Journal of Integrative Agriculture 2026, 25(3): 893–902

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Identification of a pleiotropic QTL and development of KASP markers for 100-pod weight, 100-seed weight, and shelling percentage in peanut

Xiukun Li[1*], Jing Hao[1*], Hongtao Deng[1*], Shunli Cui[1], Li Li[2], Mingyu Hou[1], Yingru Liu[1], Lifeng Liu[1#]

[1] *State Key Laboratory of North China Crop Improvement and Regulation/Key Laboratory for Crop Germplasm Resources of Hebei/North China Key Laboratory for Crop Germplasm Resources, Ministry of Education/College of Agronomy, Hebei Agricultural University, Baoding 071001, China*
[2] *School of Landscape and Ecological Engineering, Hebei University of Engineering, Handan 056038, China*

Highlights

- A total of 46 QTLs and 115 SNPs associated with 100-pod weight (HPW), 100-seed weight (HSW), and shelling percentage (SP) were identified through combined linkage and association analyses.
- The locus *qSP7*, which influences SP, was consistently detected across two populations.
- Three KASP markers tightly linked to major QTL regions were developed and can be utilized for molecular marker-assisted breeding of peanut yield.

Abstract

High yield remains the primary objective of peanut breeding. Key yield components, 100-pod weight (HPW), 100-seed weight (HSW), and shelling percentage (SP), are critical determinants of overall productivity. This study aimed to construct a high-density linkage map using resequencing data from recombinant inbred lines (RILs) derived from a cross between 'Silihong' (*Arachis hypogaea* var. *fastigiate*) and 'Jinonghei 3' (*A. hypogaea* var. *hypogaea*). The resulting map comprised 4,499 bins distributed across 20 chromosomes, spanning a total length of 1,712.32 cM with an average inter-marker distance of 0.38 cM. A total of 46 quantitative trait loci (QTLs) were identified across three environments. Major QTLs, including *qHPW5.2*, *qHPW18.1*, *qSP7.1*, *qSP8.1*, *qSP8.2*, *qSP18.1*, and *qSP18.2*, explained phenotypic variation (PVE) of 12.04, 11.41, 16.53, 24.17, 10.49, 10.82, and 29.89%, respectively. Fourteen QTLs detected across multiple environments were considered stable. Notably, one QTL region (*qHPW7*, *qHSW7.1*, and *qSP7*) was associated with all three traits, accounting for PVE values of 8.91, 9.04, and 16.53% for HPW, HSW, and SP, respectively. To validate the accuracy of QTL mapping, a genome-wide association study (GWAS) was conducted using the US mini-core collection. Across two environments, 115 single-nucleotide polymorphisms (SNPs) were significantly associated with HPW, HSW, and SP in the association panel. Six SNPs were linked to two traits, explaining an average phenotypic variation of 13.84%. Integration of both mapping populations revealed that AX-176802178, detected on chromosome 7 in the association panel and associated with SP, was located within the confidence interval of QTL *qSP7* defined by the recombined inbred lines (RIL) population. Furthermore, three KASP markers were developed and validated in peanut landraces and cultivated varieties. These findings provide valuable insights into the genetic architecture underlying HPW, HSW, and SP, and offer useful molecular tools for marker-assisted selection in peanut breeding programs.

Keywords: peanut, GWAS, KASP, QTL mapping, shelling percentage

1. Introduction

Cultivated peanut (*Arachis hypogaea* L.), an allotetraploid species (AABB, 2n=4x=40), is valued for its high-quality edible oil, protein content, and diverse bioactive compounds. It serves as a vital global source of vegetable oil and a major cash crop (Bertioli *et al.* 2016; Zou *et al.* 2016; Zhuang *et al.* 2019). Originating in the Chaco region between southern Bolivia and northwestern Argentina, it was domesticated in

Received 11 March 2024; Received in revised form 6 May 2024; Accepted 30 May 2024; Available online 27 June 2024
[#]Correspondence Lifeng Liu, E-mail: liulifeng@hebau.edu.cn
[*] These authors contributed equally to this study.

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2024.06.013

Central South America (Kunta *et al.* 2022) and is now grown in over 100 countries, providing essential nutrition (Ravi *et al.* 2011; Pawan *et al.* 2016). In 2019, the reference genome for cultivated peanut was released, with a size of approximately 2.7 Gb, corresponding to the combined genomes of the wild diploid progenitors *A. duranensis* (AA, 2n=2x=20) and *A. ipaensis* (BB, 2n=2x=20) (Robledo *et al.* 2009; Bertioli *et al.* 2016, 2019; Dash *et al.* 2016; Zhuang *et al.* 2019). This milestone advanced the understanding of peanut genetics and enabled accelerated research into the molecular basis of agronomic traits.

The genetic dissection of complex quantitative traits primarily relies on identifying quantitative trait loci (QTLs) or associated genomic regions using molecular markers. Biparental populations and association panels are typically employed for QTL detection *via* linkage analysis and linkage disequilibrium analysis, respectively (Lu *et al.* 2010). Over 200 QTLs have been reported in cultivated peanuts (*A. hypogaea* L.) (https://peanutbase.org/search/qtl). However, only a limited number of causal genes have been isolated using map-based cloning or genome-wide association study (GWAS) approaches. Integrating multiple methods can improve genotypic resolution and strengthen validation, thereby reducing false positives and accelerating the identification of candidate genes or markers. For instance, one recombinant inbred line (RIL) population and an association panel were jointly used to identify QTLs for eight ear and grain morphological traits (Zhang *et al.* 2017). The US mini-core collection and RILs were combined to uncover loci associated with growth habit-related traits in cultivated peanuts (*Arachis hypogaea* L.) (Li L *et al.* 2022). Two nested association mapping (NAM) populations were utilized to detect loci influencing 100-pod weight (HPW) and 100-seed weight (HSW) (Gangurde *et al.* 2020). Nevertheless, low-density genetic maps derived from biparental populations often lack sufficient resolution to detect QTLs with small phenotypic effects (Holland 2007). Recent advances in next-generation sequencing technologies offer substantial benefits for high-throughput genotyping and the construction of high-density genetic maps (Chen *et al.* 2014). Whole-genome resequencing, a powerful tool for genotyping and constructing ultra-high-density maps, provides an effective strategy for dissecting complex traits in large and polyploid genomes (Huang *et al.* 2009; Meuwissen and Goddard 2010).

Yield is among the most complex agronomic traits in crops, and enhancing peanut yield remains a primary breeding objective. Yield is directly influenced by component traits such as HPW, HSW, number of pods per plot, and shelling percentage (SP) (Lucia *et al.* 2005; Luo *et al.* 2019). Elucidating the genetic basis of HPW, HSW, and SP has become a key focus in recent studies. For example, a RIL population derived from the cross between Tifrunner and NC3033 was used to identify QTLs for seed weight, pod weight, and SP, resulting in six, seven, and eight QTLs for these traits, respectively (Chavarro *et al.* 2019). Another study detected four additive QTLs for HPW using a mapping population (Xuhua 13×Zhonghua 6) comprising 187 RILs (Luo *et al.* 2018). Additionally, four QTLs for HSW were identified in an RIL population from the cross between the cultivar 'Huayu 36' and the germplasm line '6-13' (Zhang *et al.* 2019). Furthermore, two major QTLs and nine candidate genes for SP were discovered in the RIL population (Yuanza 9102×Xuzhou 68-4) (Luo *et al.* 2019). Despite these findings, the underlying genetic architecture and molecular mechanisms governing HPW, HSW, and SP remain largely unknown.

In our previous study, several QTLs for HPW and HSW were identified using a genetic linkage map constructed with SSR, AhTE, SRAP, and TRAP markers (Meng *et al.* 2021). However, the physical locations of these QTLs were not precisely determined, limiting fine mapping efforts. Therefore, this study applied whole-genome resequencing to construct an ultra-high-density bin map using the same RIL population derived from the cross between 'Jinonghei 3' and 'Silihong'. Subsequently, integrated linkage analysis and GWAS were performed to identify QTLs associated with SP, HPW, and HSW. These results provide a foundation for the genetic dissection of yield-related traits and support future molecular breeding strategies in peanut.

2. Materials and methods

2.1. Plant materials

This study utilized a RIL population and an association panel. The RIL population consisted of 248 individual lines derived from the cross between Jinonghei 3 (*A. hypogaea* var. *hypogaea*) and Silihong (*A. hypogaea* var. *fastigiate*). Silihong is a widely cultivated Chinese landrace characterized by a high SP and low HPW and HSW. In contrast, Jinonghei 3 is an improved cultivar with a purple-black seed coat, exhibiting moderate HPW and HSW and a lower SP (Wang *et al.* 2019). The association panel included 103 accessions from the US mini-core collection obtained from the USDA Peanut Germplasm Collection in Griffin, GA, USA (Chen *et al.* 2014).

2.2. Field experiment design and statistical analysis

Field experiments were conducted at the Qingyuan Experimental Station of Hebei Agricultural University, Baoding, China (38.79°N, 115.56°E). The RIL population was grown across three growing seasons: 2017 (E1), 2018 (E2), and 2020 (E3). The association panel was evaluated in 2017 (E1) and 2018 (E2). In each season, all RIL lines, their parental genotypes, and the US mini-core accessions were planted using a randomized complete block design with two replications, with 10 plants per line per replicate. Within-row spacing was 10 cm, and between-row spacing was 0.35 m. A single seed was sown per hole, and border rows were established on all sides of experimental plots to minimize edge effects. Field management practices followed local agricultural guidelines.

At maturity, eight plants from the central portion of each plot were individually harvested for trait assessment. Pods were air-dried and stored at room temperature for more than 1-mon post-harvest. HPW, HSW, and SP were measured as previously described (Jiang 2006). Briefly, 100 representative, mature, and undamaged pods were randomly selected from the eight harvested plants and weighed to

calculate HPW. Similarly, 100 mature, fully filled seeds were randomly sampled from the same plants and weighed to determine HSW. SP was computed as the percentage of HSW relative to HPW following manual pod shelling. These measurements for HPW, HSW, and SP were collected using consistent protocols for both populations, and each sample was analyzed in triplicate. Data were processed and analyzed using Microsoft Excel.

2.3. DNA extraction and resequencing

RILs selected for resequencing were planted in June 2021 at the Qingyuan Experimental Station. Two young, healthy leaves from a single plant of each RIL were collected and immediately frozen at –80°C. Genomic DNA was extracted from each RIL using the DNeasy Plant Mini Kit (Qiagen, Germany). DNA integrity and purity were assessed using 1% agarose gel electrophoresis and a NanoPhotometer spectrophotometer (IMPLEN, CA, USA). DNA concentration was quantified using the Qubit DNA Assay Kit on a Qubit 2.0 Fluorometer (Life Technologies, CA, USA). A total of 250 samples, including both parental lines and RILs, were processed. Approximately 1.5 μg of genomic DNA per sample was used as input for library preparation. Sequencing libraries were constructed using the TruSeq Nano DNA HT Sample Preparation Kit (Illumina, USA) according to the manufacturer's instructions, with unique index codes assigned to each sample. DNA was fragmented by sonication to an average size of 350 bp. Fragmented DNA underwent end repair, A-tailing, and ligation with full-length adapters compatible with Illumina sequencing, followed by PCR amplification. The resulting PCR products were purified using the AMPure XP System (Beckman Coulter, USA). Library fragment size distribution was evaluated using an Agilent 2100 Bioanalyzer (Agilent, USA), and library quantification was performed *via* real-time PCR. Paired-end sequencing was carried out on an Illumina HiSeq2500 platform (Illumina, USA) at the Novogene Bioinformatics Institute, Tianjin, China.

Raw reads were demultiplexed based on barcode sequences. Raw sequencing data underwent quality control (QC) using custom C scripts under the following criteria: (1) reads containing ≥10% unidentified nucleotides (N) were removed; (2) reads with >50% of bases having Phred quality scores<5 were excluded; (3) reads with >10 nt aligned to adapter sequences (allowing≤10% mismatches) were discarded. Clean reads were aligned to the reference genome using the Burrows–Wheeler Aligner (Li and Durbin 2009) (settings: mem -t 4 -k 32 -M -R). Resulting alignment files were converted into BAM format using SAMtools software (Li *et al.* 2009) (settings: -bS -t). Variant calling across all samples was performed using SAMtools (Kai *et al.* 2010). Single-nucleotide polymorphisms (SNPs) were filtered using a Perl script. Functional annotation of SNPs and InDels was conducted using ANNOVAR (Kai *et al.* 2010) based on GFF3 annotation files for the *Ah*-Tifrunner reference genome (version 2 (gnm2): *Arachis hypogaea* cv. *Tifrunner*: assembly, annotation (gnm2.ann1)) available at https://peanutbase.org/peanut_genome.

2.4. Genotyping and construction of bin map

A bin map was generated and visualized using R software. SNP filtering was applied as follows: (1) SNPs with a maximum heterozygosity rate≥5.0% were excluded; (2) SNPs with a missing rate≥10.0% were removed; (3) SNPs with a minor allele frequency (MAF)≤0.1 were discarded. Genotypes were then converted into numerical values, and unlinked SNPs were eliminated using a sliding window approach with a window size of 10. The bin map was constructed using the similar rate method, in which SNPs with a similarity rate>0.9 were merged into a single bin. Missing genotype data within bins were imputed using a missing data threshold of 0.2 and a homozygous genotype threshold of 0.9.

2.5. QTL analysis

QTL analysis was conducted using QTL IciMapping version 4.0, employing the mixed model-based composite interval mapping method. A permutation test with 1,000 iterations was performed to establish significance thresholds. The logarithm of odds (LOD) score threshold was set at 2.5, corresponding to a 5% genome-wide significance level. A 95% confidence level was applied. QTLs were designated with an initial lowercase "q" followed by the abbreviated trait name (HPW, HSW, or SP) and the chromosome number. When multiple QTLs for the same trait were detected on a single chromosome, an Arabic numeral was appended to distinguish them. For example, *qSP8.3* denotes the third QTL for SP on chromosome 8.

2.6. Genome-wide association study

The association panel was genotyped using the Affymetrix Axiom® *Arachis* 2.0 SNP array. A total of 12,342 high-quality SNPs were retained for GWAS after applying the following filters: MAF>0.05, missing data rate<20%, and heterozygosity rate<5.0%.

GWAS was conducted using TASSEL 5.0 software under a mixed linear model (MLM), incorporating population structure parameters (Q matrix) previously estimated using STRUCTURE software. A less stringent significance threshold of $-\log_{10}(P)>2.5$ was adopted to detect significant marker-trait associations (Li *et al.* 2022).

2.7. KASP marker development and validation

KASP markers were developed based on three SNP pairs flanking four identified QTLs (Appendix A). Primers were designed by extracting 150 bp of sequence upstream and downstream of each candidate SNP from the resequencing data. KASP assays were performed in 10 μL reaction volumes containing 2.0 μL of genomic DNA (25–50 ng), 5 μL of 2× AQP master mix, 0.15 μmol L^{-1} of each forward primer, 0.4 μmol L^{-1} of the common reverse primer, and nuclease-free water (JasonGen Biotech Co., Ltd., Beijing, China). PCR amplification was carried out on an ABI QuantStudio 6 System (ABI Life Technologies, CA, USA) using the following cycling conditions: 95°C for 10 min; 10 cycles of 95°C for 20 s and 61°C (decreasing by 0.6°C per cycle) for 40 s; followed by 31 cycles of 95°C for 20 s and 55°C for 40 s. Fluorescence signals were detected during amplification, and

genotype calling was performed manually using the built-in ABI QuantStudio 6 software. Subsequently, the KASP markers were validated in a diverse peanut panel comprising 81 Chinese peanut varieties and landraces. The HPW, HSW, and SP traits of this panel were evaluated in two independent trials conducted in 2021 and 2022 in Baoding, Hebei Province, China, with two replications per trial.

3. Results

3.1. Phenotype analysis

A set of RILs and an association panel were phenotyped for HPW, HSW, and SP across three and two environments, respectively. The average HPW, HSW, and SP of the RIL population were 141.16 g, 49.94 g, and 68.39%, respectively. Across environments, Silihong exhibited mean HPW and HSW values of 150.33 and 46.23 g, which were higher and lower than those of Jinonghei 3, averaging 131.26 and 71.56 g, respectively. However, Silihong displayed a higher SP (67.17%) compared to Jinonghei 3 (62.86%) (Fig. 1). The association panel had average HPW, HSW, and SP values of 103.05 g, 40.09 g, and 40.09%, respectively (Fig. 2-A), which were lower than those observed in the RILs. Nevertheless, the phenotypic variation for all three traits was greater in the association panel. These results exhibited continuous distributions with transgressive segregation, indicating suitability for subsequent statistical and QTL analyses.

3.2. Whole-genome resequencing and the construction of a highly dense bin map

We performed whole-genome resequencing of a segregating population comprising 248 RILs and the parental lines Jinonghei 3 and Silihong using the Illumina HiSeq2500 platform. A total of 3,187.77 GB of clean sequencing data were generated, achieving approximately 5-fold coverage per RIL and 30-fold coverage for each parental line. Sequence reads were aligned to the reference genome Ah-Tifrunner-Reference-Version-2 available at https://peanutbase.org/peanut_genome using SOAP2 (Bertioli *et al.* 2019; Li *et al.* 2009b). Using SOAPsnp (Li *et al.* 2009a), we identified 3,362,335 SNPs that were homozygous and polymorphic between the two parental lines. From these, 360,171 high-quality polymorphic SNP markers were selected to construct a recombinant bin map. For genetic mapping, the recombination map was simplified into a skeleton bin map, and all chromosomes of the 248 RILs were aligned accordingly (Appendix B). The resulting genetic map consisted of 20 linkage groups corresponding to the 20 peanut chromosomes. A total of 4,499 bins were used to construct the final map, spanning 1,712.32 cM, which was generated using R software. On average, each chromosome contained 225 bins, with a mean genetic distance of 0.38 cM between adjacent bins (Table 1).

3.3. QTL mapping for HPW, HSW, and SP in RILs

A total of 46 QTLs associated with the three traits were identified (Fig. 3) across three environments. The physical intervals spanned 0.44 to 1.87 Mb, averaging 0.86 Mb. Among these, 12 QTLs were linked to HPW and mapped to chromosomes 5, 7, 8, 10, and 18, with PVE ranging from 3.68 to 18.41%. Another 12 QTLs were associated with HSW and, distributed across chromosomes 1, 5, 6, 7, 9, 13, and 18, explaining an average of 6.18% of phenotypic variance. Additionally, 24 QTLs were detected for SP on chromosomes 1, 6, 7, 8, 13, 15, and 18, with a mean PVE of 11.58%.

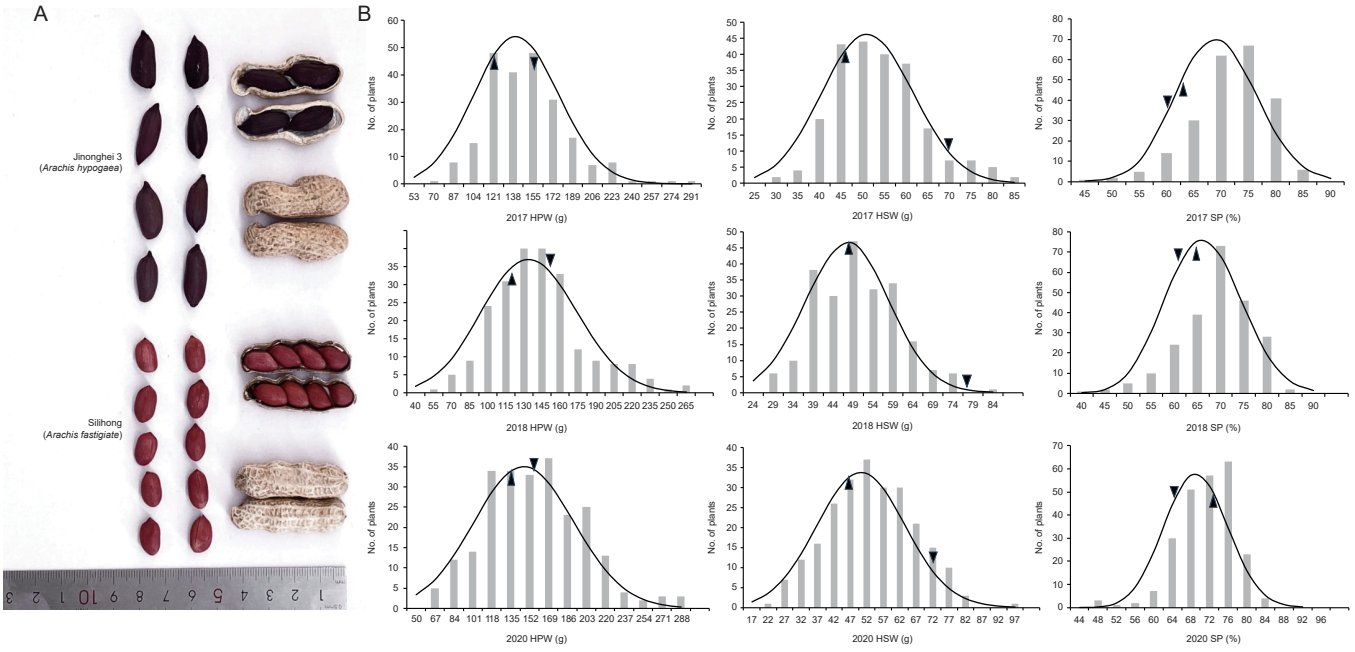


Fig. 1 Phenotypic analysis of recombinant inbred lines (RILs) and their parental lines across three different environments. A, phenotypes of mature dry pods from the parental cultivars 'Silihong' and 'Jinonghei 3'. B, variation in 100-pod weight (HPW), 100-seed weight (HSW), and shelling percentage (SP) within the RIL population across all three environments.

Table 1 Description of the peanut genetic linkage map of recombinant inbred lines (RILs)

Chr.	Bin number	Genetic distance (cM)	Average distance of adjacent markers (cM)
1	190	87.42	0.46
2	155	56.06	0.36
3	190	110.51	0.58
4	154	73.73	0.48
5	191	77.48	0.41
6	188	97.80	0.52
7	133	87.70	0.66
8	76	101.09	1.33
9	209	109.10	0.52
10	235	99.50	0.42
11	296	79.73	0.27
12	211	52.12	0.25
13	265	125.20	0.47
14	215	43.90	0.20
15	333	77.99	0.23
16	290	97.38	0.34
17	278	81.14	0.29
18	299	86.01	0.29
19	292	83.23	0.29
20	299	85.23	0.29
Total	4,499	1,712.32	0.38

Notably, *qHPW5.2*, *qHSW7.1*, and *qSP18.2* exhibited the highest PVEs for HPW, HSW, and SP, respectively.

Fourteen QTLs were stable across multiple environments (Appendix C). Of these, three were associated with HPW, three with HSW, and eight with SP. The locus *qSP18.2* had the largest effect among all stable QTLs, accounting for 29.89% of phenotypic variation. Four QTLs influenced two or more traits: one (*qHPW7*, *qHSW7.1*, and *qSP7*) was associated with all three traits, explaining 8.91, 9.04, and 16.53% of the variation in HPW, HSW, and SP, respectively. Three additional QTLs were associated with both HPW and HSW, while one co-localized for HPW and SP (Appendix C). These results suggest shared genetic control underlying the development of these traits.

3.4. GWAS for HPW, HSW, and SP

A total of 115 single nucleotide polymorphisms (SNPs) associated with the three traits were detected across environments (Fig. 2). These SNPs were located on 16 chromosomes and explained between 10.72 and 25.90% of phenotypic variance. Nine SNPs were associated with HPW, distributed on chromosomes 6, 7, 16, 18, and 19. An additional nine SNPs were linked to HSW, localized to chromosomes 6, 16, 19, and 20. Notably, 97 SNPs were associated with SP, including 39 on chromosome 5.

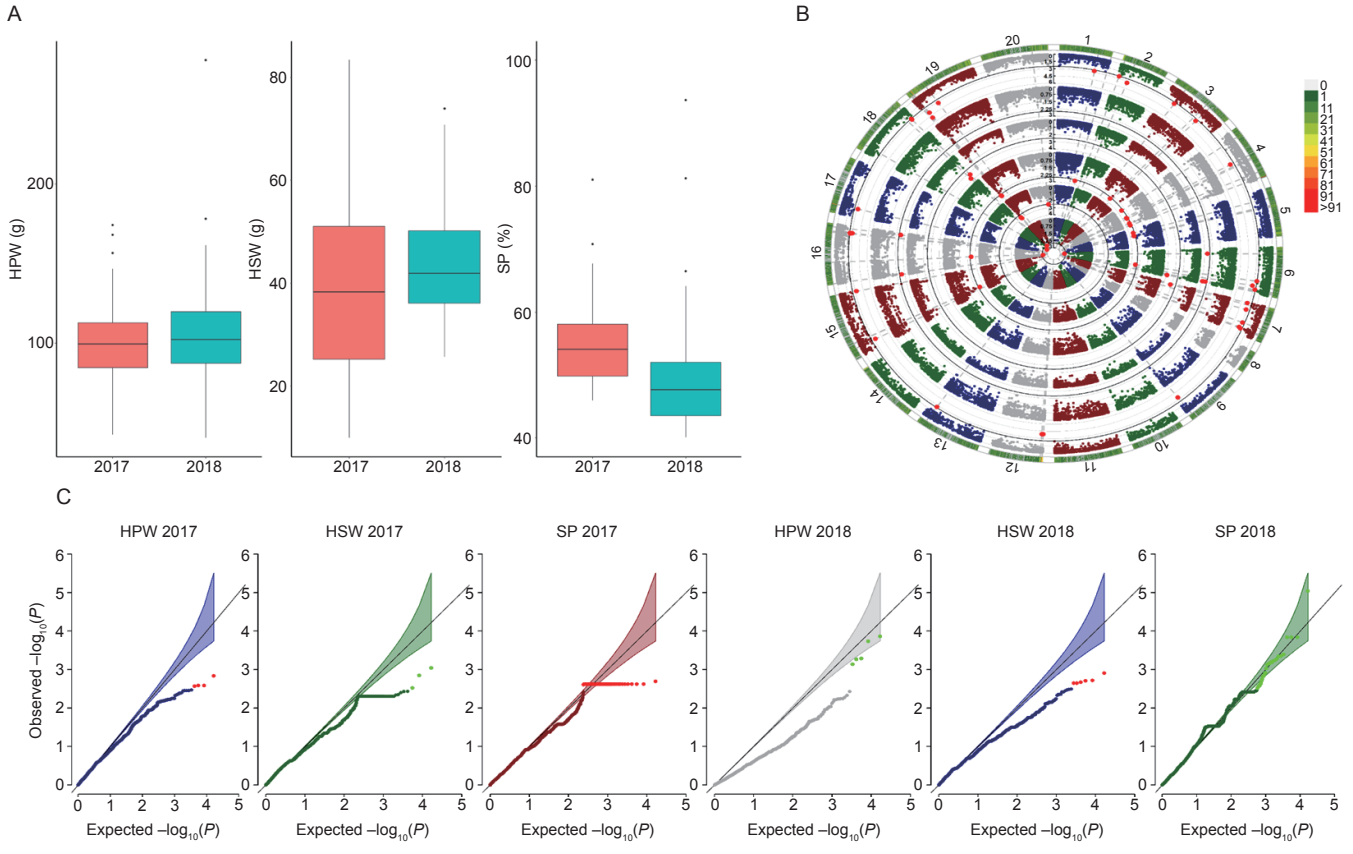


Fig. 2 Genome-wide association scans for single-nucleotide polymorphisms (SNPs) associated with 100-pod weight (HPW), 100-seed weight (HSW), and shelling percentage (SP). A, box plots show phenotypic variation for SP, HPW, and HSW in the US mini-core collection. B, concentric circles from innermost to outermost represent Manhattan plot results for HPW in 2017, HSW in 2017, SP in 2017, HPW in 2018, HSW in 2018, and SP in 2018. The horizontal solid line indicates the genome-wide significance threshold $-\log_{10}(P)=2.5$. C, Q-Q plots for the genome-wide association study (GWAS) of HPW, HSW, and SP.

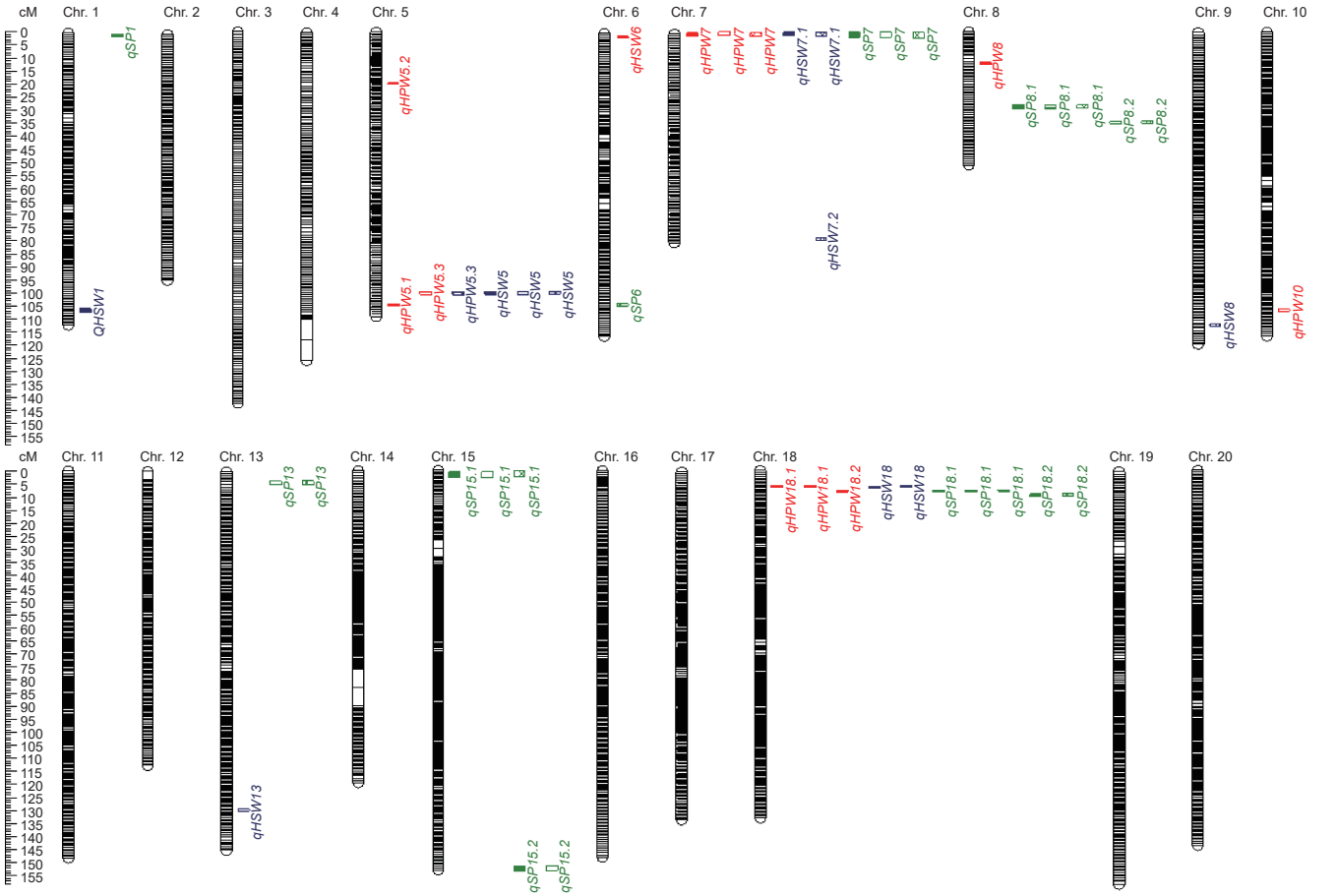


Fig. 3 Quantitative trait locus (QTL) analysis for 100-pod weight (HPW), 100-seed weight (HSW), and shelling percentage (SP) under three environments. Colored boxes represent QTLs detected for three traits under three environments.

Among the 115 SNPs, six showed pleiotropic effects on two traits and collectively accounted for an average of 13.84% of phenotypic variation across all three traits. Specifically, AX-176799328 (Chr. 6) and AX-147251673 (Chr. 16) were associated with both HPW and HSW. Furthermore, AX-147259251 and AX-147259252 (both on Chr. 18) were linked to HPW and SP, whereas AX-176821655 and AX-177640290 (Chr. 19) were associated with HSW and SP (Table 2).

Table 2 Pleiotropic QTL or single-nucleotide polymorphisms (SNPs) for 100-seed weight (HSW), 100-pod weight (HPW), and shelling percentage (SP) across multiple environments

Chr.	Flanking marker	Position (bp)	Trait	PVE (%)[1)]
5	868–869	106,334,456–106,992,335	HSW, HPW	7.07–8.15
7	1,069–1,070	562,585–1,415,079	HSW, HPW, SP	8.91–16.53
18	3,618–3,619	6,962,901–7,528,117	HSW, HPW	5.08–11.41
18	3,621–3,622	8,587,895–9,068,728	HSW, SP	4.57–10.82
6	AX-176799328	14,881,762	HSW, HPW	13.00–14.00
16	AX-147251673	1,987,025	HSW, HPW	13.00–14.00
18	AX-147259251	123,346,474	HPW, SP	16.00–18.00
18	AX-147259252	123,349,333	HPW, SP	18.00
19	AX-176821655	5,372,024	HSW, SP	14.00–15.00
19	AX-177640290	6,214,436	HSW, SP	14.00–17.00

[1)] PVE, phentypic variance explained.

3.5. Co-localization of QTLs for HPW, HSW, and SP through joint linkage mapping and GWAS

Four SNPs identified by GWAS co-localized with QTLs detected in the biparental population (Table 3). Specifically, AX-176802178, associated with SP on chromosome 7 in the association panel, resided within the confidence interval of *qSP7*, a QTL identified in the RIL population for the same trait. Similarly, two linked SNPs, AX-147224249 and AX-176805419 on chromosome 6, were associated with HPW in the GWAS and also mapped within the confidence interval of *qHSW6* defined by the RILs. Furthermore, AX-147234281, associated with SP on chromosome 9 in the association panel, fell within the confidence interval of *qHSW9*. These convergent findings across independent mapping approaches indicate that these genomic regions may harbor pleiotropic loci or regulatory hotspots influencing multiple seed-related traits.

3.6. QTL validation using KASP markers

KASP markers were developed to validate QTLs with high PVE, specifically *qHPW5.2*, *qSP18.2*, and *qSP7*, which was associated with all three traits, HSW, HPW, and SP. The

Table 3 Single-nucleotide polymorphisms (SNP) identified by genome-wide association study (GWAS) overlapping with quantitative trait loci (QTLs) mapped in biparental population

Trait[1)]	Marker	Chr.	Position	QTL	Chr.	Flanking marker	Position (bp)
HPW	AX-147224249	6	2,373,392	*qHSW6*	6	883–884	1,879,052–2,552,324
HPW	AX-176805419	6	2,456,328	*qHSW6*	6	883–884	1,879,052–2,552,324
SP	AX-176802178	7	2,403,855	*qSP7*	7	1,069–1,072	562,585–2,434,572
SP	AX-147234281	9	113,207,108	*qHSW9*	9	1,476–1,477	112,825,494–113,386,917

[1)] HPW, 100-pod weight; SP, shelling percentage.

markers exhibited polymorphism among both cultivated varieties and landraces. Notably, KASP-SP7 was significantly associated with HSW, HPW, and SP. The C and T alleles were each present at approximately 50.0% frequency across all germplasm. Materials carrying the T allele exhibited higher HSW and HPW but lower SP (Fig. 4), consistent with the additive effect of *qSP7*. For KASP-SP18, the T and C allele frequencies were approximately 83.1 and 16.9%, respectively. Genotypes with the T allele showed increased SP (Fig. 5-A). Similarly, the A and G alleles of KASP-HPW5 occurred at frequencies of about 48.2 and 51.8%, respectively, with the A allele associated with higher HPW (Fig. 5-B). These findings demonstrate that the KASP markers developed within these QTL regions can effectively screen for HSW, HPW, and SP in peanut breeding programs.

4. Discussion

4.1. Novel QTLs detected for HPW, HSW, and SP

Improving yield-related traits is essential for enhancing the yield potential of new peanut varieties. Therefore, genetic studies of traits closely associated with yield, such as seed number per pod, are both theoretically significant and practically valuable for selecting and breeding high-yielding, high-quality peanut cultivars. In this study, an RIL population and an association panel were used to dissect the genetics of three key traits. The average phenotypic performance was higher in the RIL population than in the association panel (Figs. 1 and 3-A). Both parental lines exhibited high productivity and are widely cultivated in China. In contrast, the association panel was assembled from global germplasm and displays broad phenotypic diversity. These characteristics make both populations well-suited for genetic analyses of HPW, HSW, and SP.

Among the factors influencing peanut yield, HPW, HSW, and SP have been consistently shown to correlate significantly with overall yield (Huang *et al*. 2015; Pawan *et al*. 2016; Zou *et al*. 2016; Carter *et al*. 2017; Luo *et al*. 2017; Chavarro *et al*. 2019), establishing them as key contributors to yield improvement. However, the genetic mechanisms

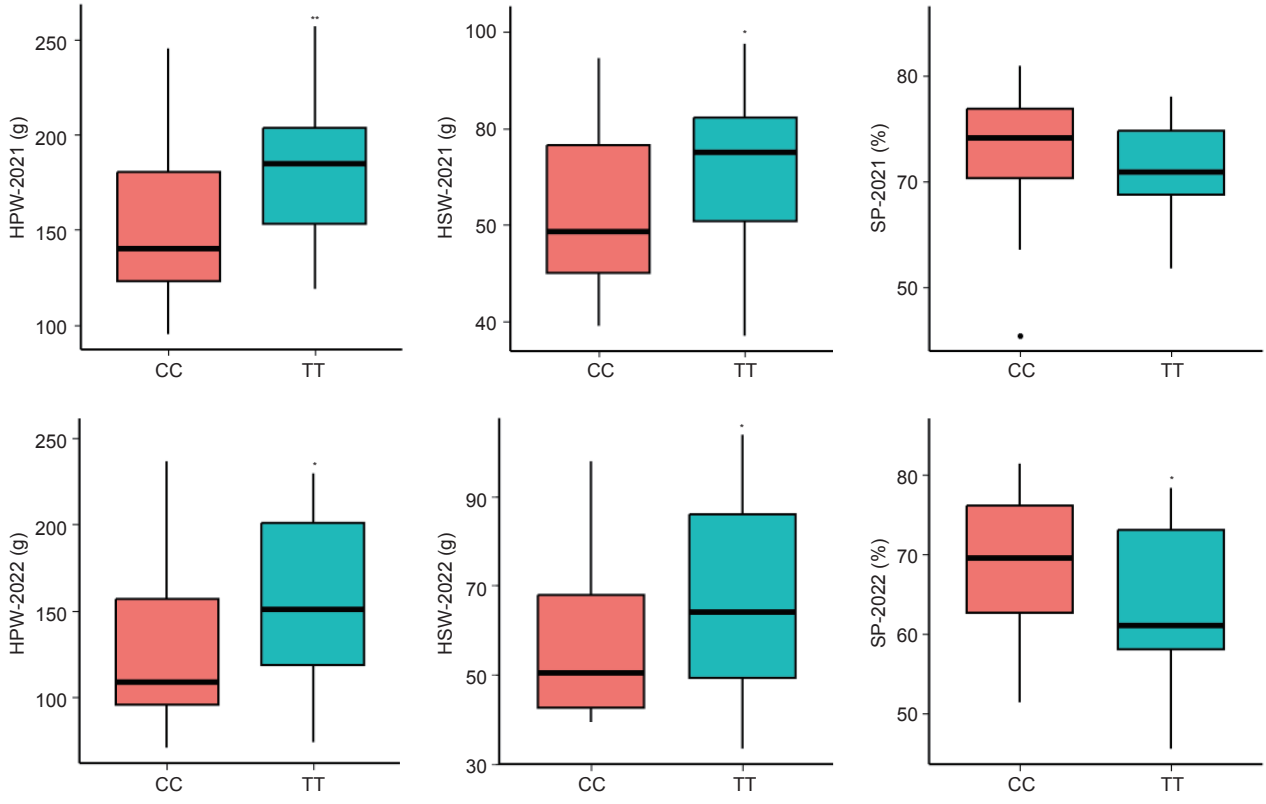


Fig. 4 Effect of KASP-SP7 on shelling percentage (SP), 100-pod weight (HPW), and 100-seed weight (HSW) in the peanut panel. * and ** indicate statistically significant differences at $P<0.05$ and $P<0.01$ (Student's *t*-test), respectively.

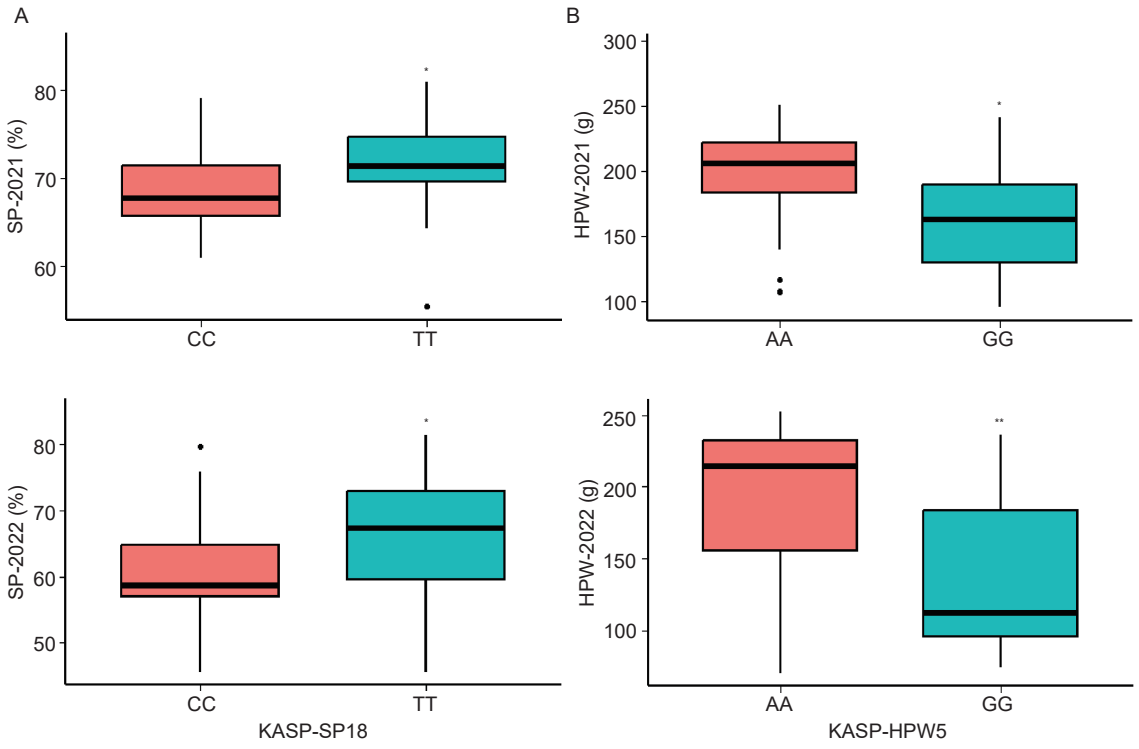


Fig. 5 Validation of KASP-SP18 and KASP-HPW5 in the peanut panel. A, effect of KASP-SP18 on shelling percentage (SP). B, effect of KASP-HPW5 on 100-pod weight (HPW). * and ** indicate significant differences at $P<0.05$ and $P<0.01$ (Student's *t*-test), respectively.

underlying the QTLs or genes controlling these traits remain poorly understood. In this study, 46 QTLs associated with HPW, HSW, and SP were identified. Some of these QTLs align with previously reported loci, while others represent novel discoveries. For example, *qHPW5.3* colocalized with HPW-related QTLs identified by Chu *et al*. (2020), and *qHPW7* was also reported by Chavarro *et al*. (2020). Similarly, for HSW, six out of 12 QTLs had been previously described (Zhang *et al*. 2019; Chavarro *et al*. 2020; Chu *et al*. 2020; Wang *et al*. 2022). Specifically, *qHSW5* colocalized with HSW-related QTLs reported by Wang *et al*. (2022) and Chu *et al*. (2020); *qHSW7.1* overlapped with QTLs identified by Chavarro *et al*. (2020), Chu *et al*. (2020), and Zhang *et al*. (2019); and *qHSW9* coincided with an HSW-related QTL reported by Chu *et al*. (2020). Regarding SP, all identified QTLs were novel except *qSP7*, *qSP8.1*, and *qSP8.2* (Li *et al*. 2022), including major-effect QTLs such as *qHPW5.2* and *qHPW18.1* for HPW, and *qSP8.1*, *qSP8.2*, *qSP18.1*, and *qSP18.2* for SP. Among the 115 significantly associated SNPs, 97 were linked to SP. Notably, three SNPs, AX-176802178 on Chr. 7, AX-147234281 on Chr. 9, and AX-176798196 on Chr. 15, were located within previously reported SP-related QTL regions (Li *et al*. 2022). Additionally, 38 SNPs clustered within a ~3-Mb region on Chr. 5, representing a newly identified locus. Previous studies identified and validated an HSW-related QTL on Chr. 16 using KASP markers across diverse peanut varieties (Wang *et al*. 2022). In this study, an HSW-associated SNP, *AX-147251673*, was identified in this same region and explained 12.93% of the phenotypic variation.

4.2. KASP markers for marker-assisted breeding of peanut HPW, HSW, and SP

Four QTLs and six SNPs showed consistent linkage or pleiotropy with HPW, HSW, and SP across multiple environments (Table 2). The QTL for HSW frequently co-occurred with that for HPW, and one QTL on Chr. 7 was associated with all three traits. Within the association panel, two SNPs (AX-147259251 and AX-147259252) were associated with both HPW and SP, while two others (AX-176821655 and AX-177640290) were associated with HSW and SP. These results highlight the strong genetic correlation among these traits and suggest pleiotropic effects, underscoring the importance of studying them collectively. Based on flanking markers, three KASP markers were developed for *qHPW7/qHSW7.1/qSP7*, *qSP18.2*, and *qHPW5.2*, respectively. Student *t*-tests in the peanut panel revealed strong associations: KASP-SP7 with HPW, HSW, and SP (Appendix D); KASP-SP18 with SP; and KASP-HPW5 with HPW, in both landraces and modern varieties (Appendix E). Furthermore, *t*-tests showed no significant difference in favorable allele frequencies between landraces and varieties for these three markers, suggesting that these alleles have not undergone artificial selection. Thus, these KASP markers offer substantial potential for marker-assisted breeding in peanut improvement programs.

4.3. Candidate gene prediction in the major QTL regions

To date, few genes associated with HPW, HSW, or SP have been cloned in peanut (Luo *et al*. 2019), primarily due to the polygenic and complex nature of these traits. In this study,

several major QTLs were identified for HPW, HSW, and SP. However, the physical intervals of these QTLs remain relatively large, necessitating fine mapping to pinpoint causal genes. Gene Ontology (GO) analysis was conducted on the genes located within the major QTL regions, *qHPW5.2*, *qHPW18.1*, *qHSW7.1*, *qSP7*, *qSP8.1*, *qSP8.2*, *qSP18.1*, and *qSP18.2*, to elucidate the genetic architecture underlying these traits. A total of 454 genes were annotated within these regions and enriched in 12 GO terms, comprising three biological process terms and nine molecular function terms (Appendix F). The most significantly enriched biological process term was peptidyl-histidine phosphorylation (GO: 0018106), while the top molecular function term was mannan synthase activity (GO:0051753). Mannan is known to play a critical role in cereal grain development, and previous studies have demonstrated its involvement in early endosperm cell development in wheat (Verhertbruggen *et al*. 2021). These findings emphasize the complexity of seed development in peanut and provide useful insights for candidate gene identification related to yield-associated traits.

5. Conclusion

In this study, a high-density linkage map was constructed by resequencing RILs derived from a cross between 'Silihong' (*A. hypogaea* var. *fastigiate*) and 'Jinonghei 3' (*A. hypogaea* var. *hypogaea*). A total of 46 QTLs and 115 SNPs associated with HPW, HSW, and SP were identified through combined linkage and association analyses. The locus *qSP7*, which influences SP, was consistently detected across two populations. Additionally, three KASP markers tightly linked to major QTL regions were developed and validated in peanut landraces and cultivated varieties. These findings enhance understanding of the genetic architecture underlying HPW, HSW, and SP in peanut and provide valuable molecular tools for marker-assisted breeding programs.

Acknowledgments

This study was financially sponsored by the National Natural Science Foundation of China (320720977), the China Agriculture Research System (CARS-13), the Hebei Agriculture Research System, China (HBCT2024040205), the S&T Program of Hebei, China (23567601H), the Peanut Modern Seed Industry Technology Innovation Team in Hebei Province, China (21326316D-2), the State Key Laboratory of North China Crop Improvement and Regulation, China (NCCIR2020RC-2), and the Startup Fund of Hebei Agricultural University, China (YJ2020044).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

Data availability

The datasets supporting the conclusions of this study are included in this manuscript. The sequences reported in this study have been deposited in the National Center for Biotechnology Information database.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2024.06.013

References

Bertioli D J, Cannon S B, Froenicke L, Huang G, Farmer A D, Cannon E K, Liu X, Gao D, Clevenger J, Dash S, Ren L, Moretzsohn M C, Shirasawa K, Huang W, Vidigal B, Abernathy B, Chu Y, Niederhuth C E, Umale P, Araujo A C, *et al.* 2016. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics*, **48**, 438–446.

Bertioli D J, Jenkins J, Clevenger J, Dudchenko O, Gao D, Seijo G, Leal-Bertioli S C M, Ren L, Farmer A D, Pandey M K, Samoluk S S, Abernathy B, Agarwal G, Ballén-Taborda C, Cameron C, Campbell J, Chavarro C, Chitikineni A, Chu Y, Dash S, *et al*. 2019. The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nature Genetics*, **51**, 877–884.

Carter E T , Rowland D L, Tillman B L, Erickson J E, Grey T L, Gillett-Kaufman J L, Clark M W. 2017. Pod maturity in the shelling process. *Peanut Science*, **44**, 26–34.

Chavarro C, Chu Y, Holbrook C, Isleib T, Bertioli D, Hovav R, Butts C, Lamb M, Sorensen R, Jackson S A, Peggy O A. 2020. Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3* (Bethesda), **10**, 2297–2315.

Chen C Y, Barkley N A, Wang M L, Holbrook C C, Dang P M. 2014. Registration of purified accessions for the U.S. peanut mini-core germplasm collection. *Journal of Plant Registrations*, **8**, 77–85.

Chen Z L, Wang B B, Dong X M, Liu H, Ren L H. 2014. An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F_2 maize population. *BMC Genomics*, **15**, 433.

Chu Y, Chee P, Isleib T G, Holbrook C C, Ozias-Akins P. 2020. Major seed size QTL on chromosome A05 of peanut (*Arachis hypogaea*) is conserved in the US mini core germplasm collection. *Molecular Breeding*, **40**, 6.

Dash S, Cannon E K S, Kalberer S R, Farmer A D, Cannon S B. 2016. *Peanutbase and Other Bioinformatic Resources for Peanut*. AOCS Press, Urbana, USA. pp. 241–252.

Gangurde S S, Wang H, Yaduru S, Pandey M K, Fountain J C, Chu Y, Isleib T, Holbrook C C, Xavier A, Culbreath A K, Ozias-Akins P, Varshney R K, Guo B. 2020. Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnology Journal*, **18**, 1457–1471.

Holland J B. 2007. Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, **10**, 156–161.

Huang L, He H, Chen W, Ren X, Chen Y, Zhou X, Xia Y, Wang X, Jiang X, Liao B. 2015. Quantitative trait locus analysis of agronomic and quality-related traits in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, **128**, 1103–1115.

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T. 2009. High-throughput genotyping by whole-genome resequencing. *Genome Research*, **19**, 1068–1076.

Jiang H. 2006. Descriptors and data standard for peanut (*Arachis* spp.). China Agriculture Press, Beijing. 11–36. (in Chinese)

Kai W, Li M, Hakon H. 2010. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, **16**, e164.

Kunta S, Chu Y, Levy Y, Harel A, Hovav R. 2022. Identification of a major locus for flowering pattern sheds light on plant architecture diversification in cultivated peanut. *Theoretical and Applied Genetics*, **135**, 1767–1777.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. Genome project data processing S: The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Li L, Cui S, Dang P, Yang X, Wei X, Chen K, Liu L, Chen C Y. 2022.

GWAS and bulked segregant analysis reveal the Loci controlling growth habit-related traits in cultivated peanut (*Arachis hypogaea* L.). *BMC Genomics*, **23**, 1–13.

Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K, Wang J. 2009a. SNP detection for massively parallel whole-genome resequencing. *Genome Research*, **19**, 1124–1132.

Li R, Yu C, Lam T W, Yiu S M, Kristiansen K, Wang J. 2009b. SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics*, **25**, 1966–1967.

Li W, Liu N, Huang L, Chen Y, Guo J, Yu B, Luo H, Zhou X, Huai D, Chen W, Yan L, Wang X, Lei Y, Liao B, Jiang H. 2022. Stable major QTL on chromosomes A07 and A08 increase shelling percentage in peanut (*Arachis hypogaea* L.). *The Crop Journal*, **10**, 820–829.

Lu Y, Zhang S, Shah T, Xie C, Hao Z, Li X, Farkhari M, Ribaut J M, Cao M, Rong T. 2010. Joint linkage-linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 19585–19590.

Lucia R, Gomes F, Celis A, Lopes A. 2005. Correlations and path analysis in peanut. *Crop Breeding* & *Applied Biotechnology*, **5**, 105–110.

Luo H, Guo J, Ren X, Chen W, Huang L, Zhou X, Chen Y, Liu N, Xiong F, Lei Y, Liao B, Jiang H. 2018. Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, **131**, 267–282.

Luo H, Pandey M K, Khan A W, Guo J, Wu B, Cai Y, Huang L, Zhou X, Chen Y, Chen W, Liu N, Lei Y, Liao B, Varshney R K, Jiang H. 2019. Discovery of genomic regions and candidate genes controlling shelling percentage using QTL-seq approach in cultivated peanut (*Arachis hypogaea* L.). *Plant Biotechnology Journal*, **17**, 1248–1260.

Luo H, Xu Z, Li Z, Li X, Lv J, Ren X, Huang L, Zhou X, Chen Y, Yu J. 2017. Development of SSR markers and identification of major quantitative trait loci controlling shelling percentage in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, **130**, 1635–1648.

Meng X, Zhang J, Cui S, Chen C Y, Mu G, Hou M, Yang X, Liu L. 2021. QTL mapping and QTL×environment interaction analysis of pod and seed related traits in cultivated peanut (*Arachis hypogaea* L.). *Acta Agronomica Sinica*, **47**, 1874–1890. (in Chinese)

Meuwissen T, Goddard M. 2010. Accurate prediction of genetic values for complex traits by whole-genome resequencing. *Genetics*, **185**, 622–623.

Pawan K, Pandey M K, Wang H, Feng S, Qiao L, Culbreath A K, Sandip K, Wang J, Corley H C, Zhuang W. 2016. Mapping quantitative trait loci of resistance to tomato spotted wilt virus and leaf spots in a recombinant inbred line population of peanut (*Arachis hypogaea* L.) from SunOleic 97R and NC94022. *PLoS ONE*, **11**, e0158452.

Ravi K, Vadez V, Isobe S, Mir R R, Guo Y, Nigam S N, Gowda M, Radhakrishnan T, Bertioli D J, Knapp S J. 2011. Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, **122**, 1119–1132.

Robledo G, Lavia G I, Seijo G. 2009. Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. *Theoretical and Applied Genetics*, **118**, 1295–1307.

Verhertbruggen Y, Bouder A, Vigouroux J, Alvarado C, Geairon A, Guillon F, Wilkinson M D, Stritt F, Pauly M, Lee M Y, Mortimer J C, Scheller H V, Mitchell R A C, Voiniciuc C, Saulnier L, Chateigner-Boutin A L. 2021. The *TaCslA12* gene expressed in the wheat grain endosperm synthesizes wheat-like mannan when expressed in yeast and *Arabidopsis*. *Plant Science*, **302**, 110693.

Wang L, Yang X, Cui S, Mu G, Sun X, Liu L, Li Z. 2019. QTL mapping and QTL × environment interaction analysis of multi-seed pod in cultivated peanut (*Arachis hypogaea* L.). *The Crop Journal*, **7**, 249–260.

Wang Z, Yan L, Chen Y, Wang X, Huai D, Kang Y, Jiang H, Liu K, Lei Y, Liao B. 2022. Detection of a major QTL and development of KASP markers for seed weight by combining QTL-seq, QTL-mapping and RNA-seq in peanut. *Theoretical and Applied Genetics*, **135**, 1779–1795.

Zhang C, Zhou Z, Yong H, Zhang X, Hao Z, Zhang F, Li M, Zhang D, Li X, Wang Z. 2017. Analysis of the genetic architecture of maize ear and grain morphological traits by combined linkage and association mapping. *Theoretical and Applied Genetics*, **130**, 1011–1029.

Zhang S, Hu X, Miao H, Chu Y, Cui F, Yang W, Wang C, Shen Y, Xu T, Zhao L, Zhang J, Chen J. 2019. QTL identification for seed weight and size based on a high-density SLAF-seq genetic map in peanut (*Arachis hypogaea* L.). *BMC Plant Biology*, **19**, 537.

Zhuang W, Chen H, Yang M, Wang J, Pandey M K, Zhang C, Chang W C, Zhang L, Zhang X, Tang R, Garg V, Wang X, Tang H, Chow C N, Wang J, Deng Y, Wang D, Khan A W, Yang Q, Cai T, *et al.* 2019. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nature Genetics*, **51**, 865–876.

Zou C, Wang P, Xu Y. 2016. Bulked sample analysis in genetics, genomics and crop improvement. *Plant Biotechnology Journal*, **14**, 1941–1955.

Executive Editor-in-Chief Xueyong Zhang
Managing Editor Ning Wang