Journal of Tropical Forest Science 36(1): 26–39 (2024)
https://doi.org/10.26525/jtfs2024.36.1.26
ISSN: 0128-1283, eISSN: 2521-9847

PINUS ARMANDII GROWTH GENE IDENTIFICATION USING GENOME-WIDE ASSOCIATION STUDY APPROACHES

Li QS[1, 2], Liu C[3], Xin YX[1, 2], Shen WX[1, 2], Zhao WZ[1, 2], Wang F[1, 2], Cao ZY[1, 2], Bai B[4] & Xin PY[1, 2, *]

[1]*Southwest Research Center for Landscape Architecture Engineering, National Forestry and Grassland Administration, Southwest Forestry University, Kunming, China*
[2]*Key Laboratory of Forest Resources Conservation and Utilization in the Southwest Mountains of China Ministry of Education, Southwest Forestry University, Kunming, China*
[3]*Dehong Tropical Agricultural Research Institute of Yunnan Province, Ruili, China*
[4]*Gansu Academy of Agricultural Sciences, Lanzhou, China*

**xpytgyx@163.com*

Submitted November 2022; accepted April 2023

In the study to identify key genes and molecular markers related to *Pinus armandii* Franch growth and to obtain shorten breeding cycles, we selected 209 single plants from six seed sources from a clonal *P. armandii* seed orchard on Zixi Mountain, Chuxiong City, Yunnan Province, China and examined tree height and diameter at breast height (DBH) traits. We then studied single nucleotide polymorphisms (SNP) markers related to tree height and DBH using Genome-Wide Association Studies (GWAS) and predicted candidate genes associated with growth traits using quantitative polymerase chain reaction (qPCR) to preliminarily verify these genes. We identified 13 SNP loci in the population, of which six and seven SNPs were related to DBH and tree height, respectively. We also identified a candidate gene caffeoyl-CoA-O-methyltransferase (CCoAOMT) which has important roles in lignin monomer synthesis. Another candidate gene cellulose synthase A2 (CesA2) was involved in cellulose synthesis in different organs, tissue, and the cell wall. The SNP marker (Marker227806) was associated with DBH, the genotypes of marker227806 in 209 samples of *P. armandii*. In total, 175 samples had homozygous (GG) and 34 had heterozygous (GT) genotypes, and average DBH values were 23.80 cm and 26.10 cm, respectively, indicating significant differences between two genotypes for DBH ($P = 0.037, < 0.05$), which could be used for optimal DBH selection. The candidate genes CCoAOMT and CesA2 were evaluated by qPCR and AT5G42190 was selected as an internal reference gene. Needle (ZY), bark (SP), root (SG), phloem (RP), and xylem (MZ) tissues from 15 samples were used to analyze gene expression, with CCoAOMT and CesA2 identified in all tissue, but expression was highest in RP and relatively low in SP tissue, with significant differences observed. Both genes were preliminarily related to *P. armandii* growth traits and may be implicated in DBH and tree height. Our study provides important information for the genetic improvement of *P. armandii*.

Keywords: *Pinus armandii* Franch, Genome-Wide Association Studies, diameter at breast height, tree height, Caffeoyl-CoA-O-methyltransferase, Cellulose synthase A2.

INTRODUCTION

Pinus armandii Franch (Five-Needle Pine) has a tall and straight shape and is the most widely distributed species in the pine family. The seeds have a high oil content and are used for oil and food production; therefore, the species has recognized ornamental and edible properties (Chen & Shen 2005, Li 2007). The tree is also an important afforestation and timber species (Yuan & Chen 2010) and is vital for timber and forestry construction in China (Zhang & Cai 1989).

Growth traits and timber properties are important indicators when evaluating timber species. Japarudin et al. (2022) measured the growth and wood traits of pure seed progeny from *Eucalyptus pellita* in eastern Malaysia and selected parent and progeny families with high general combining abilities, which provided materials for *E. pellita* genetic improvement. Nguyen et al. (2019) identified growth and wood traits in the offspring from the *Melaleuca cajuputi* half-sib family, evaluated and selected parent and offspring families with higher general combining abilities, and provided materials for *M. cajuputi* genetic improvement. Therefore, selecting germplasm resources with good growth trait performances such as tree height, diameter at breast height and

timber properties such as density are required to improve *P. armandii* Franch timber value (He 2019, Xu et al. 2014). However, multiple factors including large and highly heterozygous genomes (complex genetic backgrounds) (Chen et al. 2002) and long growth cycles have limited forest tree genome research when compared with other crops. Consequently, little research has been conducted on the genetic variations of important forest tree traits (diameter at breast height, height, growth, cone yield and material properties), thus limiting gene function-mining and metabolic pathway analysis and indirectly affecting the development and selection of superior tree resources (Wan et al. 2012). Additionally, traditional breeding methods are limited by long breeding cycles and low effectiveness, making it impossible to meet the increasing demands for greater quantities and quality of forest products (Xin et al. 2010).

With continuous improvements and developments in molecular marker technology, first-generation representative restriction fragment length polymorphisms have been developed into third-generation single nucleotide polymorphisms (SNPs) (Shao 2015). SNPs are used as new generation molecular marker technologies for plant trait marker localisation and genetic map construction as they are advantageous in terms of high density, representativeness, genetic stability and wide distribution (Jia 2003). Genome-wide association studies (GWASs) are based on genetic linkage disequilibrium, with many relevant genome-wide SNP markers developed using study material. By identifying SNP markers, loci or candidate genes with specific functions associated with important phenotypic trait variations, can be screened and the genetic effects analyzed (Duan & Zhu 2015, Loots et al. 2000). Genome-wide association studies (GWASs) are widely used in association studies examining important plant traits, especially agricultural crops; traits have included yields, plant height, spike length, flowering times and tassel stages in rice (Zaid et al. 2017), wheat (Chen et al. 2015), and maize (Farfan et al. 2015), and several genetic loci related to important traits of crops have been identified. Genome-wide association studies (GWASs) have generated data related to screening and has revealed some results related to plant disease resistance gene screening genes (Xiao et al. 2016) and metabolite synthesis in plants (Lipka et al. 2013). However, there are few reports related to GWASs in forest trees due have been published as trees having highly heterozygous genomes, among other reasons. The screening of genes related to material properties and resistance, mainly in *Populus euphratica, Catalpa bungei* and *Populus* (Porth et al. 2013), gene localization of genes related to *Eucalyptus* growth traits in *Eucalyptus* (Resende 2016) and screening of candidate genes related to the peach fruit tree domestication and improvement of the economic in peach fruit tree (Cao et al. 2019) can provide a theoretical basis insights and technical guidance for GWASs of DBH and height traits trait GWASs in *P. armandii* Franch. Therefore, in order to shorten the breeding cycles and improve the breeding efficiency of *Pinusin P. armandii*, it is important to conduct molecular marker-assisted breeding by GWASs and based on their high-quality SNPs development using specific loci amplified fragment sequencing (SLAF-seq) sequencing from the perspective of molecular biology.

MATERIALS AND METHODS

Sampling site

In 1986, the clonal *P. armandii* seed orchard in ZiXinshan Forest Farm, Chuxiong City, Yunnan Province, China was formally established and is located at E 24°58′58″–25 24°58′N at an altitude of 2200–2400 m and has an area of approximately 30 hm2. The annual average temperature is approximately 12°C, the annual average rainfall is approximately 1000 mm, and the soil is dark red. The orchard contains 124 superior *P. armandii* clones from six provenances (Table 1) (Weishan, Tengchong, Chuxiong, Nanhua, Huize, and Yiliang). The geographical location of each clone provenance is outlines (Figure 1).

Materials

Previous research materials came from the clonal *P. armandii* seed orchard on Zixi Mountain. Tree height and DBH were measured by altimetry and a tree circumference ruler. In total, 1453 individuals were measured. Then,

Table 1 *Pinus armandii* material sources

Clonal source	Longitude East	Latitude North	Number of samples (Total = 209)	Provenance number
Huize	103°03′–103°55′	25°48′–27°04′	59	HZ
Weishan	99°55′–100°25′	24°56′–25°32′	72	WS
Tenchong	98°05′–98°45′	24°38′–25°52′	4	TC
Chuxiong	100°35′–101°48′	24°30′–25°15′	55	CX
Nanhua	100°44′–101°20′	24°44′–25°21′	17	NH
Yiliang	102°58′–103°28′	24°30′–25°17′	2	YL

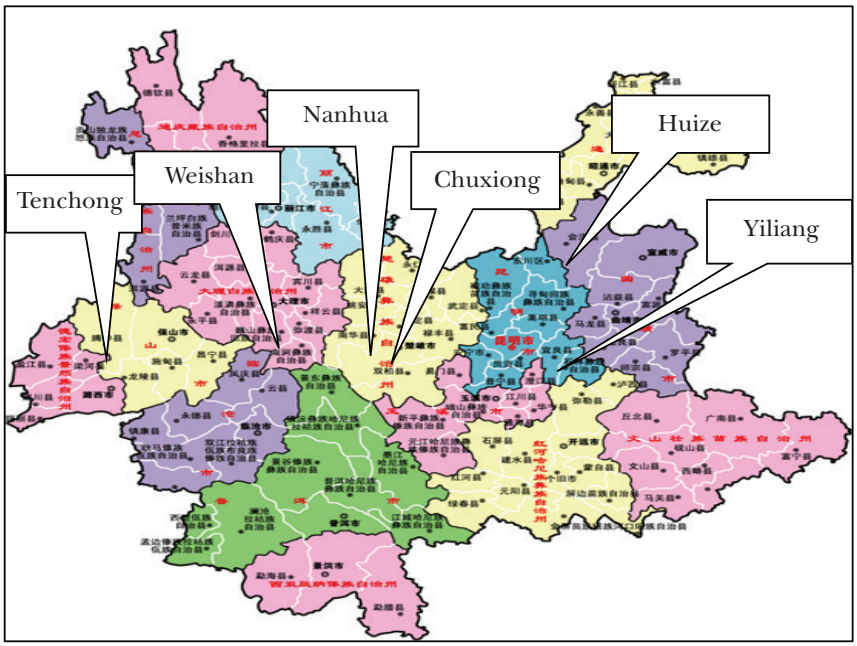

Figure 1 Map showing the geographical location of clones

based on phenotypic tree height and DBH survey data, individual plants were screened and arranged from large to small. Finally, 30 high-growth individuals, with tree heights > 15 m and DBH > 30.6 cm, and 31 low-growth individuals with tree heights < 8.7 m and DBH < 17.5 cm were selected. Then, average tree height, DBH, and cone number values were calculated and analyzed. With reference to tree height > 10.9 m and DBH > 22.2 cm, 148 trees with tree heights and DBH at middle and upper levels were selected to generate 209 samples. Additionally, three superior individuals were selected for quantitative polymerase chain reaction (qPCR) studies, and young tree parts, including needles (ZY 1-3), bark (SP 1-3), roots (SG 1-3), phloem (RP 1-3), and xylem (MZ 1-3) were collected. In total, 15 samples from each plant were collected and three biological replicates performed.

Genome-wide association studies (GWAS)

Fundamentally, GWASs are used to associate SNPs with phenotypic data. Different SNP types were combined for population analyses, and the threshold for screening important components in the environment was $P = 10^{-5}$ (1/N, where N = number of markers). Under environmental influences, SNP results did not reach the first selection threshold ($P < 10^{-4}$), but for well-repeated marker loci, we performed statistical analyses and mined phenotypic traits associated with functional genes.

GWASs were conducted as follows: 1) From breeding goals, the research population was established, the largest population was selected as the research sample, and a target trait database was established. 2) Sample DNA was extracted, quality controls performed to genotypic

requirements, and genotype data evaluated and subjected to quality controls to meet analysis requirements. 3) Appropriate statistical models (General Linear Model; Mixed Linear Model) were used to analyze associations between SNPs and target traits. 4) Further analyses and data verifications were performed after correlation analyzes. Trait Analysis by aSSociation, Evolution and Linkage (TASSEL) software (Bradbury et al. 2007) was used for association analyses based on SNP molecular marker development related to growth. The formula was: $y = X\alpha + Q\beta + K\mu + e$, where α = fixed effect vector; β = SNP substitution effect; μ = random additive genetic effect vector subject to distribution N $(0, G\sigma^2)$, G = kinship matrix derived from SNP markers; σ^2 = additive variance, and e = the residual.

Admixture software (Alexander et al. 2009) was used to calculate Q values (the overall sample structure) and SPAGeDi software (Hardy & Vekemans, 2002) was used to calculate K values (kinship between samples). A general linear model (GLM) was only required to calculate Q values, and a mixed linear model (MLM) to calculate Q and K values, where X = genotype and Y = phenotype. These approaches showed that each SNP marker locus was putatively associated with a result (Initiative 2000).

QPCR analysis of candidate gene expression

Total RNA extraction from P. armandii Franch

For qPCR studies, tender parts of three superior individuals were used, including needles (ZY 1-3), bark (SP 1-3), roots (SG 1-3), phloem (RP 1-3), and xylem (MZ 1-3). In total, 15 samples were collected from plants, with three biological replicates performed. Total *P. armandii* RNA was extracted using a plant RNA extraction kit (OMEGA R6827, Beijing Ya 'anda Biotechnology Co., Ltd, Beijing, China).

RNA concentration and quality

RNA concentrations were determined using an Ultramicro UV spectrophotometer (UL-5000, Shanghai Meixi Instrument Co., Ltd, Shanghai, Beijing). To determine impurities, approximately 1 µL of newly extracted RNA was measured at an absorbance ratio of 260 nm to 280 nm; i.e., the ideal OD260/OD280 value was 1.8–2.0.

RNA quality: On a 1% agarose gel, newly extracted RNA was subjected to electrophoresis for approximately 20 min. A gel imager (E-BOX-CX5, French Vilber Lourmat, Paris, France) was then used to observe and photograph gels. When RNA quality was satisfactory and met standard requirements, it was used for reverse transcription studies.

First-strand cDNA synthesis

First-strand cDNA was synthesized using a reverse transcription kit (Kunming Shuoyang Biotechnology Co., Ltd, Kunming, China). Briefly, 2 µL template RNA, 2 µL AccuRT Reaction Mix (4 ×), and 4 µL enzyme-free water were incubated at 42°C for 2 min or room temperature for 5 min. Then, 2 µL AccuRT Reaction Stopper (5 ×), 4 µL 5 × All-in-One RT MasterMix, and 6 µL enzyme-free water were added and incubated at 25°C for 10 min or 42°C for 15 min.

First-strand cDNA characterization

First-strand cDNA concentration and purity levels were determined using a nucleic acid protein analyzer (HM-CWF1, Hengmei Technology Co., Ltd, Zhengzhou, China). An absorbance ratio of 260 nm to 280 nm (1.8–2.0) was ideal. First-strand cDNA quality was examined using 0.8% agarose gel electrophoresis for 30 min, with gels observed on an E-BOX-CX5(Vilber Bio Imaging Company, Paris, France) gel imager.

Candidate gene primer design

After sequence alignments using the National Center for Biotechnology Information (NCBI) website, tree height and DBH related genes were identified and qPCR primers designed using Primer 3.0 (NCBI). Primer length was approximately 18–25 bp, annealing temperature was 55°C–65°C, primer melting temperature differences were not more than 5°C, the GC

content was 40%–60%, and the amplification of the PCR intermediate fragment is generally 100 ~ 400 bp. *AT5G42190*, which is related to pine trees (Xin et al. 2021), was selected as an internal reference gene for qPCR. Its primer sequence was ATGCTGGACAGGCTTTGAAC GAGTTGCTCCGAGATCTTTACA. The primer sequences are shown in Table 2.

Candidate gene validation using qPCR

Light-proof qPCR reaction was performed on an ABI 7500 Fast Real-Time PCR System (Applied Biosystems Company, Massachusetts, USA) using a kit from Kunming Shuoyang Biotechnology. A final 20 μL reaction volume was used (Table 3).

Standard qPCR amplification parameters: pre-denaturation at 95˚C for 30 s; 40 cycles at 95˚C for 5 s and 60˚C for 30 s; melting curve analysis at 95˚C for 15 s, 60˚C for 1 min and 95˚C for 15 s. Cycle threshold (Ct) values were calculated (Microsoft Excel) using the formula 2-△△Ct = 2-[△Ct (Test sample) -△Ct (Control sample)], △Ct (Test sample) = [Ct (Test sample, target gene) - Ct (Test sample, reference gene)], △Ct (Control sample) = [Ct (Control sample, target gene) - Ct (Control sample, reference gene)] (Schmittgen & Livak, 2008). Figures

Table 2 Primers for qPCR analysis

Gene Name	Primer number	Forward primer	Reverse primer
CesA2	CE1	GATAAAATCGGCCGCAACCC	CGAGCGGTTAGCGATACTGT
	CE2	ACAGTATCGCTAACCGCTCG	CAAAGGTCGGTTGCACCTTG
	CE3	TCGCCAAAGCGTAGTCAGTT	GCGGCCAGGATAGAGAGAAC
	CE4	GTACAGGCGTCACAAAGGGA	AACTGACTACGCTTTGGCGA
	CE5	GTTCTCTCTATCCTGGCCGC	GAACCCATTTCCGGGCAAAC
	CE6	CTGGTTGCTTCGGAGGTGAT	CGACACACAGGGAATCCACA
	CE7	GTGAGCTTGGTAGCAGGTGT	TGAAGAATGGGGCCTTCGAC
	CE8	CAAGGTGCAACCGACCTTTG	GCACGGCCTTACTGTTGTTG
	CE9	TCAAATGGACGACCCTGCTC	ACACCTGCTACCAAGCTCAC
	CE10	TGGTAGTCCCTGGACACGAT	AGCGGTTAGCGATACTGTGG
CCoAOMT	CC1	CTTGCATTGCCCGATGATGG	TGATCTCAATGCGGGGATCG
	CC2	TGCTATGCCAGTTCTGGACG	GATCCGTTCCACAGGGTGTT
	CC3	ACTGCCAAGCATCCCTGTTT	CCATCATCGGGCAATGCAAG
	CC4	CAACACCCTGTGGAACGGAT	ATAGACACGCCTGCAAAGGG
	CC5	CTGTGGAACGGATCTGTGGT	ACCAACTGGGATTTGGCTGA
	CC6	AGAGGAGCCGGTTAAGGTTG	GATGCTTGGCAGTCACTTCG
	CC7	CAGAGGAGCCGGTTAAGGTT	CTTCGCGGAGCTCCTTCATT
	CC8	GTCGATCCCCGCATTGAGAT	CAATAGACACGCCTGCAAAGG
	CC9	TGACAACACCCTGTGGAACG	TTGGCTGATCTCAATGCGGG

Table 3 qPCR reagents and volumes

Reagent	Volume (final 20 μL)
EvaGreen 2×qPCR MasterMix	10 μL
Forward primer	0.6 μL
Reverse primer	0.6 μL
Template DNA	1.0 μL
Enzyme-free water	7.8 μL

were plotted using Origin (OriginLab Company, Massachusetts, USA).

RESULTS

Genome-wide association studies (GWAS)

Multiple hypothesis testing corrections

A MLM was established by combining population structure (Q) and kinship matrices (K) using the Bonferroni method, after which, association markers were accurately located to reduce the probability of false positive P values. From QQ-plot (Quantile-Quantile Plot) data (Figure 2), under the assumption that no correlations existed between marker loci and phenotypic traits, observed P value and the expected P value in the graph show a deviation at the right end. Thus, DBH differences were not caused by population stratification and the MLM for GWAS was suitable.

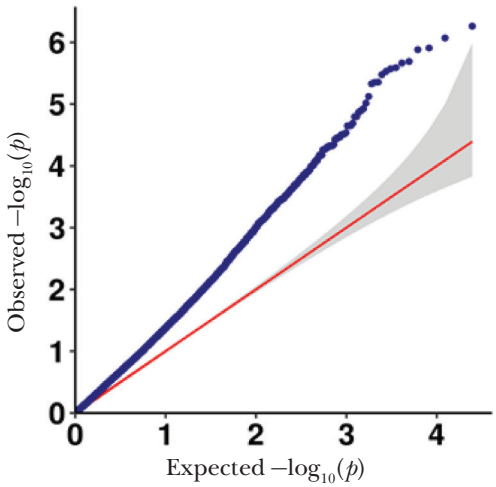

Figure 2 QQ map of genome-wide association analysis of height traits

Note: Horizontal and vertical coordinates are expected and observed values, respectively. The red line indicates the 45° centerline and the gray area indicates the 95% confidence interval of the scatter.

GWAS of height and DBH

Using 3,469,074 SNP markers from *P. armandii* populations (Wang et al. 2022), GWAS analyses of phenotypic tree height and DBH traits were conducted. Six and seven SNP loci related to DBH and tree height, respectively, were generated by the GLM. The location of the 13 SNPs on the SLAF label and corresponding base sequences are shown (Table 4). After BLAST matching using SNP sequences in NCBI, two genes were identified as potentially associated with DBH: *CesA2* (accession no. AC241331.1) associated with cellulose synthesis (similar to the Marker227806 locus), and *CCoAOMT* (accession no. AC241311.1) associated with lignin synthesis (similar to the Marker227806 locus).

Population analysis of SNP markers associated with DBH growth

A statistical analysis of 209 samples (*P. armandii* Franch with DBH phenotypic data combined with associated SNPs (Marker227806) genotypes) identified two genotypes: GG (175 samples) and GT (34 sample) (Table 5). Evaluation of the GT heterozygous genotype combined with mean DBH values showed that the mean DBH of genotype GG was 23.80 cm, and for the genotype GT, it was 26.10 cm. T test analyzes of DBH means identified a P = 0.037 value (< 0.05), indicating significant DBH differences between samples with both SNP (Marker227806) genotypes. These data indicated that this SNP was associated with DBH traits, and the favorable genotype was GG.

Tissue-specific qPCR gene analysis associated with DBH growth

Tissue-specific qPCR expression analysis of candidate genes

After screening, CE1 and CC5 primer pairs were selected for *CesA2* and *CCoAOMT*, which were both stably expressed in tissue. Using correlation analyses, *AT5G42190* was selected as an internal qPCR reference gene to verify and analyze *P. armandii* gene expression in different tissue. Materials for analyzes included needles (ZY), bark (SP), roots (SG), phloem (RP), and xylem (MZ). In total, 15 samples, incorporating three biological replicates, were used to generate Ct values, from which $2^{-\Delta\Delta}C$ values were calculated and tissue expression maps generated. We observed that CesA2, which is related to cellulose synthesis, was expressed in

Table 4 Single Nucleotide Polymorphisms (SNPs) in sequences

Trait	SLAF Sequence No.	Corresponding physical position of the SNP in the SLAF	SLAF Label
DBH	Marker 227806	12	ATCAACTAACTACTGGTGGGTAAGGGTTAGAT CTTCCTGTCAGAGACTAGTACACTTTAGGGAG TCAAAGCTCACCCCATTTTTTATGTAGTGACC AAGANNNNNNNNNNCAAAAGTAAAAATATTA AAAAATAATAAAATAATAAAACTACTTTTTCA ACTTTTATTTTCAACTTATCTCCTTAATTTTTT ACTAAAATTTCAAGTGGT
	Marker 338085	171	GAATCCTCATTTCAAACTCCTCATGTGCCTGCTGAC-CATTATGGGGGATTTTGAAGACCTTATTAATGAAAC-CCTAGACCTAATGACACTAAAACCCTAANNNNNNN-NNNCTCGTCAGATGATCCATGCTCCTAATGGAATG-CAATTGACTAAATGAGACACCTAATATGCAATGCAA-CAAGAGGATTTGATCACTGATGGAGGGATGAA
	Marker 364221	35	TCGTACTGTTGTAATTTAATTAGCCAGATTCAG-GAATATTAATATGACTATGTGTTTGGGTTGACATTT-GACTCACCCAAACAATAGTGGCAGTTAAGAGNNNNN-NNNNNTGCCGAGCTAGTGATGCTGCATAGGTTATGT-GACATGCTGCTGCTAAATAAGAACATTGCTGATGAG-CAGAGAGCTCTGAGAATGAGGTTTGTTACTGTT
	Marker 2961762	39	GCCTTTTGTGAGCTAGAATTCCGAGAATAACAAACA-CAGCCAAATATTTCCTTCCCACTGAACAACCAGAG-GAAATAATGTGAAAAACCCAGGGAAAATANNNNNNN-NNNATTATATCAGGTGTTCTTCATTTAATCGCTTTG-GAAACTAATTCTTTTAAACCCTAAATAATCTAGTATAAT-GAAAGACTCTAAAATGCCATTAACTATT
	Marker 13673505	72	GTACATATCCATAGATGAGAAGATTGTGGACAT-ACTTCAAATATAAGCTTAGTATGATATTTAATCTCTC-CCTAATTGGAACATAATGGTGAGGTATTTTNNNNN-NNNNNGAAAATGCAGGGGAAAATTATAGGATACAT-TCACCTTTGAAGCCAGAAGGACTTGGCTAAGTCAA-CAAGAGGTGAATGTACAAGGGAAATTATAGGATTG
	Marker 27848158	46	AAAAGATGCTAACGGCAATAAGGAGAAAAT-TACTTCTAGCTAGGTATGAGGAGAGATATTA-CAGAATATCTAAATAAGTGTTTACAATGCCAGCAAGT-GAANNNNNNNNNNGGAGCGAGGAAGCAAAGCAA-CACACGTTATACCAGTACAGTCTACTTACAAGACGGT-GTAGATAGCAGACATATTTATGCGAGAGATCTTTAGAC-TACAT
Tree Height	Marker 282498	15	AATTTGAAAATTTTTGTGTCCTCCTCAACACAATT-GATGGGGAGGAATTTGGAATAAATATGGTCTATA-AAAAATTAACATATTCTATTGGGAATTGGTTNNNNN-NNNNNCATAAGAATTGGTAAGAAATGCTTCAGAAAT-GGAAAAGAAAAATCTCTTAGTGAAGAAAATAAGAT-TCAATATCCATGAGATGTTGAATTGATCTTCTAA
	Marker 300606	176	ACTAGGAGAAAATTTTCTCCAAGATGGTGTTCAAC-CAACAACCAACAACCCATTTTAGCGACCAACTAGCT-CATTGTGTCATTTATCCTCCACCTGAGCCNNNNNNN-NNNCATGGATTAATGGAAAGATTTGGTGGAAAGAGT-GGGATAGATCCGTGCTTCTATCCCCTATACCATTTTTG-GATGCTACCCTTGGATCCTAGTCTCGTTG
	Marker 2961762	175	GCCTTTTGTGAGCTAGAATTCCGAGAATAACAAACA-CAGCCAAATATTTCCTTCCCACTGAACAACCAGAG-GAAATAATGTGAAAAACCCAGGGAAAATANNNNNNN-NNNATTATATCAGGTGTTCTTCATTTAATCGCTTTG-GAAACTAATTCTTTTAAACCCTAAATAATCTAGTATAAT-GAAAGACTCTAAAATGCCATTAACTATT

continued

Table 4 Continued

Trait	SLAF Sequence No.	Corresponding physical position of the SNP in the SLAF	SLAF Label
	Marker 13673505	72	GTACATATCCATAGATGAGAAGATTGTGGACATACTTCAAATATAAGCTTAGTATGATATTTAATCTCTCCCTAATTGGAACATAATGGTGAGGTATTTTNNNNNNNNNNGAAAATGCAGGGGAAAATTATAGGATACATTCACCTTTGAAGCCAGAAGGACTTGGCTAAGTCAACAAGAGGTGAATGTACAAGGGAAATTATAGGATTG
	Marker 17725109	182	ATTGGTTATAGTGTTCTGAAGTTTTTGGGCACTCATTGGCTTGGTAGTCATAGAAGAATAAATCATTTCAGAGCTTTTTTTATGAAAATCAATTTGATCTNNNNNNNNNNGTGTAATCAAAAGCTGAGAGTAATCTGTTTTGATGTCCTAGAGTTTGCTTCAGATACTGTGAATAAATATCATTCTTAGCCTTGATTTCTTTATATTTCT
	Marker 20075430	54	TTTGTACTCTTGTAGCAATTTGGTATAAGAGGTTATTTATTCTGGAGAGAATGTTTCCCCCATATATATTTCTTCGACCTTTGCTGAATGTTGGGATATGNNNNNNNNNNGGGCAGGGGGCGGGGGAGGAATGAACATTTTCTTGGGGAGAAAAAGCGAACGGTAATGCATTGCATCCCATGGGTAATCTGGCTTCATCATAGCCTCAAG
	Marker 27848158	198	AAAAGATGCTAACGGCAATAAGGAGAAATTACTTCTAGCTAGGTATGAGGAGAGATATTACAGAATATCTAAATAAGTGTTTACAATGCCAGCAAGTGAANNNNNNNNNNGGAGCGAGGAAGCAAAGCAACACACGTTATACCAGTACAGTCTACTTACAAGACGGTGTAGATAGCAGACATATTTATGCGAGAGATCTTTAGACTACAT

Table 5 Genotype information on *Pinus armandii* strains

Genotype	Sample ID	DBH (cm)	Genotype	Sample ID	DBH (cm)	Genotype	Sample ID	DBH (cm)
	XSQG-1	23.6		XS-1	32.3		XSQG-3	24.2
	XSQG-2	28.1		XS-3	34.9		XSQG-4	28.0
	XSQG-5	30.6		XS-4	34.0		XSQG-15	20.1
	XSQG-6	27.9		XS-5	39.2		XSQG-16	31.4
	XSQG-7	26.2		XS-6	36.3		XSQG-18	26.8
	XSQG-8	24.9		XS-7	35.0		XSQG-46	29.6
	XSQG-9	24.8		XS-8	33.5		XSQG-58	32.7
	XSQG-10	25.1		XS-9	35.1		XSQG-69	29.6
	XSQG-11	26.1		XS-10	33.8		XSQG-72	23.4
	XSQG-12	24.4		XS-11	34.0		XSQG-74	23.8
	XSQG-13	20.7		XS-12	34.2		XSQG-82	27.1
	XSQG-14	27.9		XS-13	31.3		XSQG-88	27.2
	XSQG-17	26.4		XS-14	36.2		XSQG-93	25.3
	XSQG-19	28.1		XS-16	40.8		XSQG-124	23.5
	XSQG-20	23.6		XS-17	34.4		XSQG-129	21.5
GG	XSQG-21	24.4	GG	XS-18	35.5	GT	XSQG-130	23.0
	XSQG-22	27.0		XS-19	33.6		XSQG-134	22.6
	XSQG-43	31.4		XS-20	32.4		XSQG-136	25.5
	XSQG-44	25.3		JSX-1	16.4		XSQG-141	26.8

continued

Table 5 Continued

Genotype	Sample ID	DBH (cm)	Genotype	Sample ID	DBH (cm)	Genotype	Sample ID	DBH (cm)
	XSQG-45	27.2		JSX-2	16.3		XSQG-147	26.0
	XSQG-47	23.1		JSX-3	17.4		XSQG-148	24.8
	XSQG-51	18.6		JSX-5	16.0		XS-2	26.8
	XSQG-52	20.8		JSX-7	16.3		XS-15	24.8
	XSQG-54	28.4		JSX-9	17.5		XS-29	21.5
	XSQG-56	21.1		JSX-10	17.1		JSX-4	17.3
	XSQG-57	22.6		JSX-11	17.4		JSX-6	17.0
	XSQG-61	25.2		JSX-14	16.2		JSX-8	16.8
	XSQG-62	20.2		JSX-15	16.7		JSX-12	16.2
	XSQG-63	22.1		JSX-16	17.0		JSX-13	16.3
	XSQG-64	26.5		JSX-17	17.3		JSX-19	16.3
	XSQG-67	24.0		JSX-18	16.0		JSX-22	16.6
	XSQG-68	26.8		JSX-20	16.7		JSX-23	16.3
	XSQG-73	26.2		JSX-21	17.3		JSX-25	16.5
	XSQG-75	26.7		JSX-24	16.4		JSX-31	16.7

all tissue (Figure 3). Expression in RP tissue was significantly higher when compared with ZY, SP, SG, and MZ tissue, while SP tissue expression was the lowest. CesA2 was confirmed as being related to secondary wall growth, which further exemplified important roles in DBH growth and development. *CCoAOMT*, which is related to lignin synthesis (Figure 4), was expressed in all tissue. Expression levels in RP tissue were significantly higher when compared with ZY, SP, SG, and MZ tissue, while SP tissue expression was relatively low. *CCoAOMT* expression in plants seriously affects lignin monomer levels and total lignin proportions, which suggest key roles in DBH growth and development.

DISCUSSION

GWAS analysis of growth traits in *P. armandii* populations

As indicated by the International Symposium on GWAS Analyses, experts envisioned a promising future for GWASs in forest genetic breeding. For example, the Institute of Forestry, Chinese Academy of Forestry has also carried out GWAS studies on *P.euphratica* and *Catalpa bungei*, mainly to obtain relevant genes in wood properties and resistance indexes of *P.euphratica* and *Catalpa bungei*, in order to obtain gene annotation results of wood properties and resistance traits of *P.euphratica* and *Catalpa bungee* (Wan et al. 2012). Currently, GWASs are widely used in association analyzes of crop traits, and have significantly improved yields, quality, and resistance in cash crops (Zaid et al. 2017, Chen et al. 2015, Farfan et al. 2015). However, GWAS applications in forest tree research are rare due to the large and highly heterozygous genomes of trees. In GWAS studies correlating growth traits and wood properties in forest trees, Porth et al. (2013) examined wood characteristics in 334 *Populus trichocarpa* individuals and identified 141 specific loci associated with 16 wood structural characteristics among 3,500 candidate genes. Resende et al. (2016) used GWAS to mine 13 SNP markers associated with growth, wood, and disease resistance traits in selected *Eucalyptus* hybrid populations. A previous GWAS of fruit shape, quality, and weight in peach trees grown under economic forestry conditions predicted several candidate genes related to fruit characteristics (Cao et al. 2016). In contrast, in our growth trait GWAS analysis in *P. armandii*, only *CCoAOMT* related to lignin synthesis was identified among the 13 SNPs related to height and DBH traits. Fewer associated SNP markers were identified when compared with previous studies (Porth et al. 2013). Additionally, although tree height traits were associated in previous studies, in our study, despite observed tree height associations,

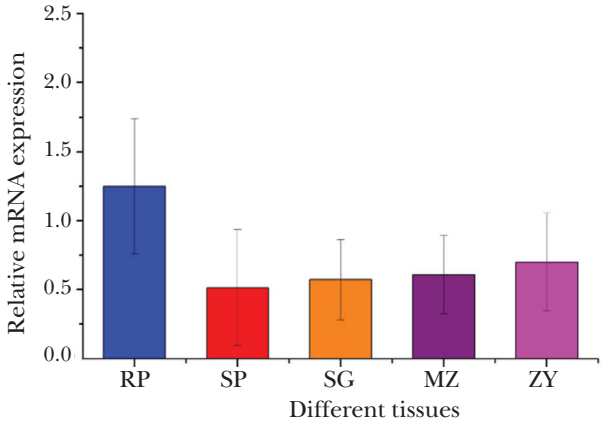

Figure 3 CesA2 expression levels in different *Pinus armandii* tissue

Note: needle (ZY), bark (SP), root (SG), phloem (RP), and xylem (MZ)

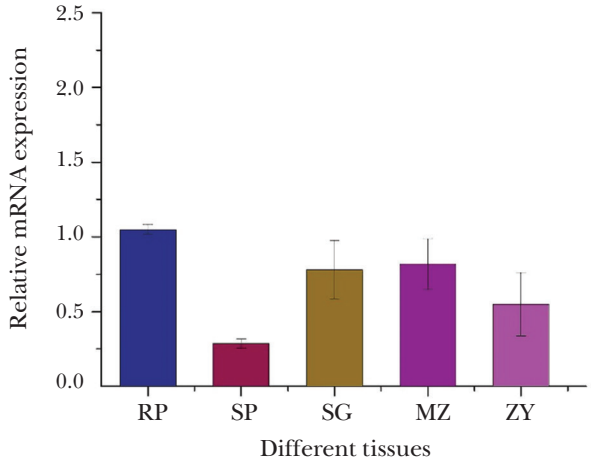

Figure 4 *CCoAOMT* expression levels in different *Pinus armandii* tissue

Note: needle (ZY), bark (SP), root (SG), phloem (RP), and xylem (MZ)

no corresponding genes were identified. One reason for this disparity may be the different materials and methods used in these studies. Coniferous species are highly heterogeneous and exhibit long growth cycles and various adaptations to growing environments, and also show characteristic differences, such as geographical structures and genetic diversity relationships in populations. In our study, only seven and six SNP loci associated with height traits and DBH traits, respectively, were associated using GLM. This model has been used to correlate traits in other tree species, including *Eucalyptus globulus* Labill, *Eucalyptus grandis* Hill, and *Pinus taeda* L. In *E. globulus* Labill, DBH correlations were assessed not only with GLM but also with a unified mixed model (Cappa et al. 2013). GLMs have also been used to study other traits, such as height and DBH in *Eucalyptus grandis* Hill, and no SNPs were associated with either height or DBH, possibly due to differences in growth between *P. armandii* Franch and Eucalyptus grandis Hill (Zhang 2017). When a GLM was used to analyze height traits in *P. taeda* L., 14 SNP loci were identified (Cumbie et al. 2011).

Correlation analysis of SNP markers associated with DBH traits

Numerous studies have reported that the underlying cause of altered or mutated genes

in plants is mutations at one or multiple loci, which in turn alter phenotypic traits (Han 2016). For example, Umemoto and Aoki (2005) speculated that maize *SSIIa* altered enzyme activity, possibly because two SNPs at its C-terminus altered amino acids and thus enzyme activity. In a rice study, Bryan et al. (2000) reported that an amino acid alteration at position 918 between rice blast resistance and susceptibility genes caused functional differences between the genes, with seven base substitutions identified and five amino acid differences generated. In our study, a statistical analysis of one candidate SNP genotype locus (Marker227806) in 209 *P. armandii* Franch samples related to DBH phenotypes, identified two genotypes in this population. In total, 175 samples had the GG genotype and 34 had the GT genotype. GT heterozygous genotype evaluations combined with mean DBH values showed that the mean DBH of genotype GG was 23.80 cm, and the mean DBH of genotype GT was 26.10 cm. T test analyses of mean DBHs revealed a P = 0.037 value (< 0.05), indicating significant DBH differences between the two SNP (Marker227806) genotypes. Thus, this SNP was associated with DBH traits and the favorable genotype was GG. This SNP marker could be used to select superior monocots. Additionally, the genotype may have arisen from purposeful artificial selection and is consistent with current genetic improvement and cross breeding trends. Nevertheless, DBH is regulated by multiple genes - a single gene may not be the main controlling factor. DBH differences are also influenced by the environment, tree variety, region, and considerable genetic heterogeneity across species. Association analyses identified significant differences (P < 0.05) in DBH traits between GG and GT genotypes at the Marker227806 locus, where GG genotype individuals showed significantly higher DBH values when compared with GT genotype individuals. From these analyzes, a mutated Marker227806 locus may be the main factor affecting DBH.

CCoAOMT is related to lignin synthesis

CCoAOMT is related to lignin synthesis in *P. armandii* and was identified in this study. After qPCR analyses, both genes were expressed in xylem, phloem, bark, needle, and root tissue. Gene expression levels in phloem tissue were significantly higher when compared with needle, bark, root, and xylem tissue, while expression levels in bark were the lowest, which further suggested that these genes may be related to DBH growth and metabolism. *CCoAOMT* expression levels in plants strongly affect lignin monomer levels and consequently total lignin levels (Zhang, 2014). Lignin-related gene cloning has been conducted in conifers, including *P. taeda, P. radiata, P. massoniana, P. pinaster, Larix gmelinii, L. kaempferi,* and *Picea abies* (Voo et al. 1995, Li et al. 1999, Chen 2010, Huan et al. 2012, Wagner et al. 2013, Craven-Bartle et al. 2013, Qiao et al. 2016). *CCoAOMT* identified in our study has only been cloned in *P. taeda, P. massoniana*, and *P. radiata*. (Li et al. 2011, Zhang et al. 2014, Wagner et al. 2013). Wagner et al. (2013) successfully cloned lignin-related *CCoAOMT* in *P. radiata*, which was shown to influence lignin content in molecular duct *P. radiata* cultures. Furthermore, in *CCoAOMT*-deficient *P. radiata*, a 20% reduction in lignin content was recorded due to G-lignin depletion and increased H-lignin levels (Huan et al. 2012). At present, most studies have focused on single genes, resulting in a lack of in-depth studies on the regulation of related large genes. Since forest tree species genomes are large and complex, and their genes interact with each other, further in-depth studies on the related genes of *Pinaceae* will help to reveal the regulatory mechanisms of lignin synthesis in this species.

CesA2 is related to cellulose synthesis

In plants, *CesA* belongs to a large gene family involved in cellulose synthesis in different organs, tissues, or cell walls (Huang et al. 2021). Different *CesA* genes not only have different expression levels (*Arabidopsis thaliana CesA1* levels > *A. thaliana CesA9* levels), expression sites (*A. thaliana CesA1* is expressed in all plant sections, while *A. thaliana CesA9* is expressed during embryo development), and expression patterns (*A. thaliana CesA7* is only expressed in xylem, while *A. thaliana CesA1* is expressed in every plant section), but are also affected by external regulatory factors (Somerville 2000, Beeckman 2002, Taylor et al. 1999). he research on CesA gene in higher plants has

never stopped. As early as the 1990s, Pear et al. (1996) not only successfully cloned the cotton CesA1 cDNA sequence (*Gossypium spp.*) but also analyzed its domains. Cellulose synthesis gene cloning in trees was initially conducted in *Populus tremuloides*. Previous studies showed that CesA1 in *P. tremuloides* was related to the secondary wall (Wu et al. 2000, Suzuki et al. 2006). The *CesA2* gene in *Populus tremuloides* has been shown to be associated with secondary walls, and it has also been shown to be involved in the formation of secondary walls in *Eucalyptus camaldulensis* Dehnh. (Lin et al. 2014). Both *CesA1* and *CesA2* genes were successfully cloned in Chinese fir and had key roles in fir growth and development (Pang 2015). In another study, the *CesA2* cDNA sequence was successfully cloned in *Betula platyphylla* Suk and that *CesA2* expression in leaves was higher when compared with the xylem (Chen 2008). In a *P. radiata* study, Krauskopf et al. (2005) not only confirmed that *CesA10* was related to cellulose synthesis but authors also isolated *P. radiata CESA2, CESA3, CESA5, CESA6, CESA7, CESA8* and *CESA11* cDNAs. In our study, while one *CesA* gene was associated with *P. armandii*, other family genes were not identified, therefore *CesA* functions in *P. armandii* require further elucidation.

CONCLUSIONS

Based on previous SNP markers and GWAS analyses, two genes were compared with respect to DBH traits: *CesA2* and *CCoAOMT* were related to cellulose and lignin synthesis, respectively. *P. armandii* SNP markers were genotyped and two (GG and GT) were found in the marker Marker227806, while average DBH data were analyzed using t tests. We observed that average DBH values corresponding to these were significantly different. Hence, this SNP marker could be used to select good monocultures. Finally, we used qPCR to analyze candidate gene expression in different *P. armandii* tissue. *CesA2* and *CCoAOMT* expression levels were highest in phloem tissue and lowest in bark, which further suggested that both genes may be related to DBH growth and metabolism.

ACKNOWLEDGEMENTS

The authors thank Southwest Research Center for Landscape Architecture Engineering of National Forestry and Grassland Administration and Key Laboratory of Forest Resources Conservation and Utilization in the Southwest Mountains of China Ministry of Education for facility and equipment use. Funding are from Yunnan Provincial Department of Science and Technology Science and Technology Plan Agriculture Joint Project (No. 202101BD070001-009) and Key R & D Project of Yunnan Provincial Department of Science and Technology (No. 2018BB005).

REFERENCES

Alexander DH, Novembre J & Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome research* 19: 1655–1664. https://doi.org/10.1101/gr.094052.109

Beeckman T, Przemeck GKH, Stamatiou G, Lau R, Terryn N, Rycke D Et Al. 2002. Genetic Complexity of Cellulose Synthase A Gene Function in *Arabidopsis* Embryogenesis. *Plant Physiology* 130: 1883–1893. https://doi.org/10.1104/pp.102.010603

Bradbury PJ, Zhang Z, Kroon DE, Casstevens DM, Ramdoss Y & Buckler ES. 2007. Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Bryan GT, Wu KS, Farall L, Jia Y, Hershey HP, Mcadams SA et al. 2000. A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene Pi-ta. *The Plant Cell* 12: 2033–2046. https://doi.org/10.1105/tpc.12.11.2033

Cao K, Li Y, Deng CH, Gardiner SE, Zhu G, Fang W et al. 2019. Comparative population genomics identified genomic regions and candidate genes associated with fruit domestication traits in peach. *Plant Biotechnology Journal* 17: 1954–1970. doi: 10.1111/pbi.13112

Cao K, Zhou Z, Wang QL, Guo J, Zhao P, Zhu G et al. 2016. Genome-wide association study of 12 agronomic traits in peach. *Nature Communications* 7: 13246.

Cappa EP, El-Kassaby YA, Garcia MN, Acuña C, Borralho NMG, Grattapaglia D et al. 2013. Impacts of Population Structure and Analytical Models in Genome-Wide Association Studies of Complex Traits in Forest Trees: A Case Study in *Eucalyptus globulus*. PLoS ONE 8: e81267. https://doi.org/10.1371/journal.pone.0081267

Chen BH. 2010. Cloning and Sequence Analysis of Cinnamoyl-CoA Reductase Gene (CCR) of *Pinus massoniana*. *Scientia Silvae Sinicae* 45: 46–53.

Chen GF, Chen JS & Tian JC. 2015. Genome-Wide Association Analysis between SNP Markers and Plant Height Related Traits in Wheat. *Acta Agronomica Sinica* 41: 1500–1509. doi: 10.3724/SP.J.1006.2015.01500

Chen PF. 2008. Cloning of cellulose synthase gene from *Betula platyphylla* and the analysis of their expression characteristic. MS thesis, Northeast Forestry University, Harbin.

Chen XY & Shen XH. 2005. *Forest tree breeding*. Higher Education Press, Beijing.

Chen YZ, Tan XF, Wang DB & Hu DB. 2002. Molecular assisted selection in forest tree breeding. Hunan Forestry Science & Technology 29: 17–20.

Craven-Bartle B, Pascual MB, Cánovas FM & Avila C. 2013. A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *The Plant Journal* 74: 755–766. https://doi.org/10.1111/tpj.12158

Cumbie WP, Eckert A, Wegrzyn J, Whetten R, Neale D & Goldfarb B. 2011. Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. Heredity 107: 105–114. https://doi.org/10.1038/hdy.2010.168

Duan ZQ & Zhu J. 2015. Research progress of genome-wide association study. *Journal of Zhejiang University* 41: 385–393. doi: 10.3785/j.issn.1008-9209.2015.03.243

Farfan IDB, Fuente GNDL, Murray SC, Isakeit T, Huang PC, Warburton M et al. 2015. Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. *Plos One* 10: e0117737. https://doi.org/10.1371/journal.pone.0117737

Hang JJ. 2016. Correlation analysis between SNP polymorphisms of wheat starch synthesis enzymes gene and starch content (*Triticum aestivum* L.). MAgr thesis, Shihezi University, Shihezi.

Hardy OJ & Vekemans X. 2002. Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular ecology notes* 2: 618–620. doi: 10.1046/J.1471-8286.2002.00305.X

He QL. 2019. Growth of *pinus armandii* in Longdonggou Region of Taibai Mountain. *Protect Forest Science and Technology* 11: 52–54. doi: 10.13601/j.issn.1005-5215.2019.11.018

Huan HV, Giang HV, Thanh NV, Zhang S, & Wang Y. 2012. Identification and functional analysis of the Pm4CL1 gene in transgenic tobacco plant as the basis for regulating lignin biosynthesis in forest trees. *Molecular Breeding 29*: 173–180. https://doi.org/10.1007/s11032-010-9535-9

Huang XZ, Jiang YX, Sun XQ, Fu ZY, Xu QJ. 2021. Research of *Cellulose Synthase* Gene in Natural Fiber Plants in China. *Heilongjiang Science* 12: 22–25.

Initiative Tag. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 408: 796-815. https://doi.org/10.1038/35048692

Japarudin Y, Meder R, Lapammu M, Waburton P, Paul-Macdonell P, Brown M et al. 2022. Developing *Eucalyptus Pellita* breeding populations for the solid wood industry of eastern malaysia. *Journal of Tropical Forest Science* 34: 347–358. doi: 10.26525/jtfs2022.34.3.346

Jia YY & Chen H. 2003. Molecular Genetic Marker of Single Nucleotide Polymorphism and its Application. *Journal of Yellow Cattle Science* 29: 42–45.

Krauskopf E, Harris PJ & Putterill J. 2005. The cellulose synthase gene PrCESA10 is involved in cellulose biosynthesis in developing tracheids of the gymnosperm *Pinus radiata*. *Gene* 350: 107–116. https://doi.org/10.1016/j.gene.2005.02.018

Li L, Osakabe Y, Joshi CP & Chiang VL. 1999. Secondary xylem-specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in *Loblolly pine*. *Plant Molecular Biology* 40: 555–565. https://doi.org/10.1023/A:1006244325250

Li SF, Su XH & Zhang BY. 2011. Research progress of forest tree gene cloning. *Chinese Bulletin of Botany* 46: 79–107. doi: 10.3724/SP.J.1259.2011.00079

Li ZF. 2007. Introduction to the conservation and utilization of *pinus armandii*. *Yunnan Agriculture* 07: 29.

Lin Y, Kao YY, Chen ZZ, Chu FH & Chung JD. 2014. cDNA cloning and molecular characterization of five cellulose synthase A genes from *Eucalyptus camaldulensis*. *Journal of Plant Biochemistry & Biotechnology* 23: 199–210. https://doi.org/10.1007/s13562-013-0202-1

Lipka AE, Gore MA, Magallanes-Lundback M, Mesberg A, Lin H, Tiede T et al. 2013. Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3 Genes Genomes Genetics* 3: 1287–1299. https://doi.org/10.1534/g3.113.006148

Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM et al. 2000. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* 288: 136–140. https://doi.org/10.1126/science.288.5463.136

Nguyen THH, Konda R, Kieu TD, Tran TC, Phung VK, Tran TH et al. 2019. Genetic improvement for wood production in *Melaleuca cajuputi*. *Journal of Tropical Forest Science* 31: 230–239. doi: 10.26525/JTFS2019.31.2.230239

Pang J, Tong ZK, Huang HH, Lin EP & Liu QY. 2015. Isolation and expression analysis of cellulose synthase genes in Chinese fir (*Cunninghamia lanceolata*). *Journal of Zhejiang A & F University* 032: 40–46. doi: 10.11833/j.issn.2095-0756.2015.01.006

Pear JR, Kawagoe Y, Schreckengost WE & Stalker DM. 1996. Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. *Proceedings of the National Academy of Sciences* 93: 12637–12642. https://doi.org/10.1073/pnas.93.22.12637

Porth I, Klapšte J, Skyba O, Hannemann J, Mckown AD, Guy Rd et al. 2013. Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. *New Phytologist* 200: 710–726. https://doi.org/10.1111/nph.12422

Qiao MX, Lin XF & Zhang WB. 2016. Isolation and Characterization of Caffeic Acid O-methyltransferase Gene from *Larix gmelinii*. *Molecular Plant Breeding* 14: 1684–1690. doi: 10.13271/j.mpb.014.001684

Resende RT, Resende MDV, Silva FF, Azevedo CF, Takahashi EK, Silva-Junior OB et al. 2016. Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytologist* 213:

1287–1300. https://doi.org/10.1111/nph.14266

Richmond TA & Somerville CR. 2000. The Cellulose Synthase Superfamily. *Plant Physiology* 124: 495–498. https://doi.org/10.1104/pp.124.2.495

Shao YL. 2015. Application Progress on molecular markers in forest tree genetics and breeding. *Agricultural Technology Service* 32: 122–123.

Suzuki S, Li L, Sun YH, Chiang VL & Chiang VL. 2006. The Cellulose Synthase Gene Superfamily and Biochemical Functions of Xylem-Specific Cellulose Synthase-Like Genes in *Populus trichocarpa*. *Plant Physiology* 142: 1233–1245. https://doi.org/10.1104/pp.106.086678

Taylor NG, Scheible WR, Cutler S, Somerville CR & Turner SR. 1999. The irregular xylem3 Locus of *Arabidopsis* Encodes a Cellulose Synthase Required for Secondary Cell Wall Synthesis. *Plant Cell* 11: 769–780. https://doi.org/10.1105/tpc.11.5.769

Umemototo T & Aoki N. 2005. Single-fiucleotid polymorphisms in rice starch synthase IIa that alter starch gelatinisation and starch association of the enzyme. *Functional Plant Biology* 32: 763–768. https://doi.org/10.1071/FP04214

Voo KS, Whetten RW, O'malley DM. & Sederoff RR. 1995. 4-Coumarate: Coenzyme a Ligase from *Loblolly Pine* Xylem: Isolation, Characterization, and Complementary DNA Cloning. *Plant Physiology* 108: 85–97. https://doi.org/10.1104/pp.108.1.85

Wagner A, Tobimatsu Y, Goeminne G, Phillips L, Flint H, Steward D Et Al. 2013. Suppression of CCR impacts metabolite profile and cell wall composition in *Pinus radiata* tracheary elements. *Plant Molecular Biology* 81: 105–117. https://doi.org/10.1007/s11103-012-9985-z

Wan ZB, Dai XG & Yin TM. 2012. Review on the Hot Topics of the Basic Studies for Forest Genetics and Breeding. *Scientia Silvae Sinicae* 48: 150–154.

Wang F, Ye P, Wang ZD, Xia MT, Shen WX, Wang HY Et Al. 2022. Development and analysis of high flux genome wide SNP markers for armand pine (*Pinus armandii*). *Molecular Plant Breeding* 20: 5733-5740. doi: 10.13271/j.mpb.020.005733

Wu L, Joshi CP & Chiang VL. 2000. A xylem-specific cellulose synthase gene from aspen *(Populus tremuloides)* is responsive to mechanical stress. *The Plant Journal* 22*:* 495–502. https://doi.org/10.1046/j.1365-313x.2000.00758.x

Xiao N, Wu Y, Pan C,Yu L, Chen Y, Liu GQ et al. 2016. Improving of rice blast resistances in japonica by pyramiding major R genes. *Frontiers in Plant Science* 7: 1918. doi: 10.3389/fpls.2016.01918

Xin PY, Cao ZY, Wang ZD, Ye PENG, Liu C, Wang F, Xin J, Dong ZH, Qu Shet AL. 2021 A development method and application of genome SNP in *Pinus armandii*. Patent, Southwest Forestry University. Yunnan.

Xin PY, Zhou J, Duan AA, Sun ZZ & He CZ. 2010. Research Progress on Genetic Improvement of *Pinus armandii* in China. *Northern Horticulture* 19: 218–222.

Xu FQ, Wang Q, Lyu X & Liu WZ. 2014. Effects of Tending Thinning on the Growth of *Pinus armandii* Plantations in Xiaolongshan Forest Region. *Journal Of Northwest Forestry University* 6: 176–179. doi: 10.3969/j.issn.1001-7461.2014.06.33

Yuan HN & Chen WH. 2010. Status of exploitation and conservation measures of *pinus armandii*. *Forestry of Shaanxi* 04: 39.

Zaid IU, Tang W, Liu E, Khan SU, Wang H, Mawuli EW et al. 2017. Genome-wide single-nucleotide polymorphisms in cms and restorer lines discovered by genotyping using sequencing and association with marker-combining ability for 12 yield-related traits in *Oryza sativa* L. Subsp. Japonica. *Frontiers in Plant Science* 8: 143. doi: 10.3389/fpls.2017.00143

Zhang FK. 2014. Cloning and analyzing of CAD and CCoAOMT genes from *Pinus massoniana*. MAgr thesis, Nanjing Forestry University, Nanjing.

Zhang YY & Cai HQ. 1989. Geographical distribution and community ecological characteristics of *Pinus armandii*. *Hunan Forestry Science & Technology* 4: 5–8.

Zhang ZY. 2017. Genome-wide association analysis with significant characters and integration on resistance to *Leptocybe invasa* in *Eucalyptus grandis*. DAgr thesis, Chinese Academy of Forestry, Beijing.