

Transcriptome and gene expression analysis revealed mechanisms for producing high oleoresin yields from Simao pine (*Pinus kesiya* var. *langbianensis*)

Wang Yi, Yuan Xiaolong, Hua Mei, Li Jiang, Wang Juan*

Key Laboratory of Forest Plant Cultivation and Utilization, Yunnan Academy of Forestry, Kunming Yunnan 650204, China

*Corresponding author: schima@163.com

Abstract

Simao pine (*Pinus kesiya* Royle ex Gordon var. *langbianensis* (A. Chev.) Gaussen) is an important oleoresin source in Yunnan Province, China. However, the molecular mechanisms of high oleoresin yield from Simao pine remain unclear. In this study, RNA-Seq was used to investigate the transcriptome of the species and compare the gene expression profiles of wounded bark from high and low oleoresin-yielding individuals, and the expression levels of genes in the methyl-erythritol 4-phosphate (MEP) pathway were detected by qPCR. A total of 1.41 billion reads were obtained and assembled into 68,881 unigenes from samples of the wounded bark of Simao pine. The KEGG analysis of differentially expressed genes showed that MEP pathway genes were obviously differentially expressed, while most genes in the mevalonate (MVA) pathway were not. The qPCR analysis showed that the MEP pathway gene expression of wounded bark from high oleoresin-yielding genotypes was higher than that of low oleoresin-yielding genotypes. The gene expression differences of 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR), and geranylgeranyl pyrophosphate synthase (GGPPs) between high and low oleoresin-yielding genotypes were significant. This implies that DXS, HDR, and GGPPs play important roles in high oleoresin production in Simao pine.

Keywords: *Pinus kesiya* var. *langbianensis*; high yield; oleoresin; transcriptome analysis; gene expression

Abbreviations: MEP_methyl-erythritol 4-phosphate; MVA_mevalonate; DXS_1-deoxy-D-xylulose-5-phosphate synthase; HDR_4-hydroxy-3-methylbut-2-enyl diphosphate reductase; GGPPs_Geranylgeranyl pyrophosphate synthase; DXR_1-deoxy-D-xylulose-5-phosphate reductoisomerase; CMS_4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase; CMK_4-diphosphocytidyl-2C-methyl-D-erythritol kinase; MCS_2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS_4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; IDI_isopentenyl diphosphate isomerase; GGPPS_geranylgeranyl diphosphate synthase; GPPS_geranyl diphosphate synthase; TPS_terpenoid synthase.

Introduction

Simao pine (*Pinus kesiya* Royle ex Gordon var. *langbianensis* (A. Chev.) Gaussen) is an important oleoresin source in Yunnan Province, China. The annual output of pine oleoresin from Simao pine is 179,100,000 kg (Dong et al., 2009), and 90% of gum turpentine output in Yunnan comes from Simao pine (Yin et al., 2005). The forestland area of Simao pine is 0.56 million hm^2 (Li et al., 2009). However, individual variation in oleoresin yield is rather higher, ranging from 3 kg to 140 kg annually (Xu et al., 2012). Variation in oleoresin yield was also found in *Pinus massoniana* and *Pinus elliottii* (Liu et al., 2013; Zhang et al., 2016). Molecular breeding can be used to increase oleoresin yields of Simao pine. However, the physiological mechanism of achieving high oleoresin yields in Simao pine remains unclear. Oleoresin is an important non-wood forestry product, and its derivatives have applications across different industries, including pharmaceutical, cosmetic, and food industries as well as the chemical industry in the manufacturing of various products, such as paint, varnishes, adhesives, insecticides, and disinfectants (Rodrigues-Corrêa et al., 2012; Silva et al., 2012; Rodrigues-Corrêa et al., 2013). Oleoresin may also become a competitive liquid biofuel

(Lieutier et al., 2004; Meylemans et al., 2012). In addition to its many important applications, oleoresin production is part of the conifer's defense systems against physical wounding, bark-feeding insects, and pathogenic fungi (Trapp and Croteau, 2001; Kolosova and Bohlmann, 2012). Pine oleoresin has typically been obtained from living trees by the repeated wounding or barks chipping methods (Rodrigues-Corrêa et al., 2013). As is well-known, pine oleoresin is synthesized by two separate pathways in conifers: the methyl-erythritol 4-phosphate (MEP) pathway and the mevalonate (MVA) pathway. Monoterpenes and diterpenes are biosynthesized by the MEP pathway, while the biosynthesis of sesquiterpenes is completed by the MVA pathway (Vranová et al., 2013). Monoterpenes and diterpenes are the main components of oleoresin in Simao pine (Song et al., 1995). Therefore, the MEP pathway plays an important role in oleoresin biosynthesis in Simao pine. The enzymes in the MEP pathway include 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (CMS), 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (CMK), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MCS), 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (HDS), and

4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR; Fig 1). The MEP pathway within plastids produces isopentenyl diphosphate like the MAV pathway. The isopentenyl diphosphate from the MEP pathway was transformed into different terpenes by a series of post-modification enzymes, such as isopentenyl diphosphate isomerase (IDI), geranylgeranyl diphosphate synthase (GGPPS), geranyl diphosphate synthase (GPPS), terpenoid synthase (TPS), and cytochrome P450-dependent monooxygenase. Although several papers have reported the roles of DXS, DXR, HDR, and TPS in oleoresin biosynthesis (Phillips et al., 2003; Byun-Mckay et al., 2006; Kim et al., 2009; Liu et al., 2015), there have been no reports about the influence of the MEP pathway and expression the MEP enzyme on high and low oleoresin yields in Simao pine. In this study, the transcriptomes of high and low oleoresin-yielding individuals of Simao pine were analyzed, and the genes involved in the MEP pathway were identified and their functions were characterized.

Results

Sequence analysis and assembly of wounded bark transcriptomes

To examine the molecular basis of oleoresin biosynthesis in Simao pine, wounded bark from high and low oleoresin-yielding Simao pines were used to construct two libraries for high-throughput sequencing using an Illumina HiSeq2000 sequencer (Illumina, Inc., San Diego, CA, USA). The high and low yield libraries produced 72.57 million and 68.65 million reads of raw data, respectively, as paired-end reads with single read lengths of ~126 bp. After removal of adaptor sequences, ambiguous reads, and low-quality reads (Q20<20), 6,169,184,280 nucleotides (7.25 Gb) and 5,822,084,160 nucleotides (7.43 Gb) of high-quality clean reads were obtained from the high and low yield libraries, respectively. After all high-quality reads were assembled, 68,881 unigenes were obtained from wounded bark with high and low oleoresin yields, and the N50 was 1,402

Functional annotation of transcriptome

Unigenes from transcriptome were annotated using several databases: National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>), Ensembl plant (<http://plants.ensembl.org/index.html>), KEGG (<http://www.genome.jp/kegg/>), and eggNO (<http://eggnogetdb.embl.de/>). Unigene sequences were first aligned to protein databases like NR, Swiss-Prot, KEGG, and COG (E-value < 0.00001) by blastx, and nucleotide database NT (E-value < 0.00001) by blastn, retrieving proteins with the highest sequence similarity with the given unigenes along with their protein function annotations. In total, 48,035 unigenes were annotated using all databases, and only 13,528 unigenes were annotated in Cluster of Orthologous Groups (COG) of proteins databases. These unigenes were distributed among 25 functional categories (Fig 3).

Transcriptome-based analysis of differentially expressed genes

A rigorous algorithm was used to identify genes that were differentially expressed between high and low oleoresin-yielding Simao pines. The genes were classified into one of three classes. Red genes were up-regulated such that gene expression in the low-yield individual exceeded that of the high-yield individual. Green genes were down-regulated such that gene expression in the high-yield individual exceeded that of the low-yield individual. Blue genes were not differentially expressed (Fig 4).

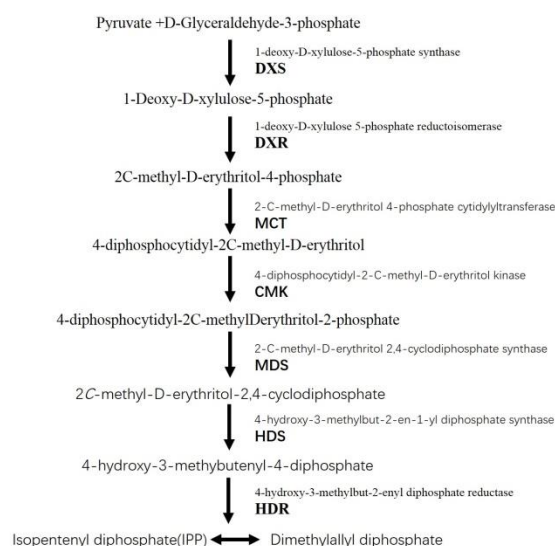
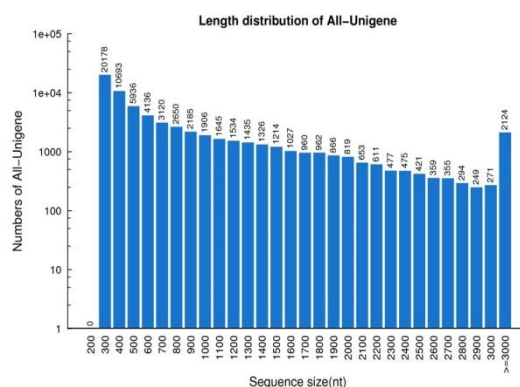
Differentially expressed genes were also analyzed for pathway enrichment. The main biochemical pathways and signal transduction pathways were analyzed to determine which pathways differentially expressed genes take part in. The terpenoid backbone biosynthesis pathway of differentially expressed genes showed that all genes involved in the MEP pathway are differentially expressed (Fig 5). According to the pathway enrichment analysis of differentially expressed genes in the MEP pathway, the expression of genes in the high-yield individual was higher than their expression in the low-yield individual. In contrast, most genes in the MVA pathway were not differentially expressed; only acetyl-CoA C-acetyltransferase and hydroxymethylglutaryl-CoA reductase genes exhibited obvious differential expression between high and low oleoresin-yielding bark samples.

Gene expression profile of the MEP pathway

To confirm the transcriptome analysis of differentially expressed genes, gene expression profiles of nine genes involved in MEP pathway were assayed by qRT-PCR (Fig 6). The qRT-PCR analysis showed that nine genes (DXS, DXR, MCT, CMK, MDS, HDS, HDR, APS, and GGPPs) were strongly expressed in wounded bark from high oleoresin-yielding Simao pines. Moreover, these genes were more highly expressed in high oleoresin-yielding individuals relative to low oleoresin-yielding individuals. In particular, gene expression of GGPPs in high oleoresin-yielding individuals was 24 times that of low oleoresin-yielding individuals. Geranylgeranyl diphosphate synthase (GGPPs) catalyzes the biosynthesis of geranylgeranyl diphosphate, a key precursor for diterpene, tetraterpene, and poly-terpene (Liang et al., 2015). GGPPs play important role in diterpenes biosynthesis and rosin (diterpenes) is main component of oleoresin. This means that GGPPs play an important role in oleoresin biosynthesis of high-yielding individuals in Simao pine. 1-Deoxy-D-xylulose-5-phosphate synthase (DXS) catalyzes the first committed step of the 2C-methyl-D-erythritol-4-phosphate (MEP) pathway, DXS is the first key enzyme in the MEP pathway (Kirby et al., 2016). And 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR) catalyzes the last step of the MEP pathway to synthesize isopentenyl diphosphate and its allyl isomer dimethylallyl diphosphate, which are common precursors for the synthesis of plastid isoprenoids (Lv et al., 2015). Oleoresin including turpentine (monoterpenes and sesquiterpenes) and rosin (diterpenes) is mainly synthesized by MEP pathway. The level of gene expression of DXS and HDR in high and low

Table 1. Primers used in qRT-PCR.

Primer name	Sequence (5'-3')	Detected gene name
tDXSF	AGAAGTATTGTGAATCCATA	1-deoxy-D-xylulose-5-phosphate synthase
tDXSR	GTATCAGTAGTAGTCATTGT	
tDXRF	GAAGAGCAGTTGTAGATGAAG	1-deoxy-D-xylulose 5-phosphate reductoisomerase
tDXRR	CCGATTGAGCCTGTAGAA	
tMCTF	CAGTATCTTCCATTATTAGG	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
tMCTR	TATACCGTCACACTATTC	
tCMKF	CCTGATTCTGTTGGTGTA	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
tCMKR	ATTACTTGCTGGTGCTTA	
tMDSF	GGAAGAGGCAACATAACG	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
tMDSR	AAGAAGTGAGGAACAAGGA	
tHDSF	AGAAGTATTGTGAATCCATA	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
tHDSR	GTATCAGTAGTAGTCATTGT	
tHDRF	GCAGTTGTTAAGTGAGAA	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
tHDRT	ATGAATGATGGAGGTGTA	
tAPSF	TTCGTCCGTATCGTGTTA	alpha-pinene synthase
tAPSR	AGTTCCTTGGTTAAGTTATCG	
tGGPPsF	CACCTTGCTCAGACATCTT	Geranylgeranyl pyrophosphate synthase
tGGPPSR	CTATCCTCAACCGTTCAG	
EF1a-F	CAACAGACTTAAGTTCAG	elongation factor 1-alpha
EF1a-R	TTACAAGATTGGTGGTAT	

**Fig 1.** MEP pathway in *P. kesiya* var. *langbianensis* adapted from Liu et al., (2015).**Fig 2.** The length distribution of Unigenes, the horizontal coordinates are Unigene lengths and the vertical coordinates are numbers of Unigenes.

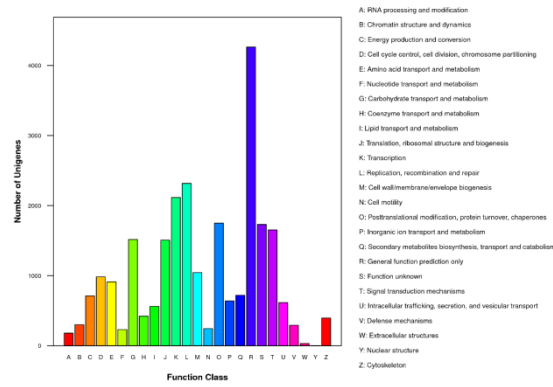


Fig 3. COG function classification of Unigenes in All-Unigene. The horizontal coordinates are function classes of cog, and the vertical coordinates are numbers of Unigenes in one class. The notation on the right is the Full name of the functions in X-axis.

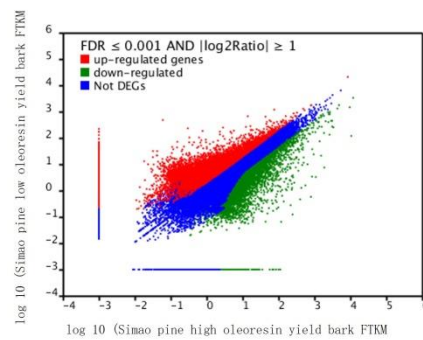


Fig 4 Expression level of high vs low oleoresin yield bark of *P. kesiya* var. *langbianensis*. The genes were classified into three classes. Red genes are up-regulated that gene expression of high oleoresin yield genotype is larger than low oleoresin yield genotype. Green genes are down-regulated that gene expression of low oleoresin yield genotype is larger than high oleoresin yield genotype. Blue genes are not differentially expressed genes. The horizontal coordinates are the expression level of high oleoresin yield genotype and the vertical coordinates is the expression level of low oleoresin yield genotype.

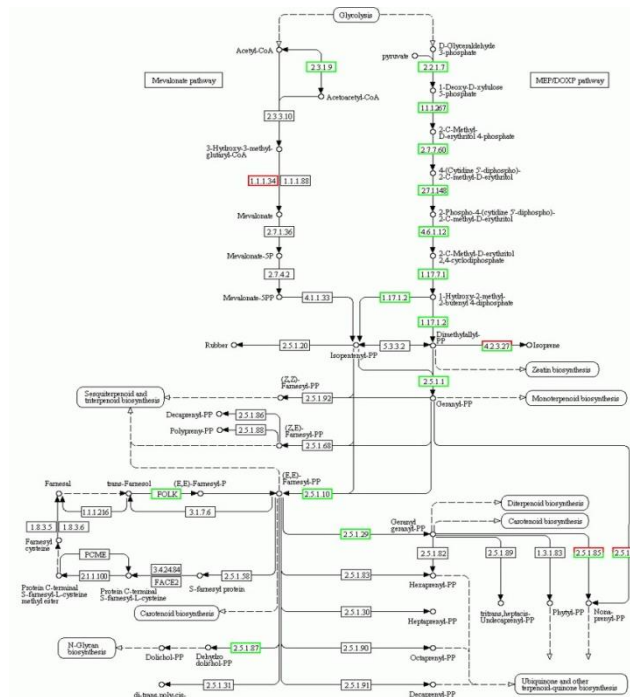


Fig 5. Differentially expressed genes significantly enriched in terpenoid backbone biosynthesis. Green box showed the gene expression in high oleoresin yield genotype was higher than in low oleoresin yield genotype. Red box showed the gene expression in high oleoresin yield genotype was lower than in low oleoresin yield genotype.

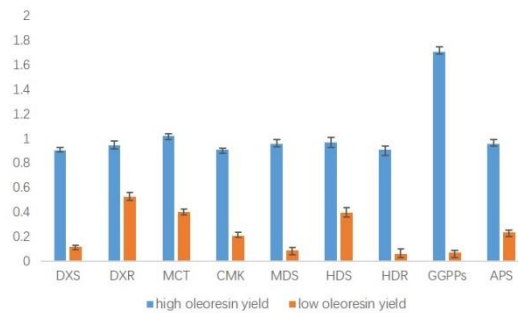


Fig 6. qRT-PCR analysis of MEP pathway-related candidate unigenes in high and low oleoresin yield of *P. kesiya* var. *langbianensis*. DXS: 1-deoxy-D-xylulose-5-phosphate synthase, DXR: 1-deoxy-D-xylulose 5-phosphate reductoisomerase, MCT: 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase, CMK: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, MDS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, HDS: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, GGPPs: Geranylgeranyl pyrophosphate synthase, APS: (+)-alpha-pinene synthase.

oleoresin-yielding individual is stronger than in low oleoresin yield individual. This suggests that DXS and HDR also play important roles in oleoresin biosynthesis of high oleoresin-yielding Simao pine.

Discussion

RNA-Seq is a straightforward and efficient method for obtaining genome-wide information on gene expression, gene regulation, and amino acid content of proteins (Wang et al., 2009; Ozsolak and Milos, 2011). Thus, a comparison of traditional expression assay methods with RNA sequencing provides comprehensive information revealing plant metabolites and physiological processes. In non-model organism's studies in particular, *de novo* transcriptome sequencing is powerful and cost-effective. Many metabolic pathways, physiological processes, and growth and development mechanisms of non-model plant have been revealed by *de novo* transcriptome sequencing (Wu et al., 2014; Lang et al., 2015). The analysis of differentially expressed genes based on transcriptomic data from several samples is a powerful tool for revealing the biological mechanisms underlying differences among individuals (Lou et al., 2014; Liu et al., 2015; Shi et al., 2015). In particular, differential gene expression analyses showed that GGPS and (-)-alpha/beta-pinene synthase were up-regulated in the high oleoresin-yielding genotype of *Pinus massoniana* (Liu et al., 2015). In this study, the differential gene expression analysis showed that nine genes in the MEP pathway were up-regulated in the high oleoresin-yielding Simao pine individual. The gene expression profile also supported the result of the differential gene expression analysis. These findings strongly suggest that the stimulation mechanisms associated with physical wounding differs between high and low oleoresin-yielding genotypes.

So far, several techniques have been developed to improve oleoresin yields of conifers, including different tapping techniques, chemical and physical induction methods, and the selection of the optimal season for tapping (Wang et al., 2006; Rodríguez-García et al., 2015; Füller et al., 2016; Rodríguez-García et al., 2016). Recently, research has focused on the molecular mechanism of oleoresin yields and attempts have been made to develop genetic markers for molecular breeding (Liu et al., 2013; Westbrook et al., 2013; Westbrook et al., 2015; Cai et al., 2017). However, there is no effective

genetic marker for breeding high oleoresin-yielding pines. As is well known, oleoresin yield is a quantitative trait, and a major current challenge in biology understands the genetic basis of variation in quantitative traits (Mackay et al., 2009). Quantitative trait loci (QTLs) are genetic loci for which functionally different alleles segregate and have significant effects on quantitative traits. Comparing traditional molecular marker development methods, candidate gene validation is likely to accelerate the pace of discovering the genes underlying QTLs (Salvi and Tuberosa, 2005). Accordingly, this is the first key step in discovering high oleoresin-yielding QTLs in order to find candidate genes related to high oleoresin-yielding genotypes. In this study, nine candidate gene related to high oleoresin yields in Simao pine were confirmed by differential gene expression analyses based on transcriptomes and real time PCR. This is also the first report demonstrating that all genes in the MEP pathway are upregulated in a high oleoresin-yielding Simao pine genotype. In particular, the gene expression differences in DXS, HDR, and GGPPs in high and low oleoresin-yielding genotypes were significant. In future work, the sequence variation of these candidate genes at different genotypes will be focused on, and new molecular markers for high oleoresin-yielding Simao pines will be developed based on sequence variation of these candidate genes.

Materials and methods

Plant materials

A population of half-sib progeny of high-resin-yielding *Pinus kesiya* var. *langbianensis* was constructed in the forest farm of BanPo Xiang, which is located in Jing Gu County, Puer, Yunnan Province, China (Li et al., 2008). A seed from a superior high-resin-yielding Simao pine was planted into flat ground composed of latosolic red soil in 2002. Another Simao pine seed from a normal resin-yielding Simao pine was also planted in the same environment. The resin-producing capacity was detected using the V-drop crossing method in Yunnan Province, China. The V-drop crossing method is a traditional tap resin collection method that involves wounding bark to induce resin production from Simao pine. Collected resin from all samples was weighed, and the resin weight determined the resin-producing capacity (Miao et al., 2016). The resin-producing capacities were detected in 2002 and 2016, respectively. After the high and low resin-yielding individuals

were confirmed, RNA extraction was conducted using bark samples from 14-year-old Simao pines; samples were frozen in liquid nitrogen immediately in July 2016.

RNA extraction, library construction, and RNA-Seq

The bark samples from high and low oleoresin-yielding individuals were ground into powder in liquid nitrogen. Total RNA was isolated using the RNeasy Plant Kit (Qiagen, Hilden, Germany). The concentration and quality was examined using a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) and 0.8% agarose gel electrophoresis. The construction of the libraries and the RNA-Seq were performed by the Beijing Genomics Institute (BGI), Beijing, China. The TruSeq RNA Sample Prep Kit (Illumina) was used for purification and fragmentation of mRNA. The first-strand cDNA was synthesized by reverse transcriptase using the cleaved RNA as a template, and the second-strand cDNA was generated using DNA polymerase I. The ends of the DNA fragments were modified and ligated with adapter sequences, and the cleaned ligation products were used as PCR templates for amplifying the fragments. cDNA libraries were obtained, and the quality was examined with a PicoGreen assay kit by Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The constructed cDNA libraries were sequenced on the Illumina HiSeq2000 instrument by BGI.

De novo transcriptome assembly and annotation

Following the sequencing of the cDNA library, the high-quality clean reads were generated by trimming the raw reads to remove adapter sequences, low-quality reads with Q-values < 20 or ambiguous bases ('N'). The clean reads were then assembled *de novo* using the Trinity platform (<http://trinityrnaseq.sf.net>). All unigenes were used as queries in searching Nr, eggNOG, GO, KO, and Swissprot databases (E-value < 10^{-5}) and functionally annotated by GO analysis with Blast2GO software (E-value < 10^{-5}). Metabolic pathways were predicted by KEGG mapping.

GO Classification of Differentially Expressed Unigene and Pathway Analysis

To compare the expression abundance between the high and low oleoresin-yielding individuals, unigene expression was calculated using the fragments per kb per million reads (FPKM) method. All high-quality reads were aligned to the assembled transcripts, and their tags were normalized into reads per kilobase of transcript per million mapped reads (RPKM) values. The differentially expressed genes between the two samples were identified by a rigorous algorithm based on the Audic and Claverie (1997) method. All differentially expressed genes were mapped to each term of the Gene Ontology (GO) database (<http://www.geneontology.org/>), and the number of genes associated with each GO term was determined. The names and numbers of genes for each GO term were listed, and then the hypergeometric test was used to find significantly enriched GO terms in differentially expressed genes compared to the genomic background. All differentially expressed genes were also mapped onto the KEGG pathway, with Q-value ≤ 0.05 considered to be significantly enriched in differentially expressed genes.

Real-time quantitative PCR analysis (RT-PCR)

Nine genes involved in MEP pathways were identified from the transcriptome of Simao pine databases. They included 1-deoxy-D-xylulose-5-phosphate synthase (DXS, accession number AY22671.1), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR, accession number: MG764427), 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT, accession number: MG764429), 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK, accession number: MG764426), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS, accession number: MG764425), 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (HDS, accession number: MG764428), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR, accession number: AY22672.1), geranylgeranyl pyrophosphate synthase (GGPPs, accession number : MG764424), and (+)-alpha-pinene synthase (APS, accession number: AY22674.1). The expression profiles were assayed using real time PCR. First, special serial primers were designed for the nine genes (Table 1). cDNA synthesis was performed with 1 μ g of total RNA from wounded bark of high and low oleoresin-yielding individuals using a reverse transcriptase kit (TaKaRa Super RT Kit; Dalian, China) according to the manufacturer's instruction. SYBR Green was used for detecting PCR products using the StepOnePlus real-time PCR system (Applied Biosystems, Foster City, CA, USA). The elongation factor 1-alpha (EF1a) gene was used as the internal control for normalization of gene expression. At least two independent biological replicates and three technical replicates of each biological replicate for each sample were analyzed by q-PCR to ensure reproducibility and reliability.

Conclusion

In this study, a comprehensive transcriptomic analysis of wounded bark samples from Simao pine was conducted, which provides a valuable resource for genetic and genomic study in the future. Differentially expressed gene analyses of transcriptomes and qPCR analyses of wounded bark from high and low oleoresin-yielding individuals showed that the genes in the MEP pathway were obviously differentially expressed. The gene expression differences of DXS, HDR, and GGPPs in high and low oleoresin-yielding genotypes were especially significant. This implies that DXS, HDR, and GGPP play important roles in high oleoresin-yielding Simao pines. Overall, our work provides new information for the molecular mechanisms of high oleoresin yields in Simao pine, which will provide a solid foundation for molecular breeding of this species.

Acknowledgments

This work was supported by a grant from the project of State Special Research Funds of Forestry Sector for Public Welfare (201304105) and Yunnan Academy of Forestry Innovation Fund Project(QN2018-01).

References

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. *Genome Res.* 7(10):986-995.

- Byun-Mckay A, Godard KA, Toudefallah M, Martin DM, Alfaro R, King J, Bohlmann J, Plant AL (2006) Wound-induced terpene synthase gene expression in *Sitka spruce* that exhibit resistance or susceptibility to attack by the white pine weevil. *Plant Physiol.* 140(3):1009-1021.
- Cai NH, Xu YL, Wang DW, Chen S, Li GQ (2017) Identification and characterization of microsatellite markers in *Pinus kesiya* var. *langbianensis* (Pinaceae). *Appl Plant Sci.* 5(2):1600126.
- Dong JX, Guo HJ, Li P, Zhao YF, Yun XH (2009) Oleoresin resources and development capacity in Yunnan. *J Northwest For Univ.* 24(1):157-160.
- Füller TN, de Lima JC, de Costa F, Rodrigues-Corrêa KC, Fett-Neto AG (2016) Stimulant paste preparation and bark streak tapping technique for pine oleoresin extraction. *Methods Mol Biol.* 1405:19-26.
- Kirby J, Dietzel KL, Wichmann G, Chan R, Antipov E, Moss N, Baidoo EEK, Jackson P, Gaucher SP, Gottlieb S, LaBarge J, Mahatdejkul T, Hawkins KM, Muley S, Newman JD, Liu P, Keasling JD, Zhao L (2016) Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in *Saccharomyces cerevisiae*. *Metab Eng.* 38:494-503.
- Kim YB, Kim SM, Kang MK, Kuzuyanma T, Lee JK, Park SC, Shin SC, Kim SU (2009) Regulation of resin acid synthesis in *Pinus densiflora* by differential transcription of genes encoding multiple 1-deoxy-D-xylulose 5-phosphate synthase and 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase genes. *Tree Physiol.* 29(5):737-749.
- Kolosova N, Bohlmann J (2012) Conifer defense against insects and fungal pathogens. In: Matyssek R, Sxhnyder H, Obwald W, Ernst D, Munch JC, Pretzsch H (ed) *Growth and defence in plants*. Springer Berlin Heidelberg. 85-109.
- Lang V, Usadel B, Obermeyer G (2015) De novo sequencing and analysis of the lily pollen transcriptome: an open access data source for an orphan plant species. *Plant Mol Biol.* 87(1-2):69-80.
- Li SG, Fu YP, Zhang KF, Jiang YD, Yao ZQ, Li M (2008) Halfsib progeny tests of high-resin-yielding *Pinus kesiya* var. *langbianensis*. *J Zhejiang For College.* 25(2):158-162.
- Li SG, Fu YP, Zhang KF, Zhao YH, Jiang YD, Yao ZQ, Li M (2009) Superior clone selection of *Pinus kesiya* var. *langbianensis* with high content of resin. *J Northeast For Univ.* 37(3):4-5.
- Liang MH, Liang YJ, Jin HH, Jiang JG (2015) Characterization and functional identification of a gene encoding geranylgeranyl diphosphate synthase from *Dunaliella bardawil*. *J Agr Food Chem.* 63(35):7805-7812.
- Lieutier F, Day KR, Battisti A, Grégoire JC, Evans HF (2004) *Bark and wood boring insects in living trees in Europe, a synthesis*. 1st ed. Springer, Netherlands.
- Liu Q, Zhou Z, Fan H, Liu Y (2013) Genetic variation and correlation among resin yield, growth, and morphologic traits of *Pinus massoniana*. *Silvae Genet.* 62(1):38-44.
- Liu Q, Zhou Z, Wei Y, Shen D, Feng Z, Hong S (2015) Genome-wide identification of differentially expressed genes associated with the high yielding of oleoresin in secondary *Xylem* of masson pine (*Pinus massoniana* Lamb) by transcriptomic analysis. *PLoS one.* 10(7):e0132624.
- Lv H, Zhang X, Liao B, Liu W, He L, Song J, Sun C, Luo H, Chen S (2015) Cloning and analysis of 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate reductase genes HsHDR1 and HsHDR2 in *Huperzia serrate*. *Acta Pharm Sin B.* 5(6):583-589.
- Lou Q, Liu Y, Qi Y, Jiao S, Tian F, Jiang L, Wang Y (2014) Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J Exp Bot.* 65(12):3157-3164.
- Mackay TF, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet.* 10(8):565-577.
- Meylemans HA, Quintana RL, Harvey BG (2012) Efficient conversion of pure and mixed terpene feedstocks to high density fuels. *Fuel.* 97(7):560-568.
- Miao FJ, Li J, Xiong Z, Yuan XL, Yang YM, Wang J (2016) Analysis on resin-producing capacity of *Pinus kesiya*. *J Fujian Agr For Univ (Nat Sci Ed).* 45(4):405-408.
- Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet.* 12(2):87-98.
- Phillips MA, Wildung MR, Williams DC, Hyatt DC, Croteau R (2003) cDNA isolation, functional expression, and characterization of (+)-alpha-pinene synthase and (-)-alpha-pinene synthase from loblolly pine (*Pinus taeda*): stereocontrol in pinene biosynthesis. *Arch Biochem Biophys.* 411(2):267-276.
- Rodríguez-García A, Martín JA, López R, Mutke S, Pinillos F, Gil L (2015) Influence of climate variables on resin yield and secretory structures in tapped *Pinus pinaster* Ait. in central Spain. *Agr Forest Meteorol.* 202:83-93.
- Rodríguez-García A, Martín JA, López R, Sanz A, Gil L (2016) Effect of four tapping methods on anatomical traits and resin yield in maritime pine (*Pinus pinaster* Ait.). *Ind Crop Prod.* 86:143-154.
- Rodrigues-Corrêa KCDS, Lima JCD, Fett-Neto AG (2012) Pine oleoresin: tapping green chemicals, biofuels, food protection, and carbon sequestration from multipurpose trees. *Food Energy Secur.* 1(2):81-93.
- Rodrigues-Corrêa KCDS, Lima JCD, Fett-Neto AG (2013) Oleoresins from pine: production and industrial uses. *Nat Prod.* 4037-4060.
- Silva ACRD, Lopes PM, Azevedo MMBD, Costa DCM, Alviano CS, Alviano DS (2012) Biological activities of α -pinene and β -pinene enantiomers. *Molecules.* 17(6):6305-6316.
- Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10(6):297-304.
- Shi Q, Zhou L, Wang Y, Li K, Zheng B, Miao K (2015) Transcriptomic analysis of *Paeonia delavayi* wild population flowers to identify differentially expressed genes involved in purple-red and yellow petal pigmentation. *PLoS one.* 10(8):e0135038.
- Song Z, Liang Z, Liu X (1995) Chemical characteristics of oleoresins from Chinese pine species. *Biochem Syst Ecol.* 23(5):517-522.
- Trapp S, Croteau R (2001) Defensive resin biosynthesis in conifers. *Annu Rev Plant Biol.* 52(1):689-724.
- Vranová E, Coman D, Grissem W (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Ann Rev Plant Biol.* 64(1):665-700.
- Wang Z, Calderon MM, Carandang MG (2006) Effects of resin tapping on optimal rotation age of pine plantation. *J Forest Econ.* 11(4):245-260.
- Wang Z, Gerstein M, Snyder M (2009) RNA-seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 10(1):57-63.
- Westbrook JW, Resende MFR, Munoz P, Walker AR, Wegrzyn JL, Nelson CD, Neale DB, Kirst M, Huber DA, Gezan SA, Peter GF, Davis JM (2013) Association genetics of oleoresin flow in

- loblolly pine: discovering genes and predicting phenotype for improved resistance to bark beetles and bioenergy potential. *New Phytol.* 199(1):89-100.
- Westbrook JW, Walker AR, Neves LG, Munoz P, Resende MFR, Neale DB, Wegrzyn JL, Huber DA, Kirst M, Davis JM (2015) Discovering candidate genes that regulate resin canal number in *Pinus taeda* stems by integrating genetic analysis across environments, ages, and populations. *New Phytol.* 205(2):627-641.
- Wu T, Luo S, Wang R, Zhong Y, Xu X, Lin Y, He X, Sun B, Huang H (2014) The first illumina-based de novo transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Mol Breeding.* 34(3):1437-1447.
- Xu MY, Deng GX, Ling WG (2012) Exploration on the utilization of *Pinus kesiya* seeds. *Seed.* 31(8):95-96.
- Yin XB, Geng SX, Ma HF, Zheng W (2005) Differences of physical and chemical characteristics of gum turpentine from simao pine. *J Nanjing For Univ (Nat Sci Ed).* 29(5):80-84.
- Zhang S, Jiang J, Luan Q (2016) Genetic and correlation analysis of oleoresin chemical components in slash pine. *Genet Mol Res.* 15(3):1-12.