Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrating multi-platform LiDAR phenotyping and GWAS reveals PSK6, AUG8, and WDL4 as key regulators of growth in Populus deltoides

Xiaoning Ge
Institute of Forest Resource Information Techniques,Chinese Academy of Forestry
Huaiqing Zhang
zhang@ifrit.ac.cn

Institute of Forest Resource Information Techniques,Chinese Academy of Forestry
Chenggong Liu
Research Institute of Forestry, Chinese Academy of Forestry
Xiaoliang Shan
Liaoning Institute of Economic Forestry, Liaoning Academy of Agricultural Sciences
Xueyan Zhu
Institute of Forest Resource Information Techniques,Chinese Academy of Forestry
Jing Zhang
Institute of Forest Resource Information Techniques,Chinese Academy of Forestry
Yang Liu
Institute of Forest Resource Information Techniques,Chinese Academy of Forestry

Research Article

Keywords: Populus deltoides, phenotype, genome-wide association study, linkage disequilibrium, population differentiation

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11462274/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Integrating multi-platform LiDAR phenotyping and GWAS reveals *PSK6*, *AUG8*, and *WDL4* as key regulators of growth in *Populus deltoides*

Xiaoning Ge[a,b], Huaiqing Zhang[a,b], Chenggong Liu[c], Xiaoliang Shan[d], Xueyan Zhu[a,b], Jing Zhang[a,b], Yang Liu[a,b]

[a] Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing, 100091, China

[b] National Forestry and Grassland Science Data Center, NFGSDC, Chinese Academy of Forestry, Beijing, 100091, China

[c] Research Institute of Forestry, Chinese Academy of Forestry, Beijing, 100091, China

[d] Liaoning Institute of Economic Forestry, Liaoning Academy of Agricultural Sciences, Dalian 116031, China

*Correspondence:

Huaiqing Zhang, zhang@ifrit.ac.cn, Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, 100091, Beijing, China. Tel: (+86) 13651166672.

Abstract

Accelerating the intelligent breeding of Populus deltoides for high-yield timber production requires high-throughput phenotyping and a precise deconstruction of the polygenic networks governing its stem and crown architecture, which is increasingly critical given the rising global demand for wood and the need to boost timber yield per unit area. In this study, UAV and backpack lidar systems were integrated to extract individual tree height and DBH to accurately capture the phenotypic indicators of *P.deltoides*. Whole-genome resequencing and genome-wide association studies were performed on 233 germplasm resources. Alignment to the reference genome yielded 4,275,812 high-quality SNPs. Population genomic analyses revealed divergent linkage disequilibrium decay patterns between the Anhui and Hubei populations, moderate genetic differentiation, and 86 selection signatures enriched in the TFIID transcription initiation complex and phenylpropanoid metabolism. Integrating FarmCPU and BLINK models, we mapped 19 genome-wide significant SNPs and identified 14 candidate protein-coding genes. Crucially, we identified *PSK6* on *Chr4* as the primary candidate regulating tree height via cambial cell division. For diameter at breast height, a strong signal on *Chr18* implicated *AUG8*, a gene essential for microtubule organizing center assembly, in radial growth. Furthermore, *WDL4* was identified as a key regulator of crown width, corroborated by transcriptome data showing significantly repressed expression in high crown-width accessions. Collectively, these findings unravel the complex genetic architecture of longitudinal and radial growth in *P. deltoides*, providing highly specific genetic targets and a robust theoretical foundation for future genomic-assisted and intelligent breeding strategies.

Key words

Populus deltoides; phenotype; genome-wide association study; linkage disequilibrium; population differentiation

Introduction

Forest tree breeding has gradually evolved from selective breeding and crossbreeding to molecular marker-assisted breeding, achieving notable progress in

some species (Wang *et al.*, 2025). However, perennial forest trees are characterized by long breeding cycles, complex genetic backgrounds, and large generation intervals; their important traits are often controlled by multiple genes and are easily influenced by environmental conditions, leading to low efficiency and poor selection accuracy of traditional breeding methods (Xu *et al.*, 2022). To accurately quantify these complex traits, high-throughput phenomics has emerged as a crucial tool (Ninomiya *et al.*, 2019). Importantly, the dynamic expression of these phenotypes is jointly driven by underlying genetic variations and fluctuating environmental conditions. Particularly in the face of rapidly changing climate conditions and growing demand for timber, which requires higher wood yield and quality, current breeding strategies struggle to achieve coordinated improvement of target traits in a short time. In recent years, genome-wide association studies (GWAS) have been gradually introduced into forest tree breeding, offering a new technical pathway to accelerate the breeding process and enhance wood production to meet increasing timber demand (Li *et al.*, 2025).

GWAS is an analytical approach used to dissect the associations between genomic variations and biologically important traits, and is widely applied in studies of plant phenotypes (Fang *et al.*, 2026; Tibbs Cortes *et al.*, 2021). Genome-wide high-density single nucleotide polymorphism (SNP) data are obtained via large-scale sequencing, and statistical approaches are then applied to screen for genetic loci significantly associated with target traits (Demirjian *et al.*, 2022; Mahinder *et al.*, 2023). Driven by advances in sequencing technologies, GWAS has greatly improved the resolution of genetic markers in forest trees (Du *et al.*, 2016). Several studies have reported key candidate genes related to wood properties, sex determination and stress responses using GWAS (Tan *et al.*, 2022; Xue *et al.*, 2020; Wei *et al.*, 2020). Jin *et al.* (2025) analyzed the bZIP family in *Populus tomentosa* and identified genes regulating xylem development under drought stress. Ding *et al.* (2025) integrated UAV-based phenotyping to study the genetic basis of growth dynamics in *Pinus elliottii*, and GWAS explained 1.1% to 14.2% of phenotypic variation. Li *et al.* (2025) combined GWAS with environmental association analysis in a *Populus* germplasm population, revealing an allelic pattern where a key candidate gene regulates spring bud differentiation via auxin biosynthesis. Although the above progress has been made with GWAS in other *Populus* species, systematic studies in *P. deltoides* remain relatively limited.

P. deltoides is native to the central and eastern United States and is characterized by fast growth, excellent wood properties, and strong adaptability. It is one of the most widely planted fast-growing timber species in mid-latitude plains worldwide. In China, *P. deltoides* resources introduced from abroad have been mainly used in crossbreeding to obtain high-yield and high-quality hybrid populations, playing an important role in industrial timber plantations and ecological shelterbelts (Ma *et al.*, 2024). However, the imbalance between wood supply and demand remains prominent, with wood imports remaining at a high level annually (Zhou *et al.*, 2025; Dong *et al.*, 2018). As a perennial forest tree, *P. deltoides* undergoes continuous stem growth throughout its entire life cycle; tree height and diameter at breast height (DBH) increase progressively with age, a process involving cambial cell division, new cell

formation and biomass accumulation (Xiao *et al.*, 2023). Although previous genetic studies in this species have focused on abiotic stress responses and seedling-stage growth regulation, high-precision GWAS for mature growth traits such as tree height, DBH, and crown width are still lacking. Moreover, early GWAS efforts in *P. deltoides* employed targeted resequencing with limited marker density, leaving much scope for genome-wide, high-resolution dissection of growth traits.

One major bottleneck for GWAS in forest trees is the labor-intensive and low-throughput nature of traditional phenotyping, which restricts statistical power when analyzing large natural populations with complex polygenic traits (Zeng *et al.*, 2021). To overcome this limitation, high-throughput phenotyping platforms based on unmanned aerial vehicles (UAV) and light detection and ranging (LiDAR) have emerged as powerful tools to accurately and non-destructively measure structural traits such as tree height, DBH, and crown width across hundreds of individuals (Tan *et al.*, 2025). These technologies provide dense, repeatable phenotypic data and enable quantification of growth dynamics over time. When combined with GWAS, such precise phenomic information can significantly enhance the detection of SNP loci significantly associated with target traits and reduce false associations (Darmagaru *et al., 2026*). In the present study, we therefore integrated UAV- and LiDAR-based phenotyping with whole-genome resequencing to conduct a GWAS for growth traits in a *P. deltoides* germplasm collection, aiming to bridge high-throughput phenomics and genomic selection (GS) in this economically important species.

In this study, 233 accessions of *P. deltoides* germplasm were used to obtain high-density SNP markers via high-throughput whole-genome resequencing. Combined with population genetic analysis and GWAS, the study aimed to dissect key candidate genes affecting growth traits. We first assessed population structure, linkage disequilibrium (LD) characteristics, and selection signals, and then established associations between genotypes and phenotypes. We expected to identify candidate genes significantly associated with tree height, DBH, and crown width, and to preliminarily dissect their potential functions. Ultimately, 14 key candidate genes were identified. Among them, the candidate gene *PSK6* associated with tree height may be involved in growth-regulatory pathways such as phytosulfokine precursor. The gene *AUG8* associated with DBH may participate in microtubule organizing center assembly, affecting cell division and radial growth. The gene *WDL4* associated with crown width may influence lateral branch expansion through microtubule dynamics regulation. By integrating GWAS with transcriptomic validation, this study will provide an important genetic foundation and gene resources for molecular breeding and intelligent breeding of *P. deltoides*.

Results

Phenotypic extraction from knapsack LiDAR and UAV data

Backpack LiDAR can effectively acquire individual tree height and DBH information within a plot. To ensure data accuracy, we concurrently integrated UAV LiDAR to obtain point cloud data. The individual tree segmentation algorithm based on ecological metabolic theory enables extraction of DBH and tree height from

backpack LiDAR data (Fig. 1a). After removing a portion of abnormally segmented data, both DBH and tree height derived from backpack LiDAR were found to follow a normal distribution. Normality tests of phenotypic data showed that the frequency distributions of tree height and DBH were symmetric and bell-shaped, satisfying the basic assumption of residual normality for linear models (Fig. 1b, c). The DBH ranged from 11.50 to 35.70 cm, with an average of 20.67 cm. The tree height ranged from 12.90 to 28.60 m, with an average of 18.67 m. Based on the UAV data and the true value of manual measurement, the correction shows that the data is reliable (Fig. 1d, e).

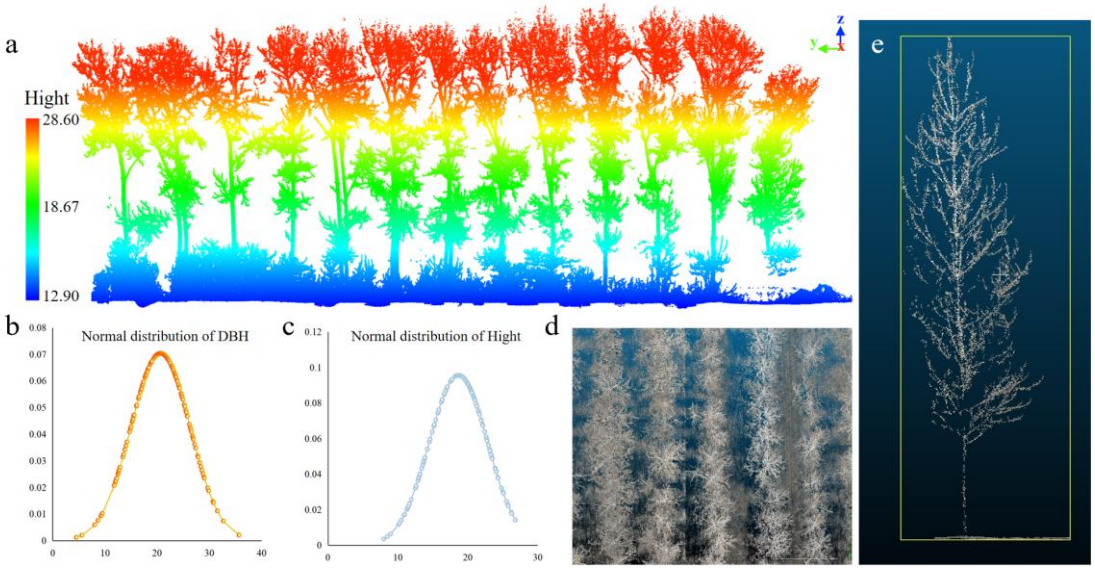


Figure 1 .Single tree segmentation and tree height and DBH extraction using backpack LiDAR and UAV data. a. Point cloud distribution results. b. Frequency distribution histogram of DBH. c. Frequency distribution histogram of tree height. d. UAV data display. e. Individual tree segmentation

Genomic variation, linkage disequilibrium and collinearity analysis

To comprehensively dissect the genomic variation landscape of *P. deltoides* germplasm resources, we first systematically analyzed the genome-wide distribution, functional annotation, and linkage disequilibrium characteristics of SNPs. Based on a total of 6,021.95 Gb of clean data from 233 *P. deltoides* germplasm accessions, we performed variant detection and obtained 4,275,812 high-quality SNPs across 19 chromosomes, CM027570.2~CM027588.2 correspond to Chr01~Chr19 respectively (Supplementary table S1). The distribution of SNPs across the 19 chromosomes was markedly heterogeneous: Chr01 harbored the largest number of SNPs, 547,901, accounting for 12.9% of the genome-wide SNPs, whereas Chr19 contained the fewest, 145,108. The Chr02~Chr18 chromosomes exhibited a relatively even distribution, ranging from 145,108 to 310,411 SNPs per chromosome, reflecting differences in gene density and recombination rates among chromosomal regions (Fig. 2a). Functional annotation of SNP (Figure 2b ; supplementary Table S2) showed that 64.79 % of SNPs were located in the intergenic region, with 2,770,392 SNPs, followed by 12.42 % in the intron region, 531,086 SNPs, 7.39 % in the upstream region, 315,939 SNPs, 6.55 % in the downstream region, 280,256 SNPs, 4.82 % in

the exon region, 205,912 SNPs, 2.29 % in the 3'UTR, 97,779 SNPs, 1.14 % in the 5'UTR, 48,779 SNPs and 0.02 % splicing sites, and 858 SNPs. The proportion of SNPs in exons was only 4.82 %, which was much lower than that in intergenic region (64.79 %) and intron region (12.42 %), reflecting the restriction of evolutionary negative selection on the variation of protein coding region.

With respect to mutation types (Fig. 2c; Supplementary table S3), C/T transitions were the most abundant, with 846,705 occurrences accounting for 19.80%, followed by G/A transitions with 845,725 occurrences accounting for 19.78%. The genome-wide transition-to-transversion ratio (Ts/Tv) was 1.91 (Ts = 2,808,781; Tv = 1,467,031), close to the theoretical equilibrium value of 2.0, confirming the high quality of the data and a low false-positive rate. LD analysis (Fig. 2d) showed that in the Anhui population, the initial r^2 value was 0.4006 and decayed rapidly to the background level of 0.0852 within approximately 2 kb, indicating a high historical recombination rate at the genome-wide level in this population. In the Hubei population, the initial r^2 value was 0.5229 and decayed slowly to the background level of 0.1713 within approximately 30 kb; the background LD was significantly higher than that in the Anhui population, suggesting that the Hubei population may have experienced more recent genetic drift or a bottleneck event. The marked difference in LD decay patterns between the two populations points to divergence in their mating systems and population histories.

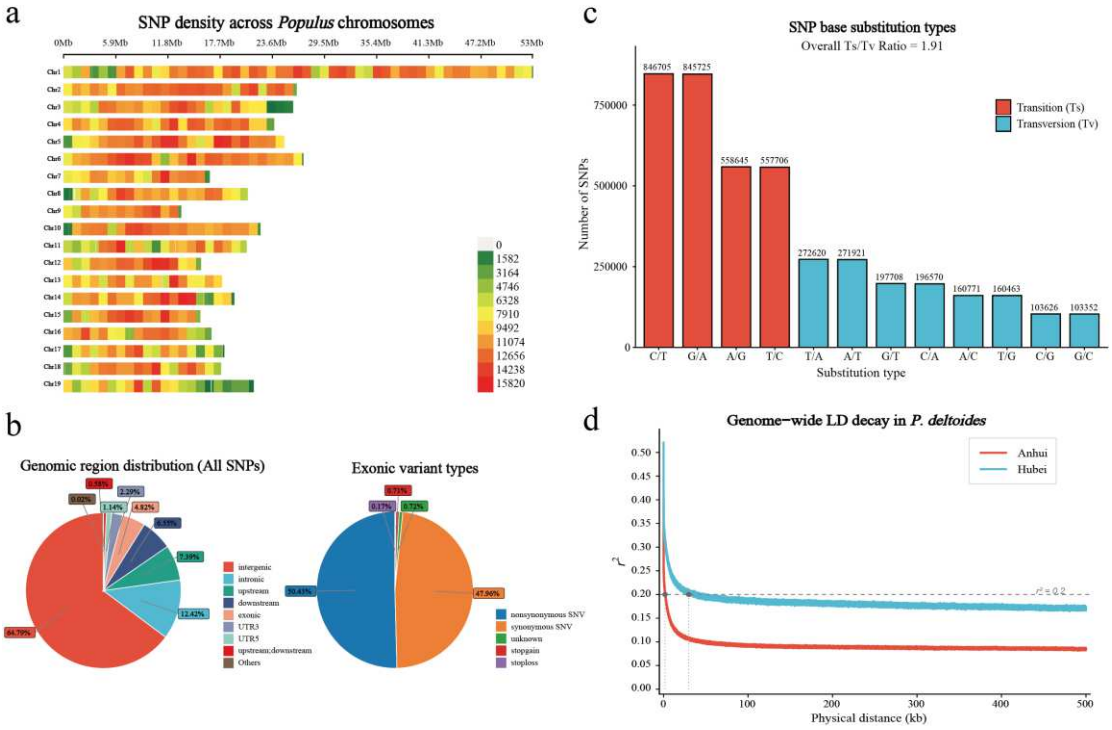


Figure 2 Genetic variation and linkage disequilibrium analysis. a. SNP density distribution across the 19 chromosomes of *P. deltoides*. The x-axis indicates the chromosome number (Chr01–Chr19), and the y-axis indicates the number of SNPs. b. Genomic distribution of SNPs across different genic regions. c. Transition/transversion (Ts/Tv) distribution. d. Linkage disequilibrium (LD) decay curves for the Anhui and Hubei populations (threshold $r^2 = 0.20$).

Population structure and kinship analysis of *P. deltoides* germplasm

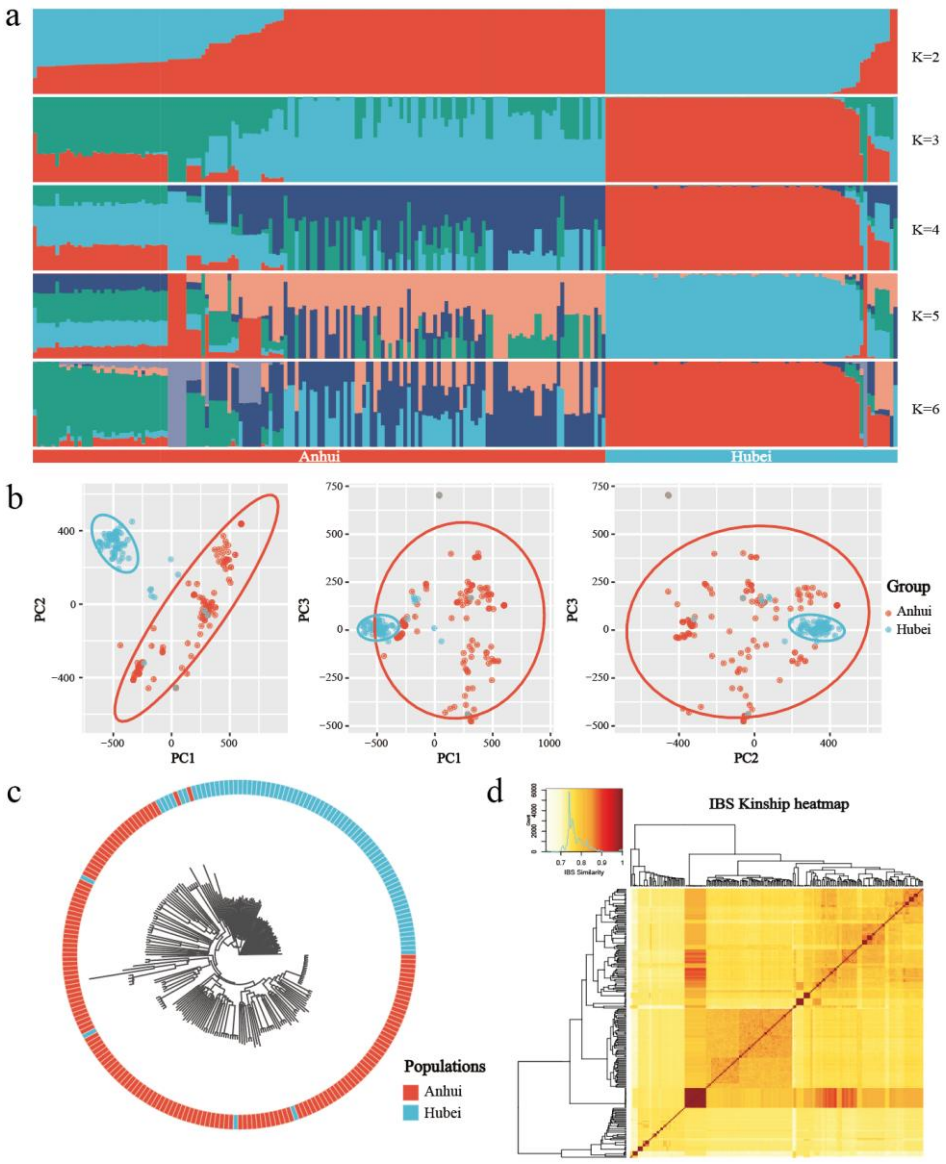


Figure 3 Population genetic structure analysis. a. Admixture population structure analysis (K = 2). Each vertical bar represents an individual sample, and the two colors denote the two genetic groups (Anhui population and Hubei population). b. PCA scatter plot. c. Neighbor-joining phylogenetic tree based on genetic distances among samples. d. IBS analysis of *P. deltoides* germplasm.

Based on the above variation characteristics, we further evaluated the population structure and genetic relatedness of *P. deltoides* germplasm to reveal the genetic differentiation patterns among samples. Admixture population structure analysis divided the 233 *P. deltoides* accessions into K = 1 to 10 genetic clusters. The cross-validation error was minimized at K = 2, indicating the presence of two major genetic groups within this natural population (Fig. 3a). The genetic compositions of the two groups were clearly separated in the plot, suggesting a pronounced genetic divergence between the Anhui and Hubei accessions. Principal component analysis (PCA) (Fig. 3b) showed that the contribution rates of PC1, PC2, and PC3 were 10.66%, 6.83%, and 5.72%, respectively. The samples clustered distinctly into two groups along the PC1 and PC2 axes, consistent with the genetic grouping, indicating

that inter-regional genetic differentiation is the primary source of variation among the samples. Phylogenetic tree analysis and identity-by-state (IBS) analysis (Fig. 3c, d) further validated these results: all germplasm accessions were divided into two major clades on the phylogenetic tree, which highly matched the grouping from Admixture at K = 2, providing additional support for the genetic differentiation pattern between the Anhui and Hubei groups.

Asymmetric selection signatures between the Anhui and Hubei populations

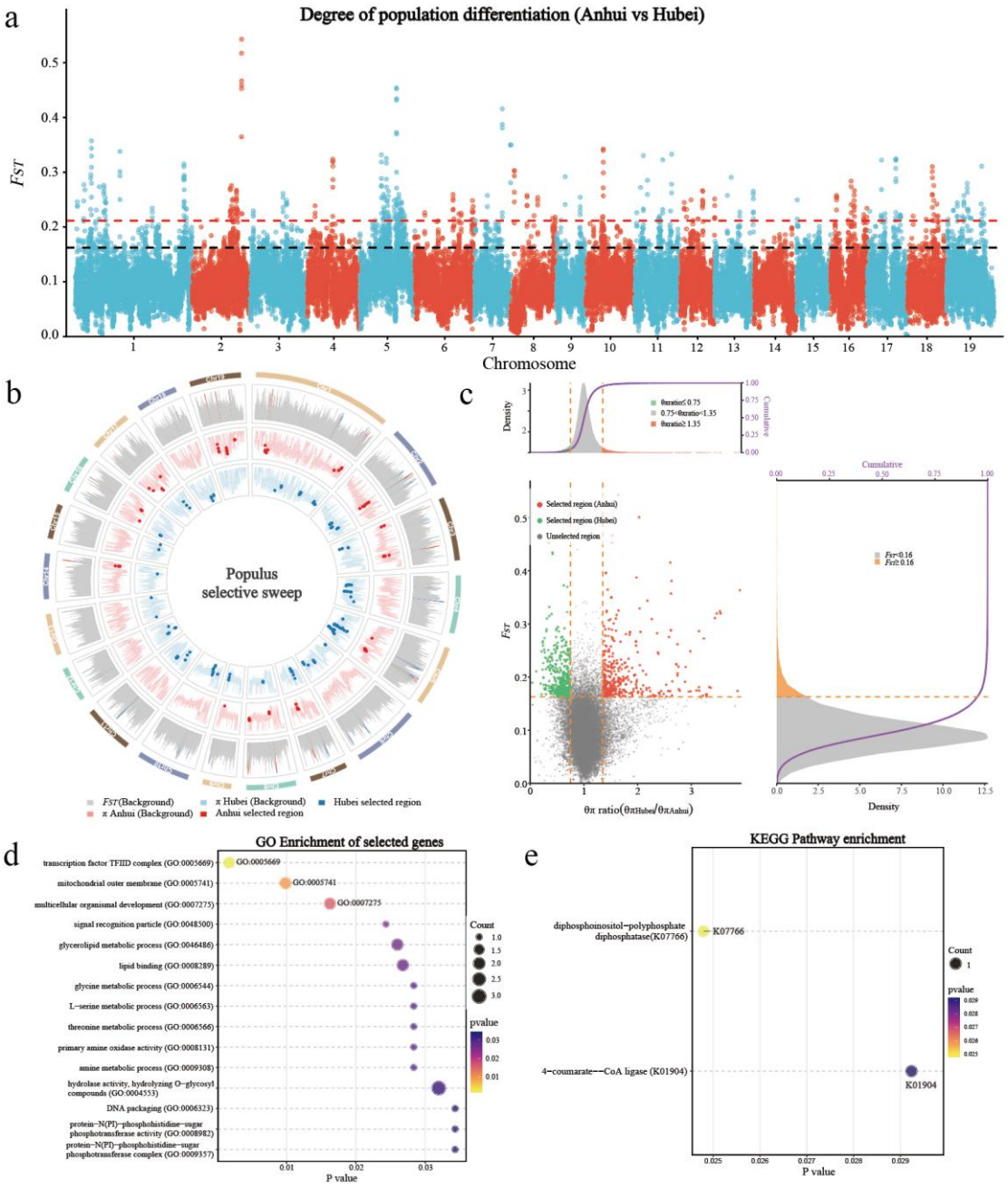


Figure 4 Genetic differentiation and selection signature analysis between the Anhui and Hubei populations. a. Genome-wide F_{st} distribution plot (10-kb sliding window). b. Bar chart comparing Tajima's D distributions. c. Distribution plot of candidate regions under selection. d. Bar chart of GO enrichment analysis. e. Bubble plot of KEGG pathway enrichment analysis.

Given the significant genetic differentiation between the two populations, we

performed selection signature analysis to investigate whether they have undergone differential adaptive selection during evolution. Using a 10-kb sliding window to calculate F_{st} values, a total of 42,118 windows were obtained across the genome, with an average F_{st} of 0.0902. By taking the top 5% of π windows, 86 candidate regions under selection were identified (Supplementary table S3). The distribution of F_{st} showed a heterogeneous pattern (Fig. 4a), clear F_{st} peaks were observed near Chr01 at approximately 7.2 Mb with an F_{st} value of 0.3089, near Chr02 at approximately 17.3 Mb with an F_{st} value of 0.2287, near Chr07 at approximately 7.3 Mb with an F_{st} value of 0.2577, near Chr10 at approximately 10.6 Mb with an F_{st} value of 0.2582, and near Chr19 at approximately 4.8 Mb with an F_{st} value of 0.2442, suggesting that these genomic regions have undergone divergent selection between the Anhui and Hubei groups.

The Tajima's D test results (Fig. 4b; Supplementary table S4) showed that Tajima's D was 2.2857 for the Anhui group and 1.7890 for the Hubei group; both values were positive, implying recent population expansion or balancing selection. The higher Tajima's D in the Anhui group suggests that it may have experienced stronger population expansion. The candidate regions under selection exhibited an asymmetric distribution (Fig. 4c). GO enrichment analysis (Fig. 4d; Supplementary table S5) revealed that genes within the candidate intervals were significantly enriched in functional categories such as TFIID transcription initiation complex (GO:0005669), mitochondrial outer membrane (GO:0005741), and multicellular organism development (GO:0007275). KEGG pathway analysis (Fig. 4e; Supplementary table S6) further indicated that the candidate genes were enriched in the phenylpropanoid metabolic pathway (4-coumarate-CoA ligase, K01904) and the phosphatidylinositol signaling pathway (diphosphoinositol-related enzyme, K07766), suggesting that the two groups may respond to their respective environmental pressures through distinct molecular pathways.

Fourteen candidate genes were identified by GWAS

To directly dissect the genetic loci controlling growth traits, we performed GWAS for tree height, DBH, and crown width using high-density SNP markers. Based on 4,241,298 high-quality SNPs, GWAS for the three growth traits was conducted using both the FarmCPU and BLINK models.

The Manhattan plot analysis (Fig. 5a) revealed that for tree height, the most significant genome-wide signal was detected on Chr4 (SCM027573.2_4006173, $P = 2.89\times10^{-18}$), with an association strength much higher than signals for the other traits on any chromosome, representing the most significant QTL identified in this study. The second and third strongest signals for tree height were detected on Chr14 (SCM027583.2_192380, $P = 1.28\times10^{-13}$) and Chr10 (SCM027579.2_1721767, $P = 6.12\times10^{-12}$), respectively. For DBH, a significant signal was identified on Chr18 (SCM027587.2_4278044, $P = 1.83\times10^{-12}$), and a weaker secondary signal was observed on Chr15 (SCM027584.2_14132933, $P = 2.80\times10^{-8}$). For crown, significant loci were detected on multiple chromosomes, including Chr1 ($P = 2.86\times10^{-9}$), Chr2 ($P = 2.08\times10^{-9}$), and Chr16 ($P = 3.10\times10^{-9}$), exhibiting a multi-peak distribution pattern consistent with the polygenic regulation of complex traits. The results from the

FarmCPU and BLINK models corroborated each other; for example, the Chr16 signal for crown width was detected by both BLINK ($P = 3.10\times10^{-9}$) and FarmCPU ($P = 7.69\times10^{-8}$), enhancing the reliability of the findings. The Q-Q plot (Fig. 5b) showed that the low P-value region deviated markedly upward from the diagonal, indicating genuine association signals rather than genome-wide inflation; the majority of SNPs followed the diagonal line, confirming effective model correction. The LD block plots (Fig. 5c, d; Supplementary table S7) clearly illustrated the linkage disequilibrium structure surrounding the candidate genes on Chr4 and Chr18. Within the 30-kb LD intervals, a total of 14 candidate protein-coding genes were screened (Table 1; Supplementary table S8).

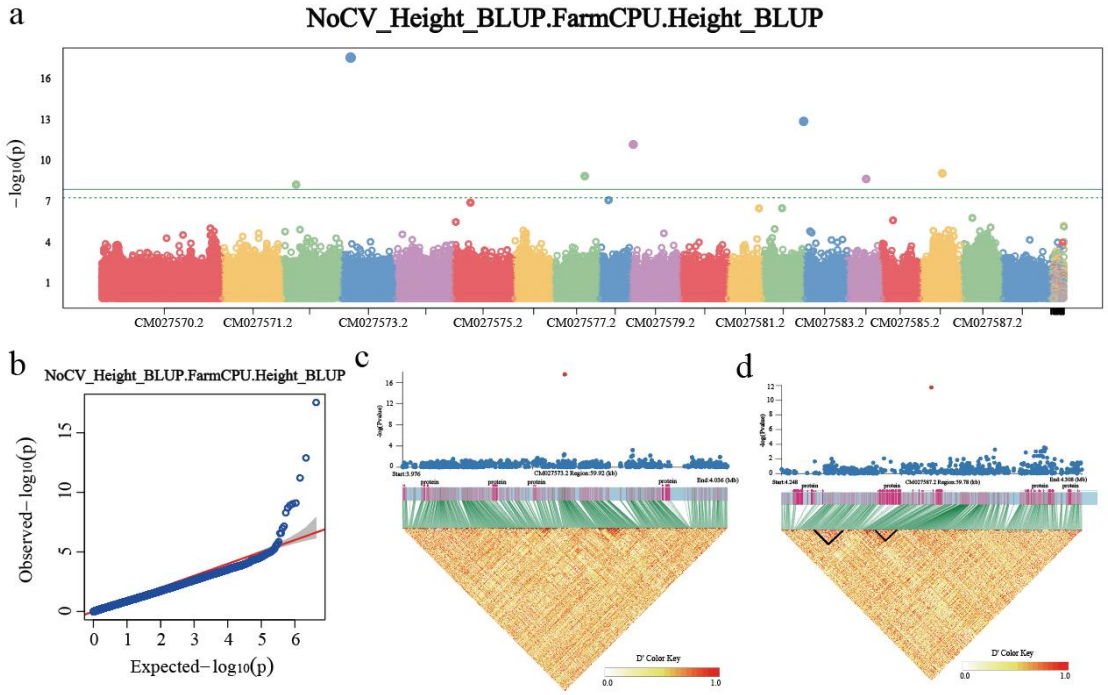


Figure 5 Genome-wide association study. a. GWAS Manhattan plots for the three growth traits (the red dashed line indicates the significance threshold of $P = 1\times10^{-5}$). b. Q-Q plot. c. LD block plot of candidate genes on Chr4 (surrounding the tree height signal SCM027573.2_4006173). The color gradient denotes r^2 values, illustrating the linkage pattern of SNPs in the Chr4 interval; the red dashed box indicates the location of the candidate gene *PSK6*. d. LD block plot of candidate genes on Chr18 (surrounding the DBH signal SCM027587.2_4278044).

Table 1 GWAS significant SNPs and candidate genes

Trait	Gene ID	Gene name	Significant SNP	P-value	Gene function
Crown	H0E87_027040	WDL4	SCM027585.2_11233729	3.10E-09	WVD2-like 4 protein
Crown	H0E87_031526	ATL80	SJACEGQ020000100.1_11615	7.76E-09	RING-H2 finger E3 ubiquitin ligase
Crown	H0E87_002485	LOL1	SCM027570.2_40651197	2.86E-09	LSD family zinc finger protein

Crown	H0E87_005777	COL9	SCM027571.2_20431250	6.96E-08	CONSTANS-LIKE 9 transcription factor
Crown	H0E87_027039	CPK	SCM027585.2_11233729	3.10E-09	Calcium-dependent protein kinase
Crown	H0E87_004971	PUB33	SCM027571.2_12539643	2.08E-09	U-box E3 ubiquitin ligase
Crown	H0E87_010468	EH3	SCM027574.2_9799698	5.55E-08	Epoxide hydrolase 3
Height	H0E87_024461	KLCR1	SCM027583.2_7330853	2.50E-10	Kinesin light chain-associated protein 1
Height	H0E87_008060	PSK6	SCM027573.2_4006173	2.89E-18	Phytosulfokine 6 precursor
Height	H0E87_008059	ADT5	SCM027573.2_4006173	2.89E-18	Prephenate dehydratase
Height	H0E87_023752	UPF0481	SCM027583.2_192380	1.28E-13	UPF0481 protein At3g47200
Height	H0E87_018257	KLCR1	SCM027579.2_1721767	6.12E-12	Kinesin light chain-associated protein 1
DBH	H0E87_016058	BGLU34	SCM027577.2_14460943	1.24E-09	Beta-glucosidase 34
DBH	H0E87_006410	bZIP-TF	SCM027572.2_6016420	5.13E-09	bZIP transcription factor
Crown	H0E87_025951	MED17	SCM027584.2_8336767	2.01E-09	RNA polymerase II
Crown	H0E87_029376	AUG8	SCM027587.2_4278044	1.83E-12	AUGMIN subunit 8
Crown	H0E87_026243	WAK-like	SCM027584.2_14132933	2.80E-08	Cell wall-associated receptor kinase

Transcriptome expression validation of candidate genes

To validate the potential functions of the above candidate genes in phenotypic differences, we selected high- and low-phenotype groups for transcriptome comparison, focusing on genes with significant expression differences. The samples were divided into high-value and low-value groups, and transcriptome sequencing was performed to compare the expression levels of the candidate genes. Among the 14 candidate genes, *WDL4* was the only one showing a statistically significant expression difference between the two groups (Fig. 6a; Supplementary table S9). In the high crown-width group, the expression level of *WDL4* was only 19% of that in the low crown-width group , indicating a significant negative correlation between *WDL4* expression level and crown-width trait (Fig. 6b, c). *WDL4* belongs to the *WVD2*-like protein family, which has been reported in *Arabidopsis thaliana* to be involved in microtubule organization and cell morphogenesis regulation. The association between downregulation of *WDL4* and increased branch elongation suggests that *WDL4* may

act as a growth suppressor in the regulatory network of crown width formation. Upon inhibition of *WDL4* expression, the altered microtubule organization state may promote cell division and elongation in lateral meristems, ultimately resulting in a larger crown width.

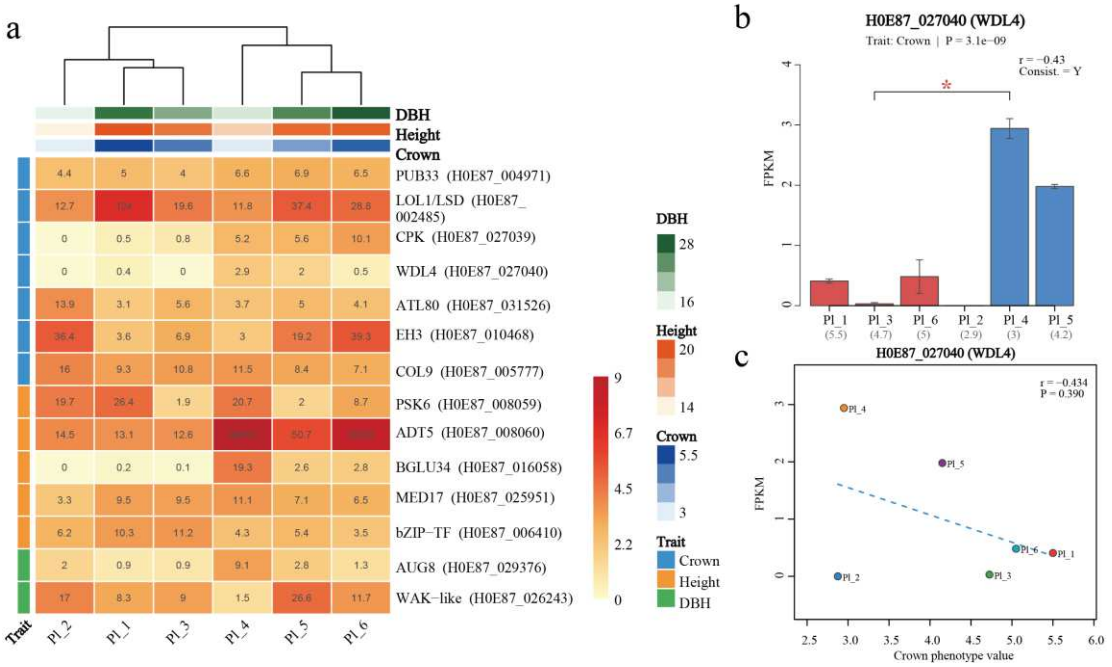


Figure 6 Transcriptome differential expression validation of the *WDL4* gene. a. FPKM expression differences of GWAS candidate genes between different phenotypic groups. b. FPKM expression differences of the candidate gene *WDL4* between high- and low-phenotype groups, with Welch's t-test statistical analysis. Red represents individuals in the large crown width group, and blue represents individuals in the small crown width group. c. Pearson correlation coefficient evaluating the linear relationship between *WDL4* expression level and the target trait. Bars represent relative expression ($2^{-\Delta\Delta Ct}$ method), error bars indicate standard error of the mean (SEM), * indicates P<0.05.

Discussion

High-quality variant data reveal inter-population LD differences

Through whole-genome resequencing of *P. deltoides* populations, a total of 4,275,812 high-quality SNPs were identified. The Ts/Tv was 1.91, which falls within the reasonable range for genomic variation in poplar and is consistent with the pattern of cytosine deamination commonly observed in forest tree genomes, thereby strongly supporting the accuracy of SNP calling and the reliability of the variant dataset (Li *et al.*, 2023). Regarding the genomic distribution of variant sites, exonic SNPs accounted for only 4.82% of the total, whereas as much as 64.79% of the variants were located in intergenic regions. This distribution pattern reflects the strong purifying selection that the poplar genome has experienced during evolution. Mutations occurring in coding regions often lead to lethality or significantly reduce fitness and are therefore gradually eliminated during evolution; variants in exonic regions are more likely to affect protein structure and function and are more stringently removed or maintained at low frequencies by natural selection (Slavov *et*

al., 2025). In contrast, intergenic and intronic regions experience weaker selection pressure and can accumulate more neutral or nearly neutral variants. Among the coding-region variants, non-synonymous mutations accounted for 50.43%. This proportion indicates that nearly half of the coding SNPs cause amino acid sequence changes and may affect protein function. Consequently, these non-synonymous sites become priority candidate targets for subsequent functional validation and molecular breeding. However, it should also be noted that some non-synonymous mutations may be subject to weak negative selection, and their functional importance needs further evaluation in combination with evolutionary conservation analysis.

LD analysis revealed significant differences in the genetic structure between the Anhui and Hubei populations. The Anhui population exhibited rapid LD decay, dropping from an initial r^2 of 0.4006 to the background level, 0.0852 within approximately 2 kb. In contrast, the Hubei population showed much slower LD decay, decreasing from 0.5229 to 0.1713 over about 30 kb, with a background LD level twice that of the Anhui population. This difference in LD decay distance reflects distinct genetic backgrounds and evolutionary histories of the two populations; LD analysis of natural *P. deltoides* populations also indicates low to moderate LD levels (Fahrenkrog *et al.*, 2017). The extremely short LD decay distance in the Anhui population suggests a larger effective population size and more frequent genetic recombination, which is a significant advantage for GWAS, enabling more precise candidate gene mapping. In contrast, the longer LD interval in the Hubei population may be related to a founder effect, local adaptive selection, or a smaller effective population size experienced by this population. The marked divergence in LD decay patterns between the two populations suggests that when performing multi-modal data integration and genomic prediction for selection, population-specific linkage structures must be fully considered: for populations with rapid LD decay, higher-density marker coverage is required; for populations with slow LD decay, attention should be paid to the potential impact of long-range linkage on estimation bias of effects (Slavov *et al.*, 2025).

Molecular evidence supports inter-population genetic differentiation

This study employed population genetic analyses to thoroughly reveal the genetic evolutionary characteristics of *P. deltoides* in China. Population structure analysis showed significant differentiation at K = 2, with a F_{st} of 0.0902 between the populations. This moderate level of differentiation is highly consistent with the introduction route of *Section Tacamahaca* poplars into China and the subsequent long-term regional adaptive selection (Shi *et al.*, 2025). Anhui and Hubei belong to different climatic zones, and the long-standing environmental heterogeneity may have driven divergent selection at key adaptive loci between the two populations. Genome-wide F_{st} scanning identified a total of 86 candidate regions under selection, which were asymmetrically distributed across chromosomes. Prominent F_{st} peaks were detected on chromosomes Chr01, Chr02, Chr07, Chr10, and Chr19, suggesting that these regions harbor key selective loci underlying population differentiation, and the genes enriched therein may be involved in critical biological processes of local adaptation (Geraldes *et al.*, 2014). Tajima's D values were positive for both

populations (2.2857 for Anhui > 1.7890 for Hubei). This molecular signature suggests the possible action of balancing selection rather than typical population expansion signals. Considering the introduction history, it is speculated that after being introduced into the Huang-Huai-Hai and Yangtze River basins, *P. deltoides* may have undergone rapid population establishment and local adaptation under the dual pressures of artificial selection and environmental adaptation.

GO enrichment analysis revealed that genes within the selected regions were significantly enriched in the TFIID transcription initiation complex (GO:0005669). The TFIID complex is a core component of transcription initiation by RNA polymerase II; its enrichment suggests that the two populations have undergone adaptive differentiation at the level of global transcriptional regulatory networks, potentially coordinating responses to regional environmental pressures by modulating the expression of multiple downstream target genes. KEGG pathway analysis further showed that the selected genes were significantly enriched in the phenylpropanoid metabolic pathway, which corresponds to ko00940, with 4-coumarate-CoA ligase as the key enzyme, entry K01904. Phenylpropanoid metabolism is a major pathway of plant secondary metabolism, producing lignin, flavonoids, phenolic acids, and other compounds that are closely related to cell wall reinforcement, stress resistance, and signal transduction (Dong *et al.*, 2021). The enrichment of this pathway indicates that the Anhui and Hubei populations may have undergone adaptive differentiation in traits such as lignification degree, pest and disease resistance, or ultraviolet radiation tolerance in response to the different ecological conditions of the two regions. In summary, population genetic differentiation is not only reflected in overall allele frequency differences but has also left selection imprints in key functional pathways such as transcriptional regulation and secondary metabolism, providing molecular evidence for understanding the regional adaptive evolution of *P. deltoides*.

GWAS dissects the polygenic regulatory architecture of growth traits

In this study, we performed GWAS for three growth traits using both the FarmCPU and BLINK models. A total of 19 genome-wide significant SNPs were detected, and 14 candidate protein-coding genes were identified. Distinct traits exhibited markedly different patterns of association signals, reflecting the diversity of their genetic architectures. An exceptionally significant signal detected on Chr4, with a P-value of 2.89×10^{-18} and a $-\log_{10}(P)$ of 17.54, is extremely rare in comparable studies. The candidate gene is *PSK6*, encoding a phytosulfokine precursor. Phytosulfokines (*PSKs*) are sulfated pentapeptide growth factors that promote cell proliferation and elongation through receptor-mediated signaling pathways (Zhang *et al.*, 2025). Previous studies have shown that *PSK6* is highly expressed in the cambial zone of poplar and directly affects stem radial growth and tree height development by regulating cambial cell division activity (Zhang *et al.*, 2024). The highly significant association on Chr4 in this study indicates that *PSK6* is likely a key core gene controlling tree height in *P. deltoides*. In addition to *PSK6*, secondary signals on Chr14 (UPF0481) and Chr10 (*KLCR1*) also contribute to tree height variation. *KLCR1*, which is kinesin light chain-related protein 1, participates in microtubule movement. For DBH, a significant signal was identified on Chr18 with a P-value of 1.83×10^{-12},

and the candidate gene is *AUG8*. As an essential subunit of the *AUGMIN* complex, *AUG8* is a microtubule plus-end binding protein that regulates microtubule dynamics and promotes microtubule reorientation (Farquharson *et al.*, 2013). In plant cambial cells, the expression level of *AUG8* is closely related to the plane and frequency of cell division. The significant association of *AUG8* with DBH reveals, at the molecular level, a direct link between radial growth and the regulation of cambial cell division, namely that *AUG8* may influence secondary xylem yield by affecting microtubule organization during cambial cell mitosis, ultimately affecting DBH.

In contrast, SNPs significantly associated with crown width were distributed across multiple chromosomes, including Chr1, Chr2, and Chr16, exhibiting a typical multi-peak distribution pattern with relatively similar significance levels ($P \approx 10^{-8}$~10^{-9}). The identified candidate genes involve diverse functions, including microtubule organization (*WDL4*), calcium-dependent protein kinase (*CPK1*), photoperiod-responsive transcription factor (*COL9*), and E3 ubiquitin ligase (*PUB33*), indicating that crown width is controlled by multiple minor genes with similar effect sizes, without a single major-effect locus. Crown growth and development are coordinately regulated by multiple pathways including cytoskeletal dynamics, calcium signaling, photoperiod response, and protein degradation. Notably, both the tree height major-effect candidate gene *PSK6* and the crown width candidate gene *WDL4* are closely related to the regulation of microtubule dynamics: *PSK6* promotes cell division through hormone signaling (Wei *et al.*, 2025), while *WDL4* regulates the direction of cell expansion by influencing microtubule organization (Schaefer *et al.*, 2023). This finding suggests that different growth traits may share some cytological bases but achieve trait-specific regulation through distinct molecular mechanisms. Compared with GWAS studies in other tree species such as *Populus tomentosa* and slash pine, the association strength of the *PSK6* locus identified in this study far exceeds previously reported levels, highlighting the central role of this gene in tree height regulation in poplar (Wang *et al.*, 2023). The polygenic architecture of crown width is consistent with findings from studies of crown traits in other forest trees, further validating the complex genetic basis of crown width (Guo *et al.*, 2025).

Molecular mechanism and functional validation of *WDL4* negatively regulating crown width

Crown width is a key architectural trait determining stand productivity and photosynthetic efficiency, yet research on its molecular mechanisms has lagged. In this study, GWAS using both the FarmCPU and BLINK models identified the candidate gene *WDL4* on Chr16. Among the 14 candidate genes, *WDL4* was the only one validated by both RNA-seq and qPCR. Experimental data showed that the expression level of *WDL4* in the high crown-width group was only 19% of that in the low crown-width group, confirming a significant negative correlation between *WDL4* expression and crown width. *WDL4* belongs to the plant-specific *WVD2*-like microtubule-associated protein family. Functional studies in *Arabidopsis thaliana* have shown that *WVD2*-like proteins regulate the direction and rate of cell elongation by binding to and promoting microtubule bundling, thereby influencing organ morphogenesis (Perrin *et al.*, 2007). It has been reported that overexpression of

At*WDL3* leads to shortened cells and dwarfed plants, whereas wdl mutants exhibit excessive cell elongation (Lian *et al.*, 2017). Based on these conserved functions, we propose a hypothetical molecular mechanism by which *WDL4* negatively regulates crown width: in low crown-width plants, *WDL4* is normally or highly expressed, promoting microtubule bundling and transverse alignment, which restricts the elongating division of lateral meristem cells and thus inhibits lateral branch expansion; in high crown-width plants, *WDL4* expression is significantly downregulated, leading to reduced microtubule organization, allowing lateral meristem cells greater freedom for division and elongation, ultimately resulting in increased crown width.

In addition to *WDL4*, we also detected an association signal for the calcium-dependent protein kinase *CPK1* in the Chr16 region harboring crown-width-associated signals. Calcium signaling is a central regulatory hub for plant growth, development, and environmental responses (Luan *et al.*, 2025). *CPK* family members can directly phosphorylate microtubule-associated proteins to regulate microtubule dynamics (Lima-Huanca *et al.*, 2025). Based on these findings, we propose an integrated calcium signaling-*WDL4*-microtubule organization regulatory network hypothesis for crown width development: environmental changes cause fluctuations in intracellular calcium concentration, activating *CPK1*; *CPK1* in turn modulates the activity or stability of *WDL4* through phosphorylation, thereby altering the microtubule organization state and ultimately affecting the direction and rate of lateral bud cell expansion, determining crown width. This hypothesis links calcium signal transduction, microtubule dynamics, and morphogenesis, providing a clear direction for subsequent validation. Regarding the reliability of the association signals, both FarmCPU and BLINK independently detected the crown-width signal on Chr16, and the mutual confirmation between the two models significantly enhances the credibility of this QTL while effectively controlling the risk of false positives that may arise from a single model. Furthermore, dual validation by RNA-seq and qPCR further confirms, at the expression level, the negative regulatory relationship between *WDL4* and crown width, providing a solid experimental foundation for subsequent molecular mechanism studies.

Limitations and future directions

Although this study identified several candidate genes significantly associated with growth traits through GWAS and validated the expression difference of *WDL4*, several limitations remain that warrant future investigation. The first limitation concerns sample size. While the 233 accessions used in this study meet the sample size requirement for moderate-scale GWAS, they are still insufficient for detecting minor-effect genes and rare variants. A large number of low-frequency variants (MAF between 0.01 and 0.05) in natural poplar populations may contribute to part of the phenotypic variation, but the current sample size makes it difficult to reliably detect them. Future GWAS should be conducted using larger germplasm collections combined with whole-genome sequencing or high-density SNP arrays to increase the statistical power of association tests and to uncover additional minor-effect loci. The second limitation is insufficient functional validation. At present, only transcriptome-level expression validation has been performed for *WDL4*, and direct

molecular genetic evidence is lacking. Follow-up studies should conduct transgenic functional validation of core candidate genes such as *PSK6*, *AUG8*, and *WDL4*. The third limitation concerns the application of GS. Traditional forest tree breeding has long cycles and high costs. GS technology, which predicts breeding values based on genome-wide markers, can significantly shorten the breeding generation interval. The prior-information genes identified in this study, including *PSK6*, *AUG8*, and *WDL4*, can serve as fixed-effect covariates in GS models to improve prediction accuracy. Future work can develop a low-density chip containing markers for these candidate genes in *P. deltoides* breeding populations, combined with annual phenotypic measurements, to validate their actual benefits in GS, thereby advancing molecular marker-assisted breeding toward intelligent and precision breeding.

Materials and methods

Sampling and phenotyping

A total of 233 accessions of *P. deltoides* germplasm were used as experimental materials. The plants were established in 2017 at the National Poplar Improved Variety Base in Shishou City, Hubei Province (29°01′53″~33°06′47″N, 108°21′42″~116°07′50″E) and at the Shahe Forest Tree Improved Variety Breeding Center of the Anhui Academy of Forestry (29°41′~34°38′N, 114°54′~119°37″E). Phenotypic traits including tree height, DBH, and crown width were acquired using a backpack LiDAR system, with manual measurements used as calibration data. Young leaves were randomly collected from each sample, immediately frozen in liquid nitrogen, and stored at −80°C for whole-genome resequencing, RNA-seq, and qPCR.

Backpack LiDAR data was acquired in December 2024 using a LiBackpack H120 scanner from Beijing Digital Green Earth Co., Ltd. This device integrates two LiDAR sensors in the horizontal and vertical directions, and is equipped with a high-precision satellite navigation system to acquire high-precision 3D point cloud data within the scanned area. Sample plot data was obtained within the study area using the S-shaped walking path recommended by the LiBackpack H120. After data acquisition, LiDAR 360 V7.0 software was used to preprocess all backpack LiDAR point cloud data, including denoising, filtering, DEM generation, point cloud normalization, and tree segmentation. UAV image data was simultaneously acquired in December 2024 using a DJI M350, with a flight altitude of 50m, a flight speed of 6m/s, and a lateral overlap rate of 50%. After data acquisition, LiDAR 360 V7.0 software was used for processing, and tree segmentation was performed based on the watershed algorithm to extract tree height and DBH information.

DNA isolation and genome resequencing

Leaf samples from each individual were used for DNA extraction. Genomic DNA was extracted using a standard phenol/chloroform method. DNA integrity was assessed by 1% agarose gel electrophoresis, while concentration and purity were measured using a NanoPhotometer and Qubit 2.0, respectively. After library construction, whole-genome resequencing was performed on an Illumina platform with paired-end 150 bp (PE150) reads. The *P. deltoides* reference genome (GCA_052148935.1,https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_05214893

5.1/) was used as the alignment reference. The raw image data obtained from high-throughput sequencing were converted into raw reads through base calling and stored in FASTQ format.

Data quality control and variant detection

Raw data were filtered using fastp software (removing adapters, reads with N > 10%, and Q ≤ 5). After alignment to the reference genome (GCA_052148935.1), variant calling was performed using GATK and BCFtools with filtering conditions: MAF ≥ 0.05, missing rate < 0.05. Functional annotation was conducted using snpEff, and SNPs were classified into Ts and Tv as well as categorized by genomic regions.

Phylogenetic and population structure analysis

The genetic distance matrix among samples was calculated using Plink software, and a phylogenetic tree was constructed for the 233 samples. The values in the matrix represent the genetic distances, with smaller values indicating closer relatedness and larger values indicating more distant relationships. Population structure of all samples was analyzed using Admixture software. The number of assumed subpopulations (K value) ranged from 1 to 10. Cross-validation errors were checked using grep, and the K value corresponding to the minimum cv error was selected.

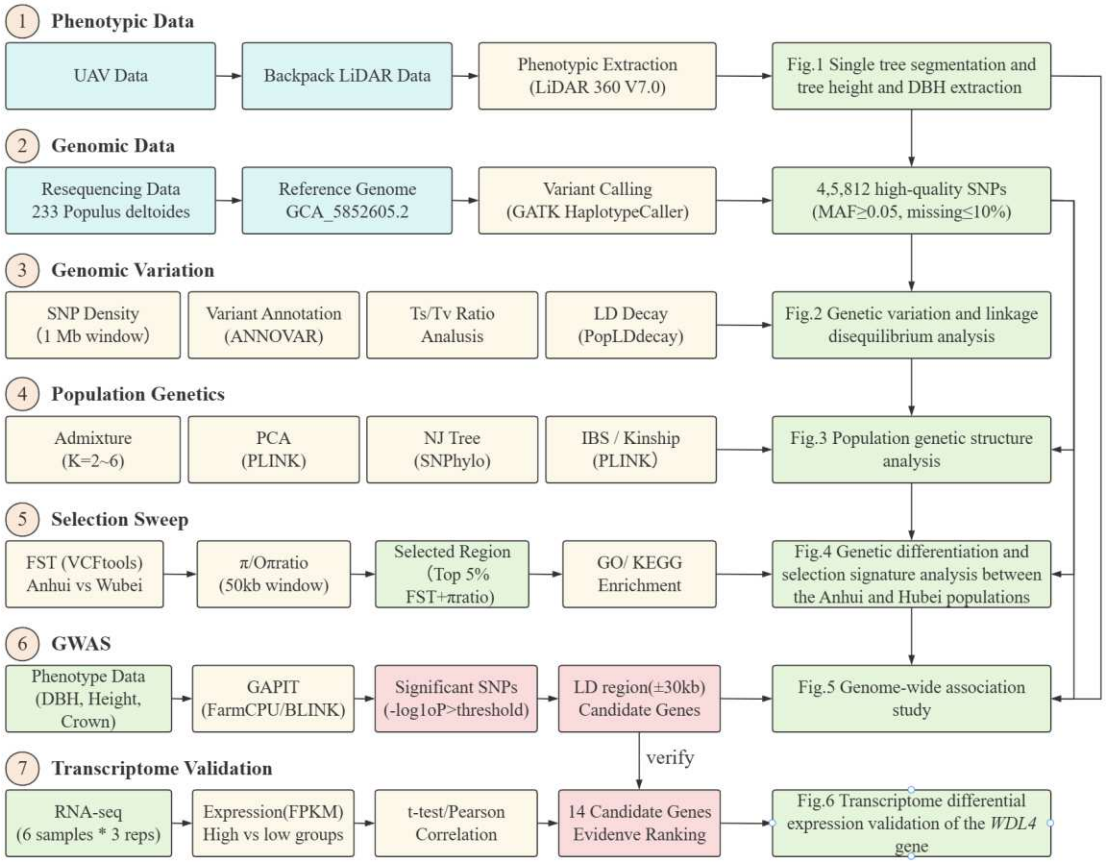


Figure 7 . The analytical method and flowchart used in the research.

Principal component analysis and population genetics

To assess population structure, PCA was performed using Plink software to cluster the sample populations, followed by drawing PCA scatter plots using R language. F_{st} was calculated using VCFtools software. The top 10% of windows (threshold value 0.1584) ranked by window F_{st} values were selected as significant

candidate windows. Additionally, a sliding window approach with a window size of 10 kb was used to calculate nucleotide diversity (π) via VCFtools. The top 5% of π windows were identified as candidate selection regions. Linkage disequilibrium coefficient (r^2) was calculated using PLINK, and the LD decay distance was determined using a threshold of $r^2 = 0.20$.

GWAS analysis

In this study, whole-genome resequencing was performed on *P. deltoides* germplasm using the Illumina PE150 high-throughput sequencing platform with a sequencing depth of 30×. The sequencing data were aligned to the *P. deltoides* reference genome (GCA_052148935.1) using the BWA-MEM algorithm, achieving a mapping rate and genome coverage of over 87%. Variant filtering was conducted using PLINK software, retaining variants with a quality score (QC) > 20 or INFO/DP > 20. SNPs were further filtered using the PLINK toolset with the following criteria: minor allele frequency (MAF) > 0.05 and minimum mapping quality (MQ) = 0.4. This yielded 4,275,812 high-quality SNP markers for subsequent GWAS analysis. GWAS for phenotypic traits of *P. deltoides* were performed using two models, FarmCPU and BLINK, with the first three principal components included as covariates and the kinship matrix K used to control for relatedness inflation (Ye *et al.*, 2024). To ensure statistical stringency, the Bonferroni correction was applied, setting the significance threshold at $P < 0.5/3{,}667{,}348 = 1.38 \times 10^{-8}$ which corresponds to $-\log_{10}(P) = 7.86$. Candidate genes were identified within a 30 kb window surrounding significant SNPs and functionally annotated using the NR, SwissProt, PFAM, GO, and KEGG databases. In addition, Manhattan plots were used to visually annotate significantly associated SNP loci, and Q-Q plots were employed to evaluate the overall fit of the association results, thereby validating model robustness and the reliability of the analysis.

Linkage disequilibrium analysis and candidate gene screening

Population LD decay distance was calculated using PopLDdecay-v3.40 software for linkage disequilibrium analysis, with a linkage disequilibrium coefficient threshold (r^2) greater than 0.2. Based on the obtained LD decay value of 10 kb as the standard, gene annotation was performed on sequences within approximately 10 kb on both sides of the variant sites using the *P. deltoides* genome database (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_052148935.1/) and the protein database (https://www.uniprot.org/). Final candidate genes were selected from those associated with target traits based on gene function annotations and previous research reports.

Transcriptome differential expression and qPCR validation

Samples were divided into high-value and low-value groups using the median of each trait as the cutoff, and each group contained three biological replicates. Agarose gel electrophoresis was used to assess RNA integrity and the presence of DNA contamination. RNA purity, including OD260/280 and OD260/230 ratios, was measured using a NanoPhotometer spectrophotometer. RNA concentration was precisely quantified using a Qubit2.0 Fluorometer, and RNA integrity was accurately assessed using an Agilent 2100 bioanalyzer. After ensuring that library quality met

the standards, sequencing was performed on an Illumina platform. Raw data were filtered using fastp v0.19.3, primarily removing adapter-containing reads; paired reads were removed if the N content in either read exceeded 10% of its bases, or if the number of low-quality bases ($Q \leq 20$) in either read exceeded 50% of its bases. All subsequent analyses were based on clean reads. Differential expression analysis was performed using DESeq2 (v1.20.0) with a negative binomial distribution model, using an adjusted P-value ≤ 0.05 and $|\log_2 \text{fold change}| \geq 1$ as the criteria for differentially expressed genes. For functional annotation, GO enrichment analysis was conducted using clusterProfiler (v3.8.1) with correction for gene length bias, and metabolic pathway annotation was performed using the KEGG database. qPCR validation was carried out using the $2^{-\Delta\Delta Ct}$ method (with Actin/GAPDH as internal reference genes) to provide a basis for subsequent mechanistic studies.

Conclusion

In this study, multi-platform lidar phenotypic analysis and whole genome resequencing were integrated for 233 *P.deltoides* germplasm resources, and the population genomics and complex genetic structure controlling the height and DBH development of P.deltoides were systematically revealed. Population analyses reveal divergent linkage disequilibrium decay patterns and moderate genetic differentiation, reflecting distinct mating systems and evolutionary histories. The marked reduction of SNPs in coding regions underscores strong negative selection acting on functional variants. Furthermore, selective sweep analysis highlights asymmetric selection in genomic regions enriched for transcription initiation and phenylpropanoid metabolism. Crucially, multi-model GWAS identified key genetic determinants of structural traits, pinpointing *PSK6*, *AUG8*, and *WDL4* as primary regulators of tree height, DBH, and crown width, respectively. These genes functionally converge on hormone signaling, cell division, and microtubule organization, with transcriptome profiling validating the negative regulatory action of *WDL4* on crown expansion. Ultimately, these integrated genomic insights provide precise molecular targets and a robust theoretical framework to advance genomic-assisted and intelligent breeding programs in *P. deltoides*.

CRediT authorship contribution statement

Xiaoning Ge: Writing - review & editing, Writing - original draft, Methodology, Investigation, Conceptualization, Formal analysis. **Huaiqing Zhang**: Writing - review & editing, Supervision, Methodology, Investigation, Conceptualization, Funding acquisition. **Chenggong Liu, Xiaoliang Shan**: Writing - review & Editing, Methodology, Data curation. **Xueyan Zhu, Jing Zhang, Yang Liu**: Supervision, Formal analysis, Data curation.

Acknowledgements

The authors would like to thank the Anhui Academy of Forestry and the Hubei Academy of Forestry for providing data on *Populus deltoides* germplasm resources.
Funding: This research was supported by the Fundamental Research Funds of CAF

[CAFYBB2024QD001-06], and the Fundamental Research Funds of CAF [CAFYBB2023PA003].

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials. Example data used in this study have been uploaded to the web site https://github.com/GXN-A/Integrating-multi-platform-LiDAR-phenotyping-Populus-deltoides.git. Whole-genome resequencing and transcriptome data have been uploaded to https://ngdc.cncb.ac.cn/gsub/, with the accession number PRJCA065898.

Conflict of interest

The authors have no conflicts of interest to declare.

References

Darmagaru S, Rahaman AF, Jabeen F, et al. Genome-wide association studies for identifying SNPs and candidate genes associated with grain quality traits in rice (*Oryza sativa* L.). Plant Genome. 2026, 19(2):e70190.

Demirjian C, Vailleau F, Berthomé R. *et al.* Genome-wide association studies in plant pathosystems: success or failure? Trends Plant Sci. 2022, 28:471-85

Ding X Y, Pelser P B, Xu C, *et al.* Leveraging close-range UAV phenotyping and GWAS for enhanced understanding of slash pine growth dynamics. Information Processing in Agriculture. 2025, 12(4):550-564.

Dong M, He J. Linking the past to the future: a reality check on cross-border timber trade from *Myanmar* (Burma) to China. Forest Policy Econ. 2018, 87:11-9

Dong N Q, Lin H X. Contribution of phenylpropanoid metabolism to plant development and plant-environment interactions. J Integr Plant Biol. 2021, 63(1):180-209.

Du Q Z, Gong C R, Wang Q S, *et al.* Genetic architecture of growth traits in *Populus* revealed by integrated quantitative trait locus (QTL) analysis and association studies. New Phytol. 2016, 209: 1067-1082.

Fahrenkrog A M, Neves L G, Resende M F R Jr, *et al.* Population genomics of the eastern cottonwood (*Populus deltoides*). Ecol Evol. 2017, 7(22):9426-9440.

Fang X J, Wu W R. Evolution of statistical genetic paradigms: from linkage analysis and candidate gene strategies to GWAS. Molecular Plant Breeding. 2026, 24(9):1-17

Farquharson K L. AUGMIN subunit8 participates in microtubule reorientation in *Arabidopsis* hypocotyls. Plant Cell. 2013, 25(6):1913.

Geraldes A, Farzaneh N, Grassa C J, *et al.* Landscape genomics of *Populus trichocarpa*: the role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. Evolution. 2014, 68(11):3260-80.

Guo C C, Yin T M, Wu H T, *et al.* Genomic selection with GWAS-identified QTL markers enhances prediction accuracy for quantitative traits in poplar (*Populus deltoides*). Commun Biol. 2025, 8(1):1242.

Jin Z Y, Li P, Huang R, *et al.* Natural variation in *PtobZIP18* confers the trade-off between stem growth and drought tolerance in *Populus*. Plant Biotechnol J. 2025, 23(10):4633-4649.

Li L Z, Huang R, Jin Z Y, *et al.* Natural variants of *AGL80.5* and *FPA.3* contribute to bud break timing in poplar by controlling auxin biosynthesis. Sci Adv. 2025, 11(51):eadz6824.

Li P, Xiao L, Du Q Z, *et al.*Genomic insights into selection for heterozygous alleles and woody traits in *Populus tomentosa*. Plant biotechnology journal. 2023, 21(10):2002-2018.

Lian N, Liu X M, Wang X H, *et al. COP1* mediates dark-specific degradation of microtubule-associated protein *WDL3* in regulating *Arabidopsis* hypocotyl elongation. Proc Natl Acad Sci U S A. 2017, 114(46):12321-12326.

Lima-Huanca L, Alvarez-Vasquez A, Valderrama-Valencia M, *et al.* In silico characterisation and determination of gene expression levels of the *CPK* family under saline stress conditions in *Chenopodium quinoa* willd. Int J Mol Sci. 2025, 26(21):10658.
Luan S. Calcium signaling in plants: Universal and unique paradigms. Cell. 2026, 189(4):1001-1023.
Ma T Y, Hou J. Trait assessment of 1122 *Populus deltoides* clones: unveiling correlations among growth, wood properties, and disease susceptibility. Forests. 2024, 15:1250
Mahinder P, Vipasha V, Meenakshi T. *et al.* Designing of future ornamental crops: a biotechnological driven perspective. Hortic Res. 2023, 10:uhad192.
Ninomiya S, Baret F, Cheng ZM. Plant Phenomics: Emerging Transdisciplinary Science. Plant Phenomics. 2019, 2019:2765120.
Perrin R M, Wang Y, Yuen C Y, *et al. WVD2* is a novel microtubule-associated protein in *Arabidopsis thaliana*. Plant J. 2007, 49(6):961-71.
Schaefer K, Cairo Baza A, Huang T, *et al. WAVE-DAMPENED2-LIKE4* modulates the hyper-elongation of light-grown hypocotyl cells. Plant Physiol. 2023, 192(4):2687-2702.
Shi Y J, Huang J L, Wan X Q, *et al.* The population chloroplast genomes of *Populus* reveal the phylogenetic relationship between three new taxa of sect. BMC Genomics. 2025, 26(1):156.
Slavov G T, Macaya-Sanz D, DiFazio S P, *et al.* Population structure limits the use of genomic data for predicting phenotypes and managing genetic resources in forest trees. Proc Natl Acad Sci U S A. 2025, 122(26):e2425691122.
Tan B. & Ingvarsson P K. Integrating genome-wide association mapping of additive and dominance genetic effects to improve genomic prediction accuracy in Eucalyptus. Plant Genome. 2022, 15, e20208.
Tan J, Liu S, Li G, et al. UAV-LiDAR high-throughput time-series phenotyping and genome-wide association analysis reveal the genetic basis of plant height in peanut (*Arachis hypogaea* L.). Plant Phenomics. 2025, 7(4):100139.
Tibbs Cortes L, Zhang Z, Yu J. Status and prospects of genome-wide association studies in plants. Plant Genome. 2021, 14, e20077.
Wang J H, Ding C J, Li W, *et al.* Advances in genetic breeding research of Chinese forest trees in 2024. Scientia Silvae Sinicae. 2025, 61(7): 35-51.
Wang Y L, Wang Z S, Zhu S, *et al.* Analysis of growth trajectories and verification of related SNPs in *Populus deltoides*. Int J Mol Sci. 2023, 24(22):16192.
Wei S Y, Wu H T, Li X P, *et al.* Identification of genes underlying the resistance to *Melampsora larici-populina* in an R gene supercluster of the *Populus deltoides* genome. Plant Dis. 2020, 104:1133-1143.
Wei Y H, Zhang Y F, Li J X, *et al.* Unravelling the role of phytosulfokine (*PSK*) in regulating rose petal growth through cell proliferation. Plant J. 2025, 123(3):e70424.
Xiao L, Fang Y Y, Zhang H, *et al.* Natural variation in the prolyl 4-hydroxylase gene *PtoP4H9* contributes to perennial stem growth in *Populus*. Plant Cell. 2023, 35(11):4046-4065.
Xu Y B, Zhang X P, Li H H, *et al.* Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. Molecular Plant. 2022, 15(11):1664-1695.
Xue L J, Wu H T, Chen Y N, *et al.* Evidences for a role of two Y-specific genes in sex determination in *Populus deltoides*. Nat. Commun. 2020, 11:5893.
Ye Q N, Zhang L, Li Q, *et al.* Genome and GWAS analysis identified genes significantly related to phenotypic state of *Rhododendron* bark. Hortic Res. 2024, 11(3):uhae008.
Zeng J, Xue A, Jiang L, *et al*. Widespread signatures of natural selection across human complex traits and functional genomic categories. Nat Commun. 2021, 12(1):1164.
Zhang D P, Di Q, Gui J S, *et al.* Tyrosylprotein sulfotransferase positively regulates symbiotic nodulation and root growth. Plant Cell Environ. 2025, 48(1):553-570.
Zhang P H, Zhao J Z, Zhang W, *et al.* Sulfated peptides: key players in plant development, growth, and stress responses. Front Plant Sci. 2024, 15:1474111.
Zhou X L, Zhang L, Zhang M, *et al.* Genomic selection for growth and wood properties in multi-generation hybrid populations of *Populus deltoides*. Hortic Res. 2025, 12(9):uhaf165.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytables.xlsx