The Crop Journal xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Crop Journal

journal homepage: www.keaipublishing.com/en/journals/the-crop-journal/

Genetic identification of *AhGH15*, a gene under convergent selection during peanut growth habit domestication and breeding

Guanghui Chen [a,*], Yan Ren [a], Na Chen [a], Jiaxin Tan [b], Heng Wang [c], Hui Wang [c], Lijun Wu [a], Shuangling Li [a], Tianying Yu [b], Manish K Pandey [d], Xiaoyuan Chi [a,*], Mei Yuan [a,*]

[a] *Shandong Peanut Research Institute, Shandong Academy of Agricultural Sciences, Qingdao 266199, Shandong, China*
[b] *School of Life Science, Yantai University, Yantai 264005, Shandong, China*
[c] *Rizhao Agricultural Technology Service Center, Rizhao 276827, Shandong, China*
[d] *Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Hyderabad, India*

ARTICLE INFO

Article history:
Received 23 November 2025
Revised 22 January 2026
Accepted 16 March 2026
Available online xxxx

Keywords:
Peanut
Domestication
Growth habit
Plant architecture
QTL

ABSTRACT

Peanut (*Arachis hypogaea*), a globally important oilseed crop, exhibits contrasting growth habits between wild species (prostrate) and cultivated varieties (erect or spreading), but the underlying mechanism is unclear. In this study, we performed quantitative trait locus (QTL) mapping of two recombinant inbred line (RIL) populations of peanut using SNP arrays and bulk segregant analysis, which identified *qGH15* on chromosome 15 as a major QTL regulating growth habit. Fine mapping using KASP markers narrowed the candidate region to a 151-kb interval, while analysis of a residual heterozygous line (RHL) further delimited *qGH15* to a 38-kb interval containing a single candidate gene, which we designated as *AhGH15*. Genotyping of a natural population revealed multiple types of polymorphisms in this gene associated with the erect habit in cultivated varieties. Phylogenetic and pedigree analyses demonstrated that these polymorphisms were recurrently and convergently selected during peanut domestication and breeding, with modern hybridization accelerating their dissemination. Transcriptome deep sequencing and gene co-expression analysis via WGCNA revealed polymorphism-specific patterns of transcriptome regulation, with co-expressed gene modules differentially active between accessions with erect or prostrate growth habits. These findings establish *AhGH15* as a key determinant of peanut growth habit and highlight its complex selection history during peanut improvement.

1. Introduction

Peanut (*Arachis hypogaea* L.), also known as groundnut, serves as a valuable source of protein and oil for human consumption and livestock feed. This important legume crop is widely grown worldwide. In China, peanut accounts for approximately one third of the total planting area of oil crops (https://www.stats.gov.cn/). The ability of peanut to fix nitrogen makes it an important crop for sustainable agriculture.

The growth habit of peanut is an essential factor affecting its yield. Significant changes in growth habits have occurred during the domestication and breeding of peanuts from wild species to cultivated varieties. Wild peanut plants generally grow in a creeping manner, with prostrate growth. Their lateral branches extend and sprawl on the ground, with flowers growing at the nodes of the branches [1]. Through domestication and breeding, most modern cultivated peanut varieties have been developed in Asia and African countries by selecting plants with high yield and compact growth, i.e., an erect growth habit. This change in growth habit has made peanut cultivation and harvesting more efficient and convenient, thereby reducing labor costs and boosting productivity. However, erect growth can increase the distance between peanut pegs and the ground, potentially reducing pod number. Establishing an ideal plant architecture adapted to environmental conditions and cultivation methods has always been an important goal in crop breeding. During domestication, multiple mutations in genes regulating plant morphology have been retained. For instance, during rice (*Oryza sativa*) domestication, a crucial mutation in the transcription factor gene *PROSTRATE GROWTH 1* (*PROG1*)

* Corresponding authors.
E-mail addresses: ghchen@sdpeanut.ac.cn (G. Chen), chi000@126.com (X. Chi), yuanbeauty@126.com (M. Yuan).

https://doi.org/10.1016/j.cj.2026.02.017

Please cite this article as: G. Chen, Y. Ren, N. Chen et al., Genetic identification of *AhGH15*, a gene under convergent selection during peanut growth habit domestication and breeding, The Crop Journal, https://doi.org/10.1016/j.cj.2026.02.017

led to the transition from prostrate to erect growth [2–4]. However, the regulatory mechanism behind the growth habit of peanut is poorly understood.

Peanuts exhibit four different growth habits based on the morphology of the first pair of lateral branches and the main stem: erect, bunch, spreading, and prostrate [1,5]. Varieties with a prostrate or spreading habit have a higher yield per plant than the other two types but require more land area for cultivation. Cultivars with a bunch or erect habit can achieve higher planting densities, but the yield per plant is significantly lower due to the greater distance between the peg and the ground. In countries with a high degree of mechanization, such as the United States, most peanut varieties are of the prostrate or spreading types, while in countries with less developed mechanization, such as China and India, most varieties are bunch or erect. This difference may be closely related to local planting and manual harvesting methods.

Several studies have suggested that key quantitative trait loci (QTL) or genes controlling peanut growth habit are located on chromosome 15; this QTL is mainly associated with branch angle [6]. Bulked segregant analysis followed by high-throughput sequencing (BSA-seq) of F_2 and F_3 populations derived from crosses between prostrate and bunch cultivars showed that the spreading and bunch growth habits are controlled by a single QTL, which was mapped to a 1.1-Mb region on chromosome 15 [5]. Using a recombinant inbred line (RIL) population derived from a cross between a runner peanut variety and an erect peanut variety, Li *et al.* identified 39 QTL related to plant architecture, consisting of 6 associated with lateral branch angle, 8 with spreading diameter, 7 with plant type index, 11 with main stem height, and 7 with lateral branch length [7]. Among these, 12 QTL are located within a 168-kb region on chromosome 15. A genome-wide association study (GWAS) using data derived from the 48 K SNP chip for peanut from 103 core germplasm accessions detected 95 single nucleotide polymorphisms (SNPs) significantly associated with growth traits on chromosome B05. A subsequent BSA-seq analysis of the F_2 population for lateral branch angle finely mapped the causal locus to a 160-kb interval (144,193,467 bp to 144,513,467 bp) on chromosome B05, identifying the candidate growth habit gene *Araip.E64SW*, which encodes an F-box protein [8]. Diagnostic markers for growth habit traits were also developed in this genomic region [9]. Pan *et al.* further narrowed the location of the QTL controlling lateral branch angle to a 136.65-kb region (chromosome 15: 157,425,103 bp to 157,561,753 bp), identifying one F-box gene and two 2OG oxygenase-encoding genes as potential candidates regulating peanut growth habit [6]. GWAS suggested that the *AhATH5WE* (Ah15g52700) gene encoding a MADS-box transcription factor might be responsible for regulating the growth habit of peanut [10].

Despite the identification of many candidate genes that might regulate peanut growth habit, to date, no gene regulating this trait has been identified. Therefore, in the current study, we developed two recombinant inbred line (RIL) populations along with a residual heterozygous line (RHL) and used a combination of 48 K SNP array, BSA-seq, and genotyping with kompetitive allele-specific PCR (KASP) markers to finely map a gene related to growth habit. After classifying lines as having a prostrate or erect growth habit primarily based on branch angle, we identified a 38-kb interval on chromosome 15 harboring the MADS-box gene *AhATH5WE,* which we renamed here as *AhGH15*, and demonstrated that it affects growth habit in peanut. Analysis of this gene in natural populations indicated that erect-type peanuts have acquired multiple mutations in this gene during domestication and breeding, and that different types of mutations might have been selected for and preserved independently.

2. Materials and methods

2.1. RIL population construction and phenotypic analysis

To construct the RIL populations, the spreading-type peanut variety 'Sunoleic 97R' [11] was crossed to two erect-type accessions, namely 'A38' and 'Mi-2'. After successive generations of self-pollination, an F_5 RIL population named SA (Sunoleic 97R × A38) and an F_9 RIL population named SM (Sunoleic 97R × Mi-2) were developed, consisting of 293 lines (SA population) and 170 lines (SM population) [12]. Plants were grown in rows spaced 80 cm apart, with 30 cm between plants within each row. To quantitatively describe peanut growth habit, numerical values were assigned to different phenotypes: phenotypes similar to the erect parent were assigned a value of 1, while spreading phenotypes were assigned a value of 4. Intermediate phenotypes were given a value of 2 or 3.

2.2. Linkage map–based QTL analysis of the SM population

A genetic linkage map was previously constructed using the SM population [12]. Here, this map was used to perform QTL analysis for peanut growth habit traits with QTL-IciMapping (ICIM-ADD), with a significance threshold of logarithm of the odds (LOD) ≥ 3.0. A QTL that explained over 10.0% of the phenotypic variation was considered to be a major QTL. The SNP data from the SM population were also used for GWAS with the rMVP tool using the mixed linear model (MLM; PCA + Kinship) model [13,14]. Specifically, the kinship matrix (K) and principal components (PCs) were calculated in rMVP during data preparation (MVP.Data with fileKin = TRUE and filePC = TRUE), and the MLM was fitted by including three PCs (PC1–PC3) to control for residual population structure (nPC.MLM = 3). Manhattan plots were generated using the R package CMplot, with all *P*-values converted to $-\log_{10}$ (*P*-values) [13].

2.3. RHL population construction and genotyping

To ensure a near-isogenic background for fine mapping, an RHL population was generated from 27 heterozygous F_3 plants of the SA population by preliminary genotyping using the flanking markers M-1455 and M-1461, which delimit the *qGH15* candidate interval. These plants yielded 1635 F_4 seeds, with each seed representing a unique RHL. In the F_5 generation, five plants per progeny were grown in each row. One plant was randomly selected from these five plants for genotyping and phenotyping as a representative of each family to minimize sampling bias and to confirm the genotype–phenotype relationship.

2.4. BSA-seq analysis

For BSA-seq analysis, two pools were obtained from genomic DNA for 20 spreading individuals (S pool) and 20 erect individuals (E pool) from the F_5 population of the SA population; the genomic DNA from each individual was mixed in equal amounts for library construction and sequencing. Libraries for the parents and the two pools were sequenced using an Illumina NovaSeq 6000 instrument at Biomarker Technologies Co., Ltd. (Beijing, China). Quality control was performed using FastQC (https://github.com/s-andrews/FastQC). Low-quality and short reads, as well as adapters, were removed using Trimmomatic [15] (https://github.com/timflutre/trimmomatic). High-quality paired-end reads were aligned to the reference genome Tifrunner.gnm2.J5K5 (https://www.peanutbase.org) [16] using BWA or Hisat2 with the parameter '-no-spliced-alignment' [17,18] (Tables S1 and S2). Reads that aligned to multiple

sites were removed using SAMtools [19], and SNP/insertion/deletion (InDel) calling and filtration were performed using GATK4.0 [20]. SNPs from both pools were filtered to remove low-quality and low-depth sites, and homozygous SNPs with no differences between the parents were excluded from analysis. BSA-seq analysis was conducted using the R package QTLseqr [21], and the Delta SNP-index and G' methods were used for QTL identification. Dot plot figures were generated using a Python script to visualize the results.

2.5. Marker development and fine mapping

For fine mapping, aligned BAM files from the two pools and the reference genome annotation in GFF3 format were opened in Integrated Genome Viewer (IGV, v2.16.1) [22,23]. Sixteen SNPs in the candidate region were selected for KASP marker development. All KASP primers are listed in Tables S3 and S4. These markers were used to genotype individuals from the SA and SM RIL populations, the RHL population, and a natural population consisting of 494 germplasm accessions [24].

2.6. RNA-seq analysis

The first pair of lateral branches from Sunoleic 97R and A38 were divided into four sections from top to bottom (Fig. S1). For each genotype × section combination, three independent biological replicates were collected from three individual plants grown at different positions under the same growth conditions. Total RNA was extracted from the samples using TRIzol reagent (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA; Cat. No. 15596026). In addition, for WGCNA, three erect cultivars with a gap within the *AhGH15* sequence (GAP thereafter) and one erect cultivar with a transposable element (TE) insertion in *AhGH15* were sampled at three stages (stages 0–2, Fig. S1) with two biological replicates per genotype × developmental stage combination. Library construction and sequencing were performed by Biomarker Technologies. After quality control, low-quality and short reads, along with adapters, were removed using Trimmomatic [15]. High-quality paired-end reads were aligned to the Tifrunner.gnm2.J5K5 and Tennessee Red (http://www.ncbi.nlm.nih.gov/nuccore/JAKNSI000000000.1) reference genomes using Hisat2 [17,18]. Differential expression analysis was performed using featureCounts and DESeq2 [25], while StringTie was used to identify previously unannotated transcripts [26]. The IGV tool [22,23] was employed to visualize and compare differences in splicing sites. WGCNA was conducted using the Python package pyWGCNA [27].

2.7. Polymorphism analysis of natural variants

To explore different polymorphism types of *AhGH15*, public NCBI SRA data were used (Table S5). Sequencing data in FASTQ format were aligned to the Tifrunner V2 reference genome after quality control. The resulting BAM files were opened with IGV to review structural variants of *AhGH15* [22,23]. Genomic DNA was extracted from each germplasm that showed a consistent phenotype across three years using a Plant DNA Extraction kit (TIANGEN, Cat: DP350). KASP markers were used to detect InDel polymorphisms in *AhGH15* across all samples. For samples with an erect phenotype but no InDel in *AhGH15*, three PCR primers were designed to detect the GAP-type mutation in the natural population in our laboratory. Regional *de novo* assembly was performed using ABySS and SPAdes [28,29]. The Tennessee Red genome sequence (GCA_022829005.1) was also used to identify structural variants in the *AhGH15* gene.

2.8. Evolutionary analysis of the genomic region containing AhGH15

Publicly available sequencing data from peanut were downloaded from NCBI [30]. Raw sequences were quality-controlled and aligned to the Tifrunner reference genome as described above. SNP calling was performed using GATK4, generating a VCF file [20]. A Python script was used to extract SNP variant data within a 2-Mb region upstream and downstream of *AhGH15*, before converting the data to PHYLIP format using the vcf2phylip script (https://github.com/edgardomortiz/vcf2phylip). A phylogenetic tree was reconstructed using MEGA with the neighbor-joining method and 1000 bootstrap replications. The tree was viewed and annotated using Evolview [31].

2.9. Pedigree analysis

Pedigree data were obtained from the National Peanut Data Center of China (https://peanut.cropdb.cn/variety/)[32]. Gene flow networks were reconstructed using Cytoscape software [33], where the two parents of each variety were defined as source nodes, and the corresponding progeny variety as the target node. Nodes were colored by genotype class, with insertion (INS) in blue, GAP in magenta, wild type in green, and unknown genotypes in white. For genotype prediction, if the two parents shared the same genotype, the progeny was assigned that genotype; otherwise, the genotype of the progeny was recorded as unknown. When the progeny genotype was known and differed from the known genotype of one parent, the genotype of the other parent (with an unknown genotype) was predicted to be the same as the progeny genotype. Predicted nodes are shown in different colors to distinguish them from experimentally confirmed nodes. Node size was scaled by parental usage frequency.

3. Results

3.1. A major QTL controls peanut growth habit

We developed two RIL populations, designated SM and SA, by crossing the spreading-type peanut cultivar SunOleic 97R (male parent) to the erect-type peanut accessions Mi-2 and A38 (female parents), respectively (Fig. 1A). Using phenotypic data collected from all RILs and a previously generated linkage map for the SM population [12], we identified a major QTL on chromosome 15 in the region between markers AX-147251167 and AX-147251194. This QTL, called *qGH15*, spanned the physical region from 156,546,816 bp to 157,156,268 bp on this chromosome, explaining 37%–49% of the phenotypic variance, with LOD scores of 25–29 (Figs. 1B, S2A; Table S6). We also used SNP data for GWAS and used the $-\log_{10}(P\text{-values})$ to draw the corresponding Manhattan plot (Figs. 1C, S2B). We used two cutoffs: $-\log_{10}(P\text{-value}) \geq 5$ as a suggestive threshold and $-\log_{10}(P\text{-value}) \geq 10$ as a high-confidence threshold. Marker AX-176811452, located between AX-147251167 and AX-147251194, had the highest $-\log_{10}(P\text{-value})$.

Since the SNP density of the SNP array was insufficient for fine mapping and marker development, we conducted BSA-seq analysis via genome resequencing of the parents and two DNA pools from plants with either extreme phenotype from the SA population. The sequencing depth was at least 30×, with over 97% of the reads being paired (Table S1). We aligned all paired reads to the Tifrunner V2 reference genome. Due to the allotetraploid nature of the peanut genome, highly homologous sequences between subgenomes can lead to multiple alignments. Although the overall

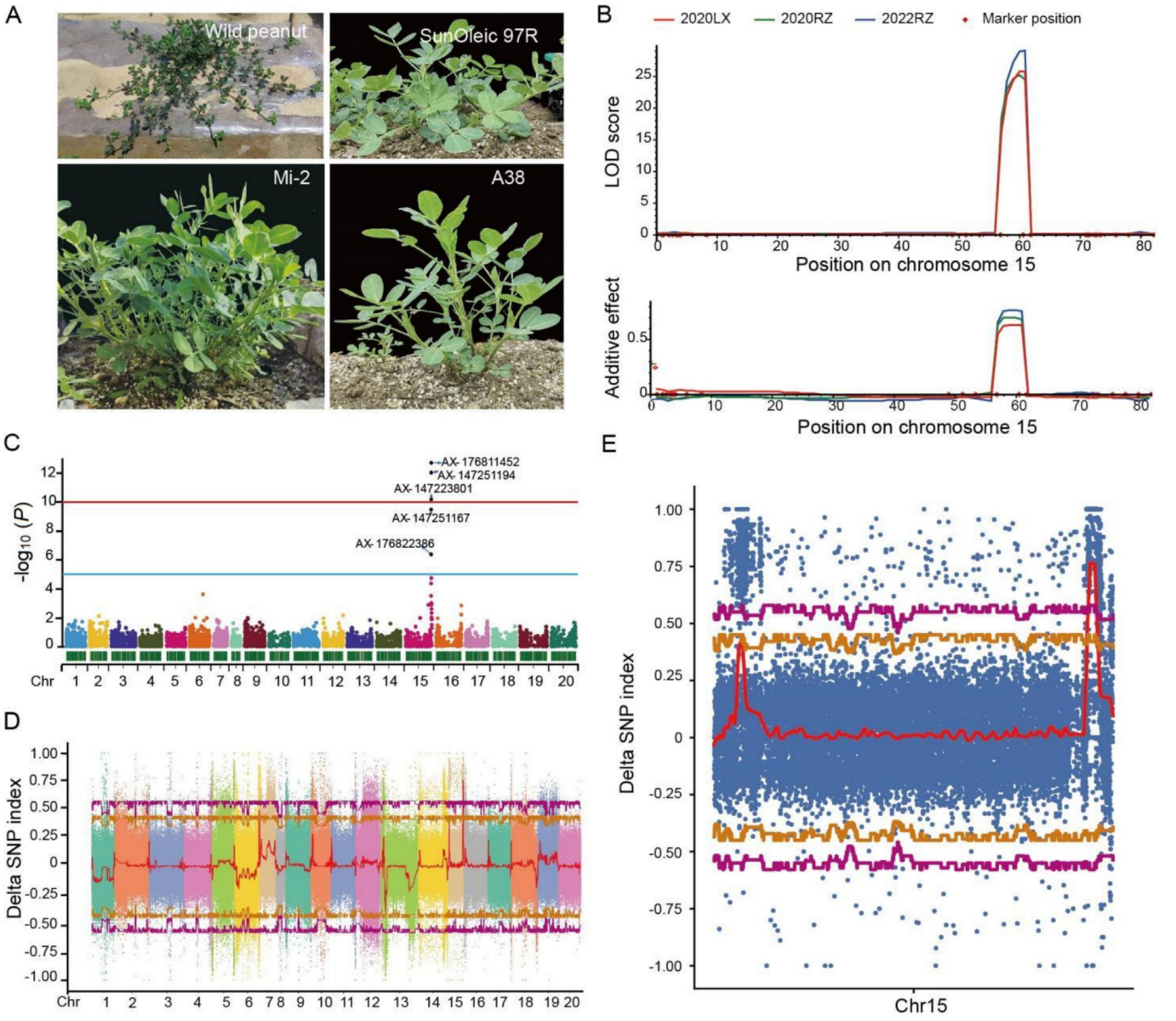

Fig. 1. *Qgh15* is a conserved QTL regulating peanut growth habit. (A) Representative photographs of peanut accessions with prostrate (Wild), spreading (SunOleic 97R), or erect (Mi-2, A38) growth habit. (B) Identification of the *qGH15* on chromosome 15 using linkage mapping. (C) Manhattan plot showing the results of a genome-wide association study (GWAS) for growth habit. The *P*-values of SNP markers along each linkage group were converted to $-\log_{10}$(*P*-value). The horizontal lines define the thresholds of significance at $P = 1.64 \times 10^{-5}$ (blue) and $P = 1.64 \times 10^{-10}$ (red). (D) Dot plot of the genomic distribution of Δ (SNP index) revealed by BSA-seq analysis. The red curve represents the average Δ (SNP index) calculated with the sliding window algorithm. The purple horizontal and orange horizontal lines represent significance levels of 0.01 and 0.05, respectively. (E) Enlarged plot of the results of BSA-seq analysis for chromosome 15.

alignment rate reached 96%, only 84% of the sequences aligned uniquely, with the remaining 12% aligning to multiple locations (Table S2). To assess whether multiple alignments affected the BSA-seq results, we ran the analysis pipeline twice using both uniquely aligned reads and all aligned reads.

Using the delta SNP-index method, we identified two QTL with 99% confidence on chromosomes 6 and 15, and an additional QTL on chromosome 13 when dropping the confidence threshold to 95% (Figs. 1D, E, S2; Table S7). The QTL on chromosome 15 was located in the 155,325,885 bp to 158,446,960-bp region. The markers flanking *qGH15* (AX-147251167 and AX-147251194) detected in the SM population were also located within this region, suggesting that *qGH15* is a common and major QTL controlling growth habit in peanut.

3.2. Fine mapping of qGH15

We combined the linkage map–based QTL mapping data from the SM population with the BSA-seq-based QTL mapping data from the SA population and specifically analyzed the genotype data of each individual in the SM population, we identified the candidate interval for *qGH15* to be between markers AX-147251167 and AX-147223801. This interval spans the physical location from 156,546,816 bp to 157,502,133 bp on chromosome 15. We developed 15 SNP markers for genotyping of each line in the SA and SM populations (Tables S3, S4) and identified eight recombination events in 22 erect plants (Fig. 2; Tables S8, S9). Ultimately, we narrowed down *qGH15* to a 152-kb interval (Chr15: 156,921,591–157,073,376) between markers M-15692 and M-1460 (Fig. 2).

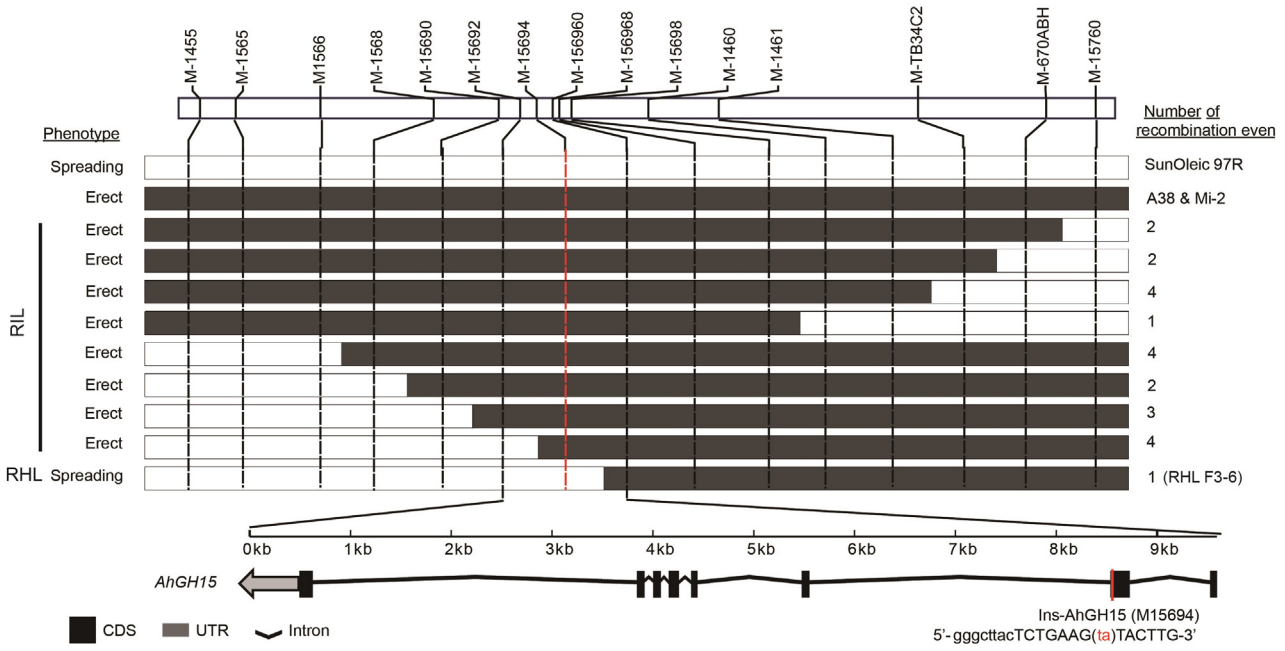

Fig. 2. Fine mapping of *AhGH15*. The bar at the top shows the relative positions of the markers along a fragment of chromosome 15. The black bars indicate genotypes identical to the erect parent, while the white bars indicate genotypes identical to the spreading parent. The phenotype (spreading; erect) is shown on the left, and the numbers on the right indicate the number of recombination events in the population. Sequences in parentheses indicate inserted nucleotides, uppercase letters represent exonic sequences, and lowercase letters represent intronic sequences.

Although we developed four additional molecular markers (M-15694, M-156960, M-156968, M-15698) within this region, we failed to detect new recombination events in the SA and SM RIL populations.

According to the Tifrunner V2 genome annotation, 14 protein-coding genes are located within this 152 kb interval. Among the SNPs and InDels in this region, seven were located within regions harboring four genes: *ATH5WE* (AH15G527000), *UV4XXD* (AH15G527400), AH15G527500, and *U0QS20* (AH15G527600); the other polymorphisms were located in intergenic regions (Table S10). After excluding SNPs located in introns and those causing synonymous changes, we identified one SNP and one InDel that could potentially result in changes in the coding sequence. In *AhATH5WE*, we identified a 2 bp insertion adjacent to the 5′ splice site of the second exon, which could cause a splicing error or frameshift mutation. Additionally, the SNP in AH15G527500 resulted in the substitution of arginine with asparagine (Table S10).

To generate plants with new recombination events for further fine mapping, we selected 27 reserved F_3 plants from the SA population that were heterozygous at markers M-1455 and M-1461. All 27 F_3 plants showed a spreading growth habit. We used their progeny population of 1635 plants as candidate RHLs. We performed genotyping and phenotyping of these plants in the F_5 generation using a single individual from each line. By genotyping the RHL population with the above markers, we detected one recombination event between markers M-15692 and M-156960 in 25 F_5 lines originating from line F3-6 (Fig. 2; Table S11). Based on this analysis, we delineated *qGH15* to a 38-kb interval between markers M-15692 and M-156960. There is only one gene within this interval: *AhATH5WE*. This gene encodes a MADS-box transcription factor and is homologous to the Arabidopsis *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*)-like family gene *AGAMOUS-LIKE 42* (*AGL42*). Notably, the erect parent in the mapping population carried a 2-bp insertion adjacent to the 3′ splice site of the second exon of this gene (Fig. 2). We used the KASP marker M-15694, which detected the 2 bp insertion in this gene, to genotype the entire RHL population. All erect-type plants were homozygous for the 2-bp insertion, strongly suggesting that *AhATH5WE* affects the growth habit of lateral branches in peanut. We have renamed this gene *Arachis hypogaea Growth Habit on chromosome 15* (*AhGH15*). The 2-bp insertion polymorphism identified in the *AhGH15* gene in this study was specifically designated as an insertion (INS) type mutation. The functional variation of *AhGH15* caused by the 2-bp insertion might be the primary factor contributing to the phenotypic changes in growth habit.

3.3. Different polymorphism types in AhGH15 in natural variants

We used the KASP marker M-15694 to survey polymorphism types including the INS type across natural populations and breeding materials (Table S4). Among the 284 erect varieties examined, 158 contained the INS-type polymorphism (Fig. 3B; Table S12). Notably, erect varieties lacking the INS-type polymorphism included 'Xuzhou 68-4', a widely used backbone germplasm in China, and 'Silihong', a popular Valencia-type variety in China. This observation suggests that other types of polymorphisms are present in *AhGH15* that might have been independently selected and retained during breeding.

To explore this notion, we aligned sequencing data from NCBI for the two above varieties to the reference genome, revealing new types of polymorphisms in *AhGH15*. In 'Xuzhou 68-4', we identified a 1.86-kb gap with no aligned reads in the second intron of *AhGH15* (Fig. 3A). PCR amplification followed by sequencing using newly designed primers flanking this gap (Table S13) confirmed that the gap is caused by a 1865-bp deletion (Figs. S3, S4). We named this polymorphism type the "GAP" type. We used these primers to detect this polymorphism in a natural population. Among the accessions lacking the INS polymorphism, 85 out of 104 contained the GAP-type polymorphism (Fig. 3B).

Neither the INS-type nor the GAP-type polymorphism was found in the genome of 'Silihong' or 'S3', an erect peanut variety with black seed coats grown in our laboratory. However, we

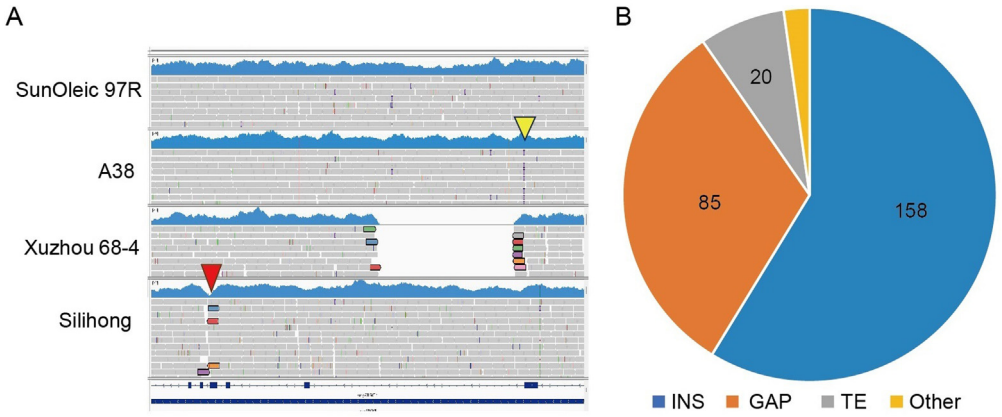

Fig. 3. Different polymorphism types in *AhGH15* across a natural population. (A) Alignment of sequencing data from SunOleic 97R (wild type), A38 (INS type), Xuzhou 68-4 (GAP type), and Silihong (TE type) to the reference genome. The triangles represent the polymorphism sites. (B) Pie chart showing the percentage of each polymorphism type among erect peanut accessions.

detected an unaligned region near the fifth exon of *AhGH15* in the alignment data (Fig. S5). This pattern typically indicates the presence of a large insertion causing structural variation. To determine the sequence in this region, we performed regional *de novo* assembly using pairs of reads characterized by one read aligning to *AhGH15*, while their paired second read aligned elsewhere in the genome; we also used reads adjacent to these pairs of reads (Fig. S6). *De novo* assembly generated two contigs, revealing a 1117-bp putative TE inserted into the fifth intron of *AhGH15*. This polymorphism was also present in the genome of 'Tennessee Red', an erect Valencia-type peanut cultivar with a publicly available *de novo* assembled genome (NCBI). Sequence analysis showed that the 1117-bp TE in Silihong and S3 is a typical hAT-type DNA transposon containing 13-bp terminal inverted repeats (TIRs) and 8-bp flanking direct repeats (FDRs), which are characteristic features of "cut-and-paste" DNA transposons (Fig. S7) [34]. This TE is widespread across the peanut genome (Table S14). We developed a PCR marker for the precise detection of this "TE" type mutation (Fig. S8; Table S12). Twenty varieties exhibited this polymorphism in our natural population, most being Valencia types.

3.4. Independent selection of AhGH15 during peanut breeding

In addition to the three major polymorphisms mentioned above, we also identified numerous rare mutation types of *AhGH15* in Chinese landraces and in publicly available resequencing data from NCBI (Fig. S9A–G). These polymorphisms included early translation stop codons caused by SNPs (e.g., in accession ICG14466, Fig. S9A, B) and exon deletions (e.g., in accession ICG15190 and ICG9037, Fig. S9C–G). Among the germplasm resources we collected, GAP-type polymorphisms were present in approximately 40% of erect-type peanuts (Fig. 3B). However, a geographical survey of all types of polymorphisms based on publicly available sequencing datasets revealed that this GAP type is absent in varieties from outside China (Fig. S9H).

We reconstructed a phylogenetic tree using SNP data within a 2-Mb interval centered on AhGH15 across different peanut germplasm resources. The polymorphism types were closely associated with specific accessions: the INS-type polymorphism appeared to be more prevalent in A. hypogaea subsp. fastigiata var. vulgaris, while the TE-type polymorphism was mainly found in subsp. fastigiata var. fastigiata. Notably, the GAP-type polymorphism clustered within the subsp. hypogaea clade (Fig. 4), which includes spreading landraces from China, such as 'Jintangshenwozi', 'Hexiandazidou', and 'Luyidapayang'. Accessions harboring the rare mutation types identified in Chinese landraces were also clustered within the subsp. hypogaea clade as well (Fig. 4). Notably, the erect variety 'Jihuatian 1', carrying the INS polymorphism, was also present in the subsp. hypogaea clade. Further analysis of the SNP patterns in 'Jihuatian 1' suggested that this clustering may have resulted from a recombination event between an INS-type and a GAP-type polymorphism or wild-type peanut varieties (Fig. S10).

To explore why numerous materials with the GAP-type polymorphism are present in China, we conducted pedigree analysis of the GAP-type and some INS-type erect peanut varieties that we had collected (Table S15) using Cytoscape software. Based on pedigree maps, among domestic peanut varieties, most GAP-type varieties were likely derived from the backbone variety 'Xuzhou 68-4', which is widely used as a parental variety in peanut breeding in China. This variety may have originally come from the Chinese landrace 'Shuyangdazhanyang'. Varieties such as 'Yuhua15', 'Yuhua7', and 'Xuzhou 68-4' may have played key roles as critical nodes during the expansion of the proportion of GAP-type varieties (Fig. 5).

3.5. Different polymorphism types lead to alternative expression of AhGH15 through distinct mechanisms

Since cultivated peanut is an allotetraploid plant, genes from each of its two subgenomes share a high degree of sequence similarity. However, the growth habit of peanut appears to be controlled by a single gene. To investigate this apparent contradiction, we compared the coding sequence of *AhGH15* to that of its homologous gene in the A subgenome, namely *AhGH05*, annotated as *AhH0ZZJY* in the Tifrunner genome. We identified a frameshift mutation caused by a 1-bp deletion in the homologous locus (Fig. S11), resulting in a truncated transcript shorter than that of *AhGH15*.

In addition to the INS-type polymorphism within the exon of *AhGH15*, we identified 13 SNPs in the intergenic region between markers M-15692 (SNP-XAW92V) and M-156960 (SNP-R9TI2M). To investigate whether these SNPs in the intergenic region affect the expression of their adjacent genes, we measured *AhGH15* expression levels in the first lateral branch internode of Sunoleic 97R and A38 using transcriptome deep sequencing (RNA-seq). A heatmap based on FPKM values showed that only two genes exhibit significant differences in expression as a function of variety and

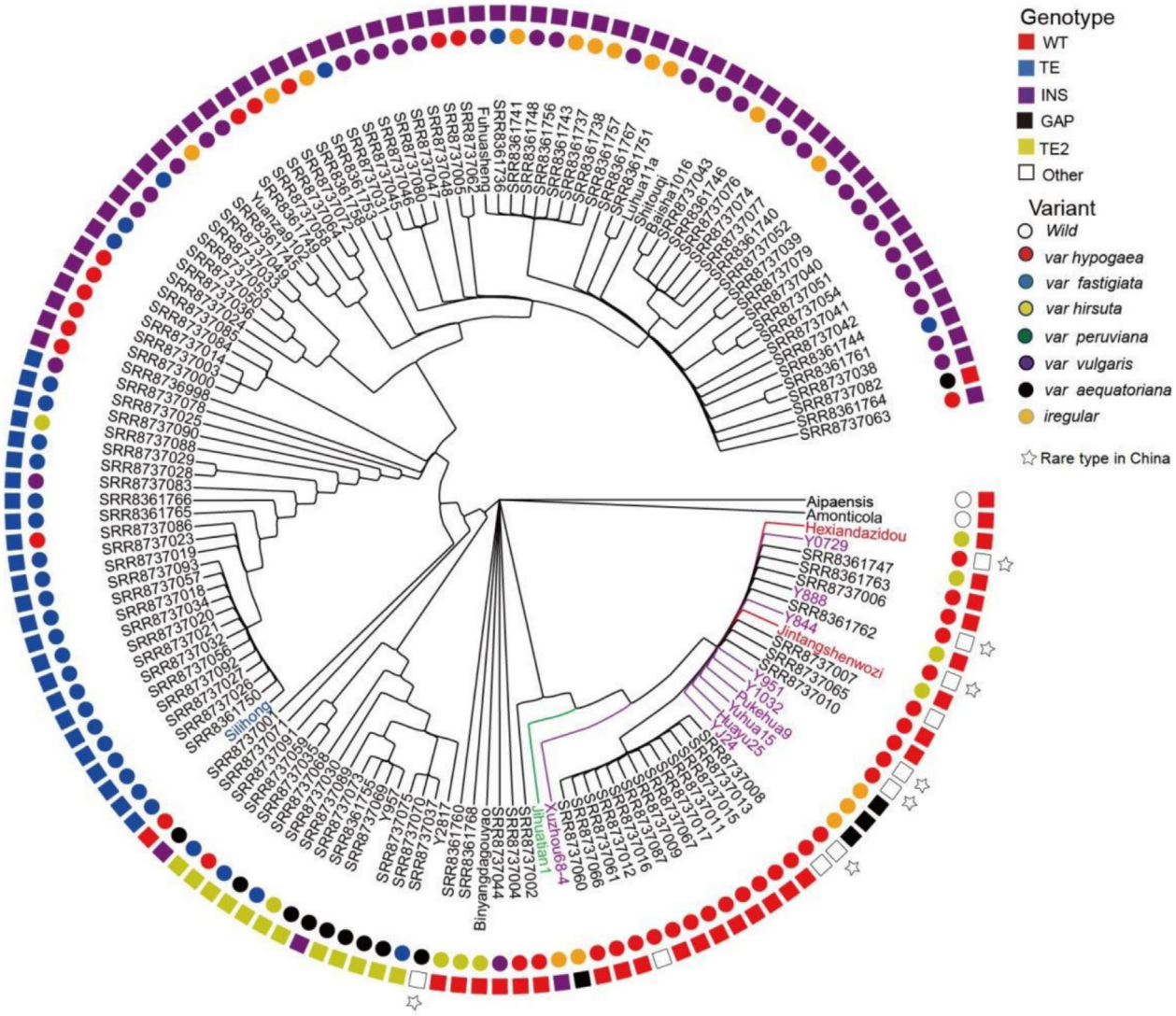

Fig. 4. Phylogenetic tree reconstructed using SNPs in the 2-Mb genomic region centered on *AhGH15*. *Arachis ipaensis* and *Arachis monticola* were selected as outgroups. Clades with purple text and white stars represent the GAP-type and rare type erect variants in China, while red clades indicate spreading-type landraces in China. The clade marked with green represents Jihuatian 1. The TE type 'Silihong' is highlighted with blue text and blue stars.

developmental stage (Fig. S12A). Of these two, only *AhGH15* displayed a growth habit–related expression pattern, with high expression in the internodes of basal lateral branches from spreading-type peanut (Figs. S1, S12B, C). *AhGH15* also showed a spatial expression gradient along the internodes of lateral branches, with high expression in internodes close to the main stem and low expression in internodes farther away from the main stem (Figs. S1, S12B, C).

Given that the INS-type polymorphism is located near the 3′ splice site of the second exon, we examined whether it affects splicing by merging the aligned BAM files from the SunOleic 97R and A38 transcriptomes into two files for comparison. In A38, the intron between the second and third exons of *AhGH15* was not properly spliced, leading to the loss of the first three exons in the transcript (Fig. 6). This splicing defect resulted in a truncated transcript in plants with an erect growth habit, likely leading to a loss or significant change in gene function. The GAP-type polymorphism in *AhGH15* is caused by a large fragment deletion in the first intron that does not affect the coding sequence. However, very few reads were aligned to the coding sequence of *AhGH15* in the RNA-seq data (Fig. 6), suggesting that the intron region is crucial for its expression. For the TE-type polymorphism, the TE insertion altered the splicing site for the fifth exon, resulting in a 47-bp frameshift (Figs. 6, S13).

3.6. WGCNA suggests the involvement of AhGH15 in lateral branch development

We collected 36 samples for RNA-seq from both spreading peanuts and erect peanuts with different polymorphism types across various stages of lateral branch development and performed weighted gene co-expression network analysis (WGCNA). Two co-expression modules (Linen and Lightsalmon) were highly correlated with the phenotypic traits of spreading or erect growth, respectively, as well as with the genotype at *AhGH15* (Pearson's correlation coefficients of 0.99 and −0.94, respectively) (Fig. 7A). The genes belonging to the Linen module exhibited low expression in the samples from the erect type and high expression in those from the spreading type (Fig. 7B), while the genes of the Lightsalmon module showed the opposite expression pattern (Fig. 7C).

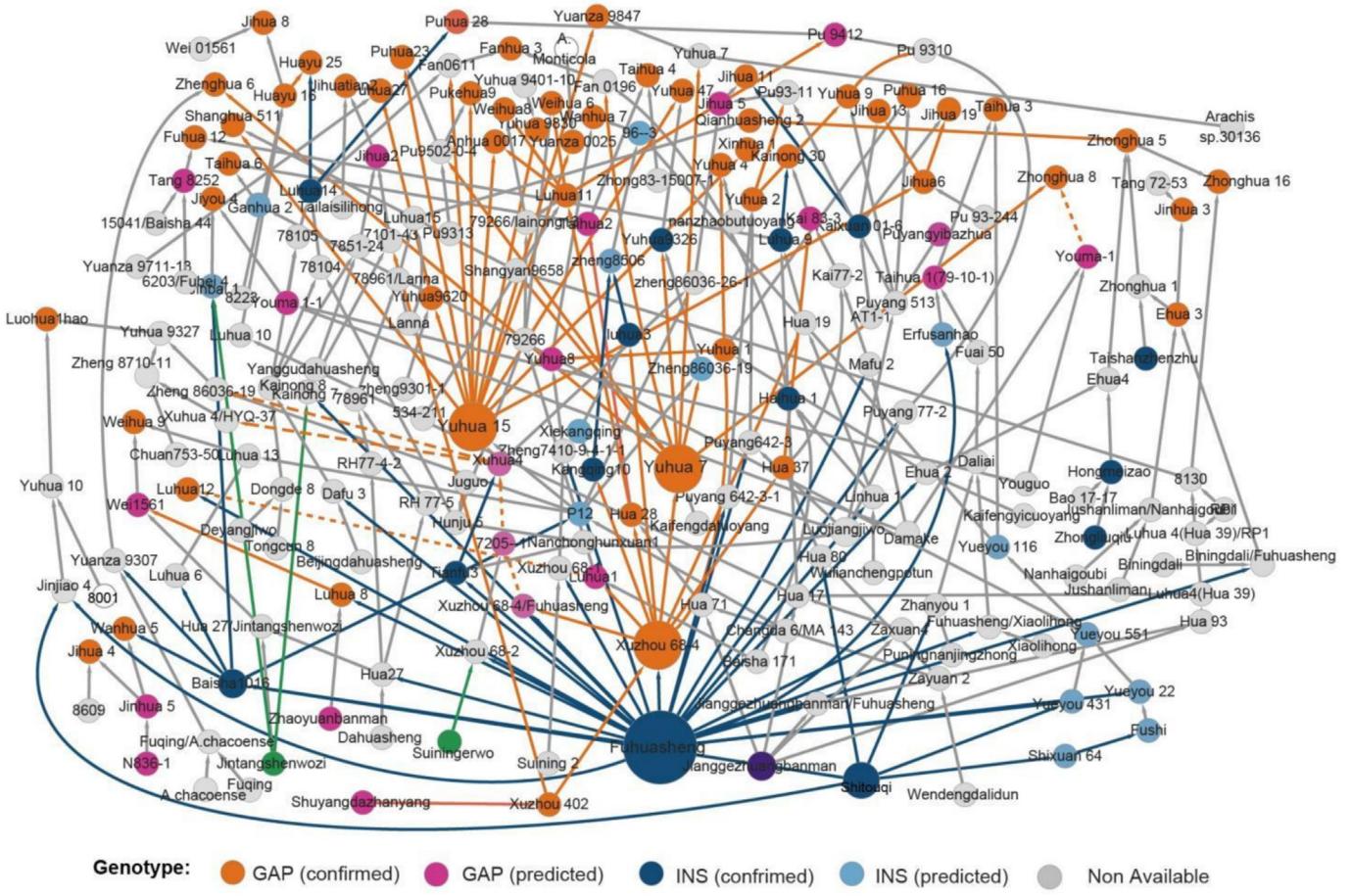

Fig. 5. Pedigree analysis of varieties with the GAP-type polymorphism in China. Colored nodes represent the confirmed and predicted genotypes of each variety. Lines with arrows indicate the transmission of genotypes to the next generation, with solid lines showing confirmed transmissions and dashed lines representing predicted transmissions. GAP-type polymorphisms are shown with orange lines and INS-type polymorphisms with blue lines. The size of each node is proportional to the number of edges connected to it.

Notably, *AhGH15* is part of the Linen module (Fig. S14; Table S16), further supporting the strong correlation between *AhGH15* and the growth habit of peanuts.

The Linen module comprises 1466 genes that are potentially co-expressed with *AhGH15* (Table S16). These genes include downstream genes that might be directly or indirectly regulated by AhGH15, as well as genes encoding proteins that regulate *AhGH15* expression. Annotation of homologous genes in Arabidopsis suggested that *AhGH15* might be involved in multiple developmental and environmental signaling pathways. These pathways include the brassinosteroid (BR) metabolism and signaling pathway (*BSK3*, *BAK1*, *BEN1*, *CAS7*, *EBS7*), gibberellin (GA) pathway (*GA20OX*, *GA2OX*, *MYB62*, *GID1B*), auxin signaling pathway (ARF transcription factor genes), and pathways involved in light signaling and shoot apical meristem development (*CRY2*, *PHYA*, *STM*) (Figs. 7D, S15). Vascular bundle development, which plays a direct role in shaping lateral branch growth, is also likely influenced by *AhGH15*. Several genes co-expressed with *AhGH15* are involved in vascular bundle formation, such as *NST1* and *NAC007*, which encode key transcription factors in xylem formation, vascular bundle development, and secondary cell wall formation. Additionally, genes encoding key enzymes for lignin and cellulose biosynthesis, including *COMT1*, *LAC2*, *CESA*, and *CSLA02*, were co-expressed with *AhGH15*. Analysis of the promoter regions of these co-expressed genes indicated that many contain one or more CArG boxes (C[C/A/T][A/T]$_6$[A/G]G), which are specific binding sequences for MADS-box transcription factors, suggesting that AhGH15 might directly regulate the expression of these genes.

4. Discussion

4.1. AhGH15 is associated with growth habit in peanut

Increasing planting density is a key approach to improving crop yield. During domestication, polymorphisms and mutations that promote traits suitable for dense planting are often selected and preserved. *PROG1* has played an important role in rice domestication, contributing to erect growth habit, increased grain number, and higher yields when inactivated [2–4]. Indeed, 182 cultivated rice varieties share the same polymorphisms in *PROG1*, which likely became fixed during domestication [3]. Several other genes, such as *TILLER ANGLE CONTROL 1* (*TAC1*), *LAZY1* (*LA1*), *IDEAL PLANT ARCHITECTURE 1* (*IPA1*), and *Teosinte branched 2* (*OsTb2*), also control tiller angle and branching in rice, highlighting a conserved regulatory network across plant species [4,13,35,36]; homologs of these genes regulate plant architecture in a similar manner in Arabidopsis [13,37].

Although key genes controlling growth habit have been identified in rice and other species, the specific genes that control the changes in growth habit that occurred during peanut domestication remain unclear, in part because cultivated peanuts are tetraploid, and the high genomic redundancy results in the lack of

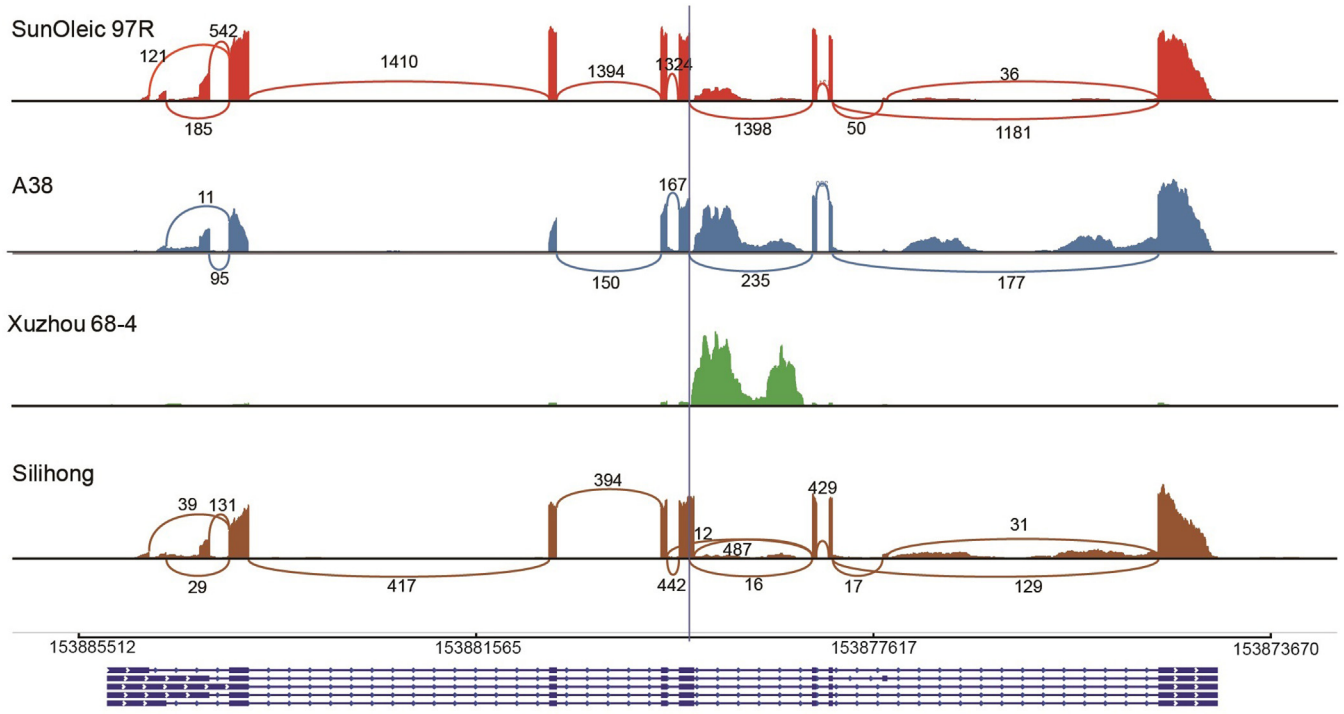

Fig. 6. Sashimi plots of *AhGH15* transcripts in four types of peanuts. Genomic reads were converted into read densities, and junction reads were plotted as arcs. The numbers represent the number of reads aligned to the junction spanning the exons.

phenotypic expression for mutations in single genes. Previous QTL mapping studies using RIL populations derived from crosses between peanut accessions with different growth habits have pointed to a key region on chromosome 15 associated with the upward growth of lateral branches [5,9]. Different researchers have selected different candidate genes within this region [5–9]. In the current study, we used two RIL populations and SNP array–based mapping to identify a major QTL on chromosome 15, explaining approximately 37%–49% of the phenotypic variation in growth habit in different environments (Table S6), which is consistent with other reports [5–9]. We fine-mapped the gene affecting peanut growth habit to a region between markers M-15692 and M-156960 on chromosome 1, an interval that contains only one gene, *AhATH5WE*, which we renamed *AhGH15*. An analysis of a natural population revealed that *AhGH15* was the only gene within this genomic region with different types of polymorphisms in erect-type peanuts (Fig. S9). Although genetic complementation experiments were not conducted due to technical limitations, the combined evidence strongly suggests that *AhGH15* is the key gene affecting peanut growth habit.

4.2. Erect-type peanuts originated from multiple independent polymorphisms and parallel selection

The re-domestication of wild species is an important approach to addressing many current agricultural challenges [38–40]. During crop domestication and breeding, polymorphisms and mutations in genes that are beneficial for yield often undergo multiple instances of parallel selection [41]. Furthermore, the types of polymorphisms utilized during the early stages of domestication constitute a large proportion of the extent of genetic variation in crops [3]. In rice, the green revolution gene *Semidwarf 1* (*SD1*) possesses several distinct alleles, with each variant potentially creating plants of different heights. Among these, three variants are present in *japonica* rice, and seven are present in *indica* rice [42]. Artificial selection has played an important role in the fixation of these polymorphisms [43].

Similarly, in this study, we discovered that *AhGH15* exhibits multiple distinct types of polymorphisms in peanut (Fig. S9). In the mapping population, the erect parent carried a 2-bp insertion at the 3′ side of the exon 2 splice site of *AhGH15*, leading to incorrect mRNA splicing and the production of truncated transcripts. Approximately 60% of the natural erect variants in our collection contained this 2-bp insertion polymorphism. We also identified other types of polymorphisms, such as GAP-type and TE-type polymorphisms, in *AhGH15* across different peanut germplasms. Transcriptome analysis showed that different allelic forms of *AhGH15* lead to structural changes or lower expression of the transcript variants of this gene through distinct mechanisms, which may ultimately result in the functional inactivation of *AhGH15* (Fig. 6). Phylogenetic analysis showed that the different polymorphism types, INS, GAP, TE, and other rare polymorphisms, clustered with different variants in the phylogenetic tree (Fig. 4), suggesting that erect-type peanuts have different origins. Furthermore, analysis of the global distribution of these polymorphism types revealed that GAP-type erect peanuts are widely distributed in China, with significantly more of these varieties than in other countries. According to phylogenetic analysis, GAP-type peanuts likely originated from subspecies *Arachis hypogaea* var. *hypogaea* (Fig. 4). Pedigree analysis indicated that the rapid expansion of GAP-type erect peanuts in China is closely related to the breeding of ‘Xuzhou 68-4’ and its derivative varieties, such as Yuhua 15, Yuhua 7, and Hua 28, which have been widely used in modern breeding programs (Fig. 5). The analysis also suggested that the GAP-type polymorphism in ‘Xuzhou 68-4’ might have originated from the landrace Shuyangdazhanyang of the subspecies *hypogaea* (Fig. 5). These findings indicate that *AhGH15* has been a crucial target of parallel selection for peanut growth habit. Different breeders have

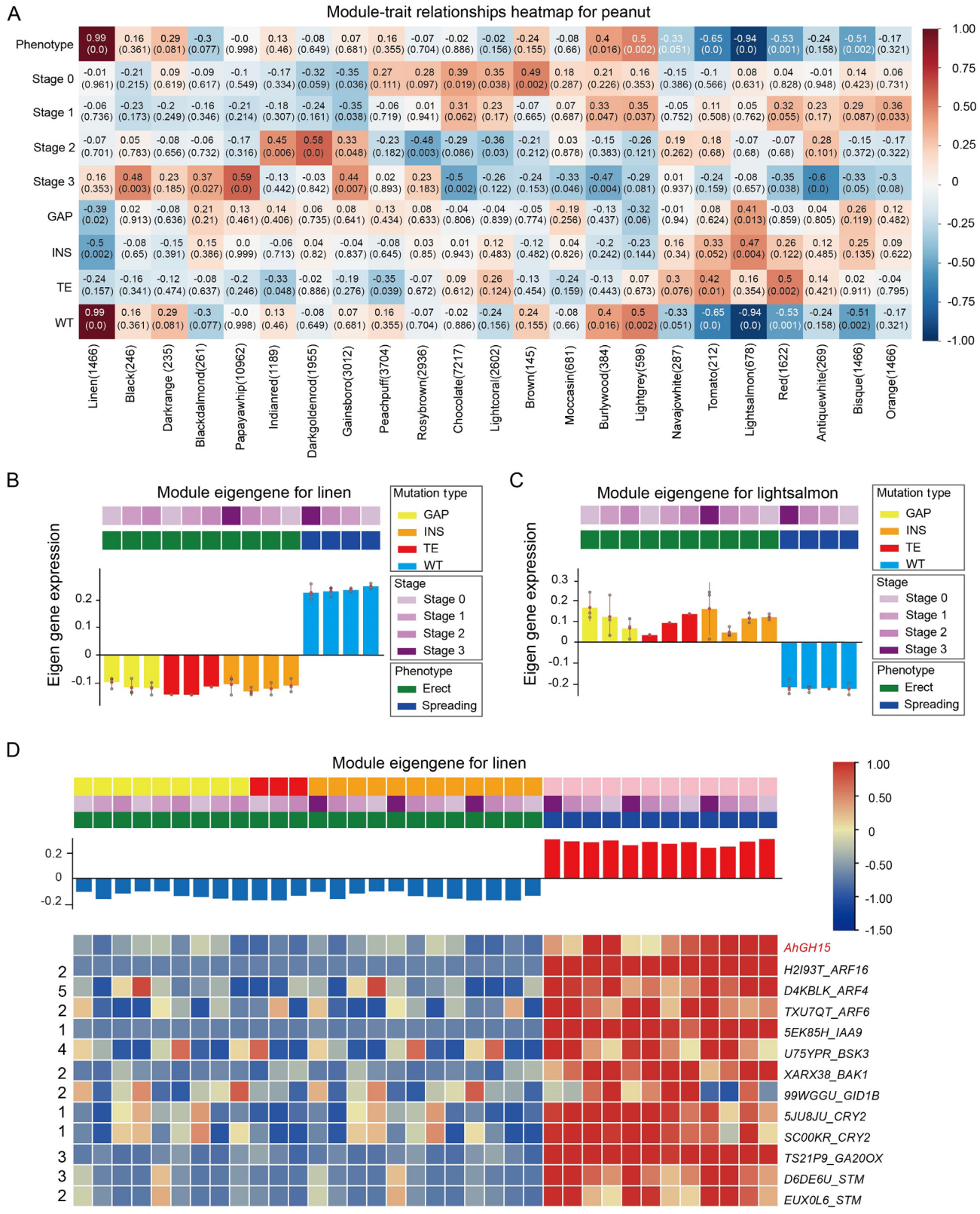

Fig. 7. WGCNA identifies gene modules that are correlated with plant growth habit. (A) Identification of the Linen and Lightsalmon modules, which comprise genes whose expression levels are correlated (Linen) or anti-correlated (Lightsalmon) with the growth habit phenotype. (B, C) Expression levels of genes in the Linen (B) and Lightsalmon (C) modules. (D) Genes co-expressed with *AhGH15*; the numbers on the left represent the number of CArG boxes in the 2-kb region upstream or downstream of the gene.

independently selected various polymorphisms within the same gene during domestication and breeding to shape plant architecture in peanut.

4.3. AhGH15 may participate in lateral branch development through a complex gene network

AhGH15 encodes a MADS-box transcription factor and is homologous to Arabidopsis *FOREVER YOUNG FLOWER* (*FYF*, also reported as *AGL42*), a member of the SOC1-like family [44–46]. Arabidopsis *SOC1* integrates developmental and plant hormonal cues and regulates the expression of genes involved in shoot development and architecture, including MADS-box target genes such as *AGL42* and flowering/meristem-associated regulator genes [44,47]. The SOC1-related pathway is also connected to GA signaling [48]. Both genetic factors and environmental inputs (e.g., gravity) are thought to regulate stem and branch angles and plant architecture [49,50]. Consistent with these observations, *AhGH15* was highly expressed in the lateral branches of spreading peanuts, supporting its potential involvement in lateral branch development and variation in growth habit.

Our transcriptome analysis further indicated that *AhGH15* expression is tightly associated with peanut growth habit, supporting its role as a key regulator of this trait. WGCNA of 36 RNA-seq samples identified two modules (Linen and Lightsalmon) that were strongly associated with the spreading or erect phenotype and with the *AhGH15* genotype (correlation coefficients of 0.99 and −0.94, respectively; Fig. 7A). *AhGH15* was assigned to the Linen module (Fig. S14; Table S16), which exhibited higher expression in the spreading type and lower expression in the erect type of peanut (Fig. 7B), whereas the Lightsalmon module showed the opposite trend (Fig. 7C). Functional annotation of genes in the Linen module that were co-expressed with *AhGH15* suggested the coordinated involvement of multiple pathways known to affect shoot growth and architecture, including BR metabolism/signaling (e.g., *BSK3*, *BAK1*, *BEN1*, *CAS7*, *EBS7*), the GA pathway (e.g., *GA20OX*, *GA2OX*, *MYB62*, *GID1B*), auxin signaling, and light or meristem-related regulators (e.g., *CRY2*, *PHYA*, *STM*) (Figs. 7D, S16). While WGCNA does not demonstrate direct regulation, these module-level associations suggest that *AhGH15* may be involved in integrating phytohormone- and environment-responsive transcriptional programs during lateral branch development.

In addition to *AhGH15* underlying *qGH15*, we also identified one candidate QTL on chromosome 6 (*qGH06*) with 99% confidence, and two more QTL on chromosomes 13 (*qGH13b*) and 15 (*qGH15b*) with 95% confidence (Fig. S2). Within the *qGH13b* region, we identified the previously reported gene *BR-SIGNALING KINASE 1* (*AhBSK1*), a core gene in BR-regulated lateral branch dispersion [26]. In addition to *qGH13*, *qGH6* exceeded the 99% confidence threshold in BSA-seq, but it was not detected by QTL mapping in the SM population. This discrepancy may have arisen because BSA-based QTL detection can be influenced by bulk construction, and its confidence level may not necessarily reflect the robustness of a locus across multiple environments. When pools are built from plants with extremes in either the erect or prostrate phenotype, variation in lateral branch dispersion may be co-selected during bulking. A similar discrepancy was observed in our late leaf spot study, where a high-confidence BSA signal was not consistently corroborated by linkage mapping [12]. These genes might help regulate growth habit, but their effects alone may be insufficient to switch the plant from a prostrate to an erect architecture. Beyond *AhGH15*, other candidate genes in the *qGH15* interval have been proposed to influence lateral branch angle in peanut [9]. Notably, these genes are located in the interval between markers M-TB34C2 and M-670ABH in our linkage map. We identified structural variations in this region in peanuts with INS-type and TE-type polymorphisms that were absent from peanuts with GAP-type polymorphisms (Fig. S16). Analysis of peanut accessions with different types of erect architecture showed that the expression of these candidate genes in this interval was not strongly correlated with the erect growth habit (Fig. S17).

In conclusion, our findings support the notion that *AhGH15* is a critical gene and a domestication hotspot controlling growth habit in peanut, and that the erect phenotype was repeatedly and independently selected by different breeders during peanut improvement. Genetic complementation studies in the future will allow further exploration of the roles of *AhGH15*.

CRediT authorship contribution statement

Guanghui Chen: Writing – original draft, Investigation, Funding acquisition, Conceptualization. **Yan Ren:** Methodology, Data curation. **Na Chen:** Methodology, Data curation. **Jiaxin Tan:** Data curation. **Heng Wang:** Resources. **Hui Wang:** Resources. **Lijun Wu:** Data curation. **Shuangling Li:** Supervision. **Tianying Yu:** Methodology. **Manish K Pandey:** Writing – review & editing. **Xiaoyuan Chi:** Writing – review & editing, Resources. **Mei Yuan:** Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We express our gratitude to Dianwen Zhang at Donggang Peanut Research Institute (Rizhao, Shandong, China) for help with field management during peanut cultivation. This research was funded by the Natural Science Foundation of Shandong Province (ZR2024MC038), the National Natural Science Foundation of China (32072107), the Agricultural Scientific and Technological Innovation Project of SAAS (CXGC2021B33, CXGC2024G20), the Taishan Scholars Program (tstp20240523), and the China Agriculture Research System of MOF and MARA (CARS-13).

Appendix A. Supplementary data

Supplementary data for this article can be found online at https://doi.org/10.1016/j.cj.2026.02.017.

References

[1] A. Krapovickas, W.C. Gregory, D.E. Williams, C.E. Simpson, Taxonomy of the genus *Arachis* (Leguminosae), Bonplandia 16 (2007) 7–205.
[2] J. Jin, W. Huang, J.P. Gao, J. Yang, M. Shi, M.Z. Zhu, D. Luo, H.X. Lin, Genetic control of rice plant architecture under domestication, Nat. Genet. 40 (2008) 1365–1369.
[3] L. Tan, X. Li, F. Liu, X. Sun, C. Li, Z. Zhu, Y. Fu, H. Cai, X. Wang, D. Xie, C. Sun, Control of a key transition from prostrate to erect growth in rice domestication, Nat. Genet. 40 (2008) 1360–1364.
[4] Y. Wu, S. Zhao, X. Li, B. Zhang, L. Jiang, Y. Tang, J. Zhao, X. Ma, H. Cai, C. Sun, L. Tan, Deletions linked to *PROG1* gene participate in plant architecture domestication in Asian and African rice, Nat. Commun. 9 (2018) 4157.
[5] G. Kayam, Y. Brand, A. Faigenboim-Doron, A. Patil, I. Hedvat, R. Hovav, Fine-mapping the branching habit trait in cultivated peanut by combining bulked segregant analysis and high-throughput sequencing, Front. Plant Sci. 8 (2017) 467.
[6] J. Pan, X. Zhou, N. Ahmad, K. Zhang, R. Tang, H. Zhao, J. Jiang, M. Tian, C. Li, A. Li, X. Zhang, L. He, J. Ma, X. Li, R. Tian, C. Ma, M.K. Pandey, R.K. Varshney, X. Wang, C. Zhao, BSA seq and genetic mapping identified candidate genes for branching habit in peanut, Theor. Appl. Genet. 135 (2022) 4457–4468.
[7] L. Li, X. Yang, S. Cui, X. Meng, G. Mu, M. Hou, M. He, H. Zhang, L. Liu, C.Y. Chen, Construction of high-density genetic map and mapping quantitative trait loci for growth habit-related traits of peanut (*Arachis hypogaea* L.), Front. Plant Sci. 10 (2019) 745.

[8] L. Li, S. Cui, P. Dang, X. Yang, X. Wei, K. Chen, L. Liu, C.Y. Chen, GWAS and bulked segregant analysis reveal the Loci controlling growth habit-related traits in cultivated peanut (*Arachis hypogaea* L.), BMC Genomics 23 (2022) 403.
[9] Y. Fang, X. Zhang, H. Liu, J. Wu, F. Qi, Z. Sun, Z. Zheng, W. Dong, B. Huang, Identification of quantitative trait loci and development of diagnostic markers for growth habit traits in peanut (*Arachis hypogaea* L.), Theor. Appl. Genet. 136 (2023) 105.
[10] Z. Zheng, Z. Sun, F. Qi, Y. Fang, K. Lin, S. Pavan, B. Huang, W. Dong, P. Du, M. Tian, L. Shi, J. Xu, S. Han, H. Liu, L. Qin, Z. Zhang, X. Dai, L. Miao, R. Zhao, J. Wang, Y. Liao, A. Li, J. Ruan, C. Delvento, R. Aiese Cigliano, C. Maliepaard, Y. Bai, R.G.F. Visser, X. Zhang, Chloroplast and whole-genome sequencing shed light on the evolutionary history and phenotypic diversification of peanuts, Nat. Genet. 56 (2024) 1975–1984.
[11] C.C. Holbrook, T.G. Isleib, D.A. Knauft, D.W. Gorbet, Registration of 'SunOleic 97R' peanut, Crop Sci. 37 (1997) 1392.
[12] G. Chen, L. Sheng, L. Wu, L. Yin, S. Li, H. Wang, X. Jiang, H. Wang, Y. Shi, F. Zhan, X. Chi, C. Qu, Y. Ren, M. Yuan, Identification of novel QTLs for resistance to late leaf spot in peanut by SNP array and QTL-seq, J. Integr. Agric. 24 (2025) 3772–3788.
[13] L. Yin, H. Zhang, Z. Tang, J. Xu, D. Yin, Z. Zhang, X. Yuan, M. Zhu, S. Zhao, X. Li, X. Liu, rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study, Genom. Proteom. Bioinform. 19 (2021) 619–628.
[14] A.L. Price, N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick, D. Reich, Principal components analysis corrects for stratification in genome-wide association studies, Nat. Genet. 38 (2006) 904–909.
[15] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114–2120.
[16] D.J. Bertioli, J. Jenkins, J. Clevenger, O. Dudchenko, D. Gao, G. Seijo, S.C.M. Leal-Bertioli, L. Ren, A.D. Farmer, M.K. Pandey, S.S. Samoluk, B. Abernathy, G. Agarwal, C. Ballen-Taborda, C. Cameron, J. Campbell, C. Chavarro, A. Chitikineni, Y. Chu, S. Dash, M. El Baidouri, B. Guo, W. Huang, K.D. Kim, W. Korani, S. Lanciano, C.G. Lui, M. Mirouze, M.C. Moretzsohn, M. Pham, J.H. Shin, K. Shirasawa, S. Sinharoy, A. Sreedasyam, N.T. Weeks, X. Zhang, Z. Zheng, Z. Sun, L. Froenicke, E.L. Aiden, R. Michelmore, R.K. Varshney, C.C. Holbrook, E.K.S. Cannon, B.E. Scheffler, J. Grimwood, P. Ozias-Akins, S.B. Cannon, S.A. Jackson, J. Schmutz, The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*, Nat. Genet. 51 (2019) 877–884.
[17] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760.
[18] D. Kim, J.M. Paggi, C. Park, C. Bennett, S.L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, Nat. Biotechnol. 37 (2019) 907–915.
[19] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R.S. Durbin, Genome project data processing, the sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[20] A. McKenna, M. Hanna, E. Banks, A.Y. Sivachenko, K. Cibulskis, A. Kernytsky, K. V. Garimella, D.D. Altshuler, S. Gabriel, M.J. Daly, M.A. DePristo, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. 20 (9) (2010) 1297–1303.
[21] B.N. Mansfeld, R. Grumet, QTLseqr: an R package for bulk segregant analysis with next-generation sequencing, Plant Genome 11 (2018) 180006.
[22] J.T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat. Biotechnol. 29 (2011) 24–26.
[23] J.T. Robinson, H. Thorvaldsdottir, A.M. Wenger, A. Zehir, J.P. Mesirov, Variant review with the integrative genomics viewer, Cancer Res. 77 (2017) e31–e34.
[24] J. Xu, X. Jiang, X. Yin, X. Zhao, N. Chen, L. Pan, C. Fu, Y. Jiao, J. Ma, M. Yuan, X. Chi, Genome-wide association analysis in peanut accessions uncovers the genetic basis regulating oil and fatty acid variation, BMC Plant Biol. 25 (2025) 651.
[25] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550.
[26] M. Pertea, G.M. Pertea, C.M. Antonescu, T.C. Chang, J.T. Mendell, S.L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads, Nat. Biotechnol. 33 (2015) 290–295.
[27] N. Rezaie, F. Reese, A. Mortazavi, PyWGCNA: a Python package for weighted gene co-expression network analysis, Bioinformatics 39 (2023) btad415.
[28] S.D. Jackman, B.P. Vandervalk, H. Mohamadi, J. Chu, S. Yeo, S.A. Hammond, G. Jahesh, H. Khan, L. Coombe, R.L. Warren, I., Birol, ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter, Genome Res. 27 (2016) 768–777.
[29] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S.K. Pham, A.D. Prjibelski, A. Pyshkin, A. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455–477.
[30] Q. Lu, L. Huang, H. Liu, V. Garg, S.S. Gangurde, H. Li, A. Chitikineni, D. Guo, M.K. Pandey, S. Li, H. Liu, R. Wang, Q. Deng, P. Du, R.K. Varshney, X. Liang, Y. Hong, X. Chen, A genomic variation map provides insights into peanut diversity in China and associations with 28 agronomic traits, Nat. Genet. 56 (2024) 530–540.
[31] B. Subramanian, S. Gao, M.J. Lercher, S. Hu, W.H. Chen, Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees, Nucleic Acids Res. 47 (2019) W270–W275.
[32] Y. Shen, Z.G. E, Y.H. Liu, Z. Chen, Database construction of Chinese peanut varieties and their genealogy, Chin. J. Oil Crop Sci. 37 (2015) 571.
[33] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.
[34] M.L. Martin, L.G.P. Jose, DNA transposons: nature and applications in genomics, Curr. Genomics 11 (2010) 115–128.
[35] J. Jiang, L. Tan, Z. Zhu, Y. Fu, F. Liu, H. Cai, C. Sun, Molecular evolution of the *TAC1* gene from rice *Oryza sativa* L, J. Genet. Genomics 39 (2012) 551–560.
[36] T. Yoshihara, M. Iino, Identification of the gravitropism-related rice gene *LAZY1* and elucidation of *LAZY1*-dependent and -independent gravity signaling pathways, Plant Cell Physiol. 48 (2007) 678–688.
[37] X. Che, B.L. Splitt, M.T. Eckholm, N.D. Miller, E.P. Spalding, BRXL4-LAZY1 interaction at the plasma membrane controls Arabidopsis branch angle and gravitropism, Plant J. 113 (2023) 211–224.
[38] Y. Hu, T. Zhao, Y. Guo, M. Wang, K. Brachhold, C. Chu, A. Hanson, S. Kumar, R. Lin, W. Long, M. Luo, J.F. Ma, Y. Miao, S. Nie, Y. Sheng, W. Shi, J. Whelan, Q. Wu, Z. Wu, W. Xie, Y. Yang, C. Zhao, L. Lei, Y. Zhu, Q. Zhang, 100 essential questions for the future of agriculture, Mod. Agric. 1 (2023) 4–12.
[39] S. Huang, P. Wei, Next-generation directed evolution in plants, Mod. Agric. 2 (2024) e70002.
[40] Y. Liu, B. Jiao, W. Qian, Synthetic gene drives in plants: development strategies, potential applications, and ethical considerations, Mod. Agric. 2 (2024) e30.
[41] M. Wang, W. Li, C. Fang, F. Xu, Y. Liu, Z. Wang, R. Yang, M. Zhang, S. Liu, S. Lu, T. Lin, J. Tang, Y. Wang, H. Wang, H. Lin, B. Zhu, M. Chen, F. Kong, B. Liu, D. Zeng, S. A. Jackson, C. Chu, Z. Tian, Parallel selection on a dormancy gene during domestication of crops from multiple families, Nat. Genet. 50 (2018) 1435–1441.
[42] Y. Peng, Y. Hu, Q. Qian, D. Ren, Progress and prospect of breeding utilization of green revolution gene *SD1* in rice, Agriculture 11 (2021) 611.
[43] S. Jiao, Q. Li, F. Zhang, Y. Tao, Y. Yu, F. Yao, Q. Li, F. Hu, L. Huang, Artificial selection of the Green Revolution gene *Semidwarf* 1 is implicated in upland rice breeding, J. Integr. Agric. 23 (2024) 769–780.
[44] C. Dorca-Fornell, V. Gregis, V. Grandi, G. Coupland, L. Colombo, M.M. Kater, The Arabidopsis SOC1-like genes *AGL42*, *AGL71* and *AGL72* promote flowering in the shoot apical and axillary meristems, Plant J. 67 (2011) 1006–1017.
[45] L. Parenicova, S. de Folter, M. Kieffer, D.S. Horner, C. Favalli, J. Busscher, H.E. Cook, R.M. Ingram, M.M. Kater, B. Davies, G.C. Angenent, L. Colombo, Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world, Plant Cell 15 (2003) 1538–1551.
[46] M.K. Chen, W.H. Hsu, P.F. Lee, M. Thiruvengadam, H.I. Chen, C.H. Yang, The MADS box gene, *FOREVER YOUNG FLOWER*, acts as a repressor controlling floral organ senescence and abscission in Arabidopsis, Plant J. 68 (2011) 168–185.
[47] J. Lee, M. Oh, H. Park, I. Lee, SOC1 translocated to the nucleus by interaction with AGL24 directly regulates *LEAFY*, Plant J. 55 (2008) 832–843.
[48] J. Moon, S.S. Suh, H. Lee, K.R. Choi, C.B. Hong, N.C. Paek, S.G. Kim, I. Lee, The *SOC1* MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*, Plant J. 35 (2003) 613–623.
[49] N. Wei, C. Tan, B. Qi, Y. Zhang, G. Xu, H. Zheng, Changes in gravitational forces induce the modification of *Arabidopsis thaliana* silique pedicel positioning, J. Exp. Bot. 61 (2010) 3875–3884.
[50] N. Yamaguchi, A. Yamaguchi, M. Abe, D. Wagner, Y. Komeda, LEAFY controls Arabidopsis pedicel length and orientation by affecting adaxial–abaxial cell fate, Plant J. 69 (2012) 844–856.