SNPs on the A genome [48]. In this study, through SLAF-seq sequencing technology, genetic diversity analysis and development of SNP molecular markers were conducted on 197 hot peppers and sweet peppers germplasm resources. A total of 639,815 SLAF tags were obtained, among which 86,381 polymorphic SLAF tags, and 18,145,155 population SNPs were used for genetic background analysis.

Utilizing SNP molecular markers for cluster analysis of the genetic structure and phylogenetic relationships of hot peppers and sweet peppers, the 197 hot peppers and sweet peppers germplasm resources were divided into 9 clusters, and a genetic relationship phylogenetic tree was drawn, clarifying the phylogenetic relationships of the 197 materials.Interestingly, hot pepper were clustered together with sweet pepper, which showed that their evolutionary relationship was very close. No obvious relationship was detected between the accessions and their geographic origins. Indicate these lines might have undergone introgression or gene flow during pepper breeding in China. This finding was consistent with those in previous reports on pepper [49, 50].

The selection of best statistical models can accurately evaluate the associations between markers and phenotypes. As the availability of genotype data increases, comprehensive statistical models are required to distinguish true biological associations from the false positives arising from population structure and LD [51]. Numerous statistical models are available to calculate the significance of the associations between SNPs and traits [52]. In this study, Four analysis models (EMMAX,FaST-LMM,GLM and LMM) were used and the best models were selected for different traits. Effectively controlled the false positive rate, and provided suitable statistical power to identify significant marker-trait associations. A large number of significantly associated SNP loci with agronomic traits were obtained. A correlation network analysis diagram was drawn between various agronomic traits, preliminarily determining the relationships

between the 25 agronomic traits of hot peppers and sweet peppers, and the positions of 15 agronomic traits in the chromosomes were annotated. Traits such as Flower color, Male sterile, and Style color, Anther color, and Cotyledon color formed aggregation areas, where multiple agronomic traits could be controlled. These SNP aggregation areas can be critical for developing molecular markers and enhancing the efficiency of multi-trait breeding in hot peppers and sweet peppers.

In summary, the 197 hot peppers and sweet peppers germplasm resources were divided into 9 clusters, clarifying their phylogenetic relationships and obtaining many SNP loci related to agronomic traits, providing a wealth of biological information for studying hot peppers and sweet peppers agronomic traits. In addition, the correlations between different agronomic traits and SNP aggregation areas were comprehensively explored, laying a theoretical foundation for multi-trait breeding of hot and sweet peppers. In future research, the genetic effects of SNPs will be analyzed to clarify the relationships between different traits and the pleiotropy of candidate genes further.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11454-8.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.
Supplementary Material 8.

Authors' contributions

For Data curation, YNM and HXZ; Investigation,ZZ,HXZ and ZHY; Methodology, YNM and LBY; Writing – original draft, YNM; Writing – review & editing, YNM,YQF and LBY.

Funding

This work was supported by the 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).

Data availability

All data generated or analyzed during this study are included in this published article and its Supplementary information files. The raw data are freely available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1225870.

Declarations

Ethics approval and consent to participate

All plant materials (not endangered or species) comply with local institutional guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institute of Economic Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang 050051, China. [2]Hebei Province Engineering Research Center for Vegetables, Shijiazhuang 050051, China. [3]College of Horticultural Science & Technology, Hebei Normal University of Science & Technology, Qinhuangdao 066000, China.

Received: 10 December 2024 Accepted: 5 March 2025
Published online: 17 March 2025

References

1. Silva WC, de Carvalho SIC, Duarte JB. Identification of minimum descriptors for characterization of Capsicum spp. germplasm. Hortic Bras. 2013;31:190–202.
2. Danojević D, Medić-Pap S. Different multivariate analysis for fruit traits in sweet pepper breeding. Genetika. 2018;50(1):121–9.
3. Wang N, Fang L, Xin H, et al. Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol. 2012;12:1–15.
4. Bourgeois YXC, Lhuillier E, Cézard T, et al. Mass production of SNP markers in a nonmodel passerine bird through RAD sequencing and contig mapping to the zebra finch genome. Mol Ecol Resour. 2013;13(5):899–907.
5. Nakamura A, Osonoi T, Terauchi Y. Relationship between urinary sodium excretion and pioglitazone-induced edema. J Diab Invest. 2010;1(5):208–11.
6. Huang X, Han B. Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol. 2014;65(1):531–51.
7. Sauvage C, Segura V, Bauchet G, et al. Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits. Plant Physiol. 2014;165(3):1120–32.
8. Zhao J, Artemyeva A, Del Carpio DP, et al. Design of a Brassica rapa core collection for association mapping studies. Genome. 2010;53(11):884–98.
9. Zhang P, Zhu Y, Wang L, et al. Mining candidate genes associated with powdery mildew resistance in cucumber via super-BSA by specific length amplified fragment (SLAF) sequencing. BMC Genomics. 2015;16:1–14.
10. Upadhyaya HD, Bajaj D, Narnoliya L, et al. Genome-wide scans for delineation of candidate genes regulating seed-protein content in chickpea. Front Plant Sci. 2016;7:302.
11. Sakiroglu M, Brummer EC. Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. Theor Appl Genet. 2017;130:261–8.
12. Nimmakayala P, Abburi VL, Saminathan T, et al. Genome-wide diversity and association mapping for capsaicinoids and fruit weight in Capsicum annuum L. Sci Rep. 2016;6(1):38081.
13. Nimmakayala P, Abburi VL, Saminathan T, et al. Genome-wide divergence and linkage disequilibrium analyses for Capsicum baccatum revealed by genome-anchored single nucleotide polymorphisms. Front Plant Sci. 2016;7:1646.
14. Ahn YK, Manivannan A, Karna S, et al. Whole genome resequencing of Capsicum baccatum and Capsicum annuum to discover single nucleotide polymorphism related to powdery mildew resistance. Sci Rep. 2018;8(1):5188.
15. Phan NT, Sim SC. Genomic tools and their implications for vegetable breeding. J Hortic Sci Technol. 2017;35(2):149–64.
16. Ingvarsson PK, Street NR. Association genetics of complex traits in plants. New Phytol. 2011;189(4):909–22.
17. Chai X, Xu X, Wang D, et al. Mapping and candidate gene screening of Cladosporium fulvum resistance gene Cf-12 in tomato (Solanum lycopersicum) by high-throughput sequencing. Plant Breed. 2020;139(5):977–87.
18. Tang Y, Zhang G, Jiang X, et al. Genome-wide association study of glucosinolate metabolites (mGWAS) in Brassica napus L. Plants. 2023;12(3):639.

19. Huang K, Shi Y, Pan G, et al. Genome-wide association analysis of fiber fineness and yield in ramie (Boehmeria nivea) using SLAF-seq. Euphytica. 2021;217:1–16.
20. Zhang Q, Sun T, Wang J, et al. Genome-wide association study and high-quality gene mining related to soybean protein and fat. BMC Genomics. 2023;24(1):596.
21. Gong J, Peng Y, Yu J, et al. Linkage and association analyses reveal that hub genes in energy-flow and lipid biosynthesis pathways form a cluster in upland cotton. Comput Struct Biotechnol J. 2022;20:1841–59.
22. Zhang Y, Wang L, Xin H, Li D, Ma C, Ding X, Hong W, Zhang X. 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. 2013;13:141.
23. Zhang T, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nat Biotechnol. 2015;33(5):531–7.
24. Sun X, Liu D, Zhang X, Li W, Liu H, Hong W, Jiang C, Guan N, Ma C, Zeng H, et al. SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using highthroughput sequencing. PLoS One. 2013;8:e58700.
25. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
26. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
27. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92.
28. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655–64.
29. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2002;2(4):618–20.
30. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21(2):263–5.
31. Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547–9.
32. Trees RP. The neighbor-joining method: a new method for. Mol Biol Evol. 1987;4(4):406–25.
33. Eickmeyer K, Huggins P, Pachter L, et al. On the optimality of the neighbor-joining algorithm. Algorithms Mol Biol. 2008;3:1–11.
34. Kim HJ, Han JH, Kwon JK, et al. Fine mapping of pepper trichome locus 1 controlling trichome formation in Capsicum annuum L. CM334. Theor Appl Genet. 2010;120:1099–106.
35. Lee HR, Kim KT, Kim HJ, et al. QTL analysis of fruit length using rRAMP, WRKY, and AFLP markers in chili pepper. Hortic Environ Biotechnol. 2011;52:602–13.
36. Lee J, Park SJ, Hong SC, et al. QTL mapping for capsaicin and dihydrocapsaicin content in a population of Capsicum annuum 'NB 1'× Capsicum chinense 'Bhut Jolokia.' Plant Breed. 2016;135(3):376–83.
37. Dwivedi N, Kumar R, Paliwal R, et al. QTL mapping for important horticultural traits in pepper (Capsicum annuum L.). J Plant Biochem Biotechnol. 2015;24:154–60.
38. Wu Kun WK, Liu HongYan LHY, Yang MinMin YMM, et al. High-density genetic map construction and QTLs analysis of grain yield-related traits in sesame (Sesamum indicum L.) based on RAD-Seq technology. BMC Plant Biol. 2014;14:1–14.
39. Wang J, Wang Z, Du X, et al. A high-density genetic map and QTL analysis of agronomic traits in foxtail millet [Setaria italica (L.) P. Beauv.] using RAD-seq. PloS One. 2017;12(6):e0179717.
40. Lee JH, Natarajan S, Biswas MK, et al. SNP discovery of Korean short day onion inbred lines using double digest restriction site-associated DNA sequencing. PLoS One. 2018;13(8):e0201229.
41. Hong Y, Zhang M, Zhu J, et al. Genome-wide association studies reveal novel loci for grain size in two-rowed barley (Hordeum vulgare L.). Theoret Appl Genet. 2024;137(3):58.
42. Cheng L, Dong X, Liu Q, et al. SLAF-Seq technology-based genome-wide association and population structure analyses of ancient Camellia sinensis (L.) Kuntze in Sandu County, China. Forests. 2022;13(11):1885.
43. Mandozai A, Moussa AA, Zhang Q, et al. Genome-wide association study of root and shoot related traits in spring soybean (Glycine max L.) at seedling stages using SLAF-seq. Front Plant Sci. 2021;12:568995.
44. Xie D, Dai Z, Yang Z, et al. Genome-wide association study identifying candidate genes influencing important agronomic traits of flax (Linum usitatissimum L.) using SLAF-seq. Front Plant Sci. 2018;8:2232.
45. Wu Y, Shi H, Yu H, et al. Combined GWAS and transcriptome analyses provide new insights into the response mechanisms of sunflower against drought stress. Front Plant Sci. 2022;13:847435.
46. Jiang T, Wen C, Tian W, et al. SNP molecular markers development and genetic diversity analysis of Forsythia suspensa based on SLAF-seq technology. Med Plant Res. 2022;12.
47. Lee HY, Ro NY, Patil A, et al. Uncovering candidate genes controlling major fruit-related traits in pepper via genotype-by-sequencing based QTL mapping and genome-wide association study. Front Plant Sci. 2020;11:1100.
48. Wassan GM, Khanzada H, Zhou Q, et al. Identification of genetic variation for salt tolerance in Brassica napus using genome-wide association mapping. Mol Genet Genomics. 2021;296:391–408.
49. Ro N, Oh H, Ko HC, Yi J, Na YW, Haile M. Exploring genomic regions associated with fruit traits in pepper: insights from multiple GWAS models. Int J Mol Sci. 2024;25(21):11836.
50. Wu L, Wang P, Hao Y, Cheng Q, Hua Q, Liu J, Li T, Ai Y, Yang W, Sun L, Shen H. Genome-wide correlation of 36 agronomic traits in the 287 pepper accessions obtained from the SLAF-seq-based GWAS. Int J Mol Sci. 2019;20(22):5675.
51. Wang H, Qin F. Genome-wide association study reveals natural variations contributing to drought resistance in crops. Front Plant Sci. 2017;8:1110.
52. Ogura T, Busch W. From phenotypes to causal sequen ces: using genome wide association studies to dissect the sequence basis for variation of plant development. Curr Opin Plant Biol. 2015;23:98–108.

Publisher's Note