Industrial Crops & Products 170 (2021) 113827

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

Identification of the diterpenoid biosynthesis genes and their expression status in relation to oleoresin yield of masson pine

Lina Mei [a,b], Youjin Yan [b], Zhengchun Li [a,b], Jiaxin Ran [a], Luonan Shen [a], Rongju Wu [a], Qiandong Hou [a], Tianjiao Shen [a], Xiaopeng Wen [a,b,*], Zhangqi Yang [c], Yuanheng Feng [c]

[a] Key Laboratory of Plant Resources Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Institute of Agro-bioengineering, Guizhou University, Guiyang, 550025, China
[b] Institute for Forest Resources & Environment of Guizhou/ College of Forestry, Guizhou University, Guiyang, 550025, China
[c] Guangxi Academy of Forestry, Nanning, Guangxi, 530000, China

ARTICLE INFO

Keywords:
Masson pine (*Pinus massoniana* L.)
Transcriptome
GC–MS
Diterpenoid biosynthesis
High-yield oleoresin germplasm
Molecular-assisted selection

ABSTRACT

Conifer oleoresin is a highly precious raw material for industrial sectors and a precursor for biofuels. Diterpenoid, particularly diterpenoid resin acids (DRAs), is one of the most abundant oleoresin components. Masson pine (*Pinus massoniana* L.) is one of the most important native species for oleoresin yield in China. To better understand the biosynthesis of diterpenoid in high oleoresin-yield masson pine and develop methods for its production in needle and trunk xylem, we used the combination of SMRT full-length transcriptome and Illumina RNA sequencing technologies to establish a transcriptome dataset. Approximately 20,560 and 17,950 full-length non-redundant transcripts were assembled in trunk xylem and needle, respectively. Of those, 29 and 20 DEGs were identified involving in the Terpenoid Backbone Biosynthesis pathway (TBB) (KO 00900) and the Diterpenoid Biosynthesis pathway (DB) (KO 00904) in trunk xylem, as well as 35 and 12 DEGs in the TBB and the DB in the needle, respectively. Compared with the low oleoresin-yield germplasm, 13 up-regulated DEGs with complete open reading frame (ORF) sequence were obtained by Venn analysis in trunk xylem and needle of a high oleoresin-yield clone. Subsequently, the expression patterns of the 16 diterpenoid biosynthesis-related genes (including three types of *DXS* genes) were validated by quantitative real-time PCR in both 12-year-old and 1-year-old high and low oleoresin-yield masson pine clones. The results showed that the 13 genes in the higher yield oleoresin trees were proved to be higher expression in comparison with the lower ones irrespectively of the tree age; as well as *MCS* (2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase), *HDS* (1-hydroxy-2-methyl-2-(*E*)-butenyl4-diphosphate synthase), and *IDI* (Isopentenyl-diphosphate Delta-isomerase) of the methylerythritol 4-phosphate (MEP) pathway demonstrated a higher expression compared to that of other TBB genes. Both cytosol-localized MVA and plastid-localized MEP pathways gave significant contributions to diterpenoid accumulation, and the latter played a more dominant role. To clarify the potential relationships between oleoresin yield and the expression patterns of 16 genes from the TBB and DB pathways, the Pearson correlation analysis was performed using three tissues (young needle, old needle, and trunk xylem) from the 60-year-old tree. Remarkably, *HMGR* (Hydroxy methylglutaryl-CoA reductase) and *ent-CDS* (Ent-copalyl diphosphate synthase)

Abbreviations: MEP pathway, 2-C-methyl-D-erythritol 4-phosphate pathway; DXS, 1-deoxy-d-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MCS, 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase; HDS, 1-hydroxy-2-methyl-2-(E)-butenyl4-diphosphate synthase; HDR, 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate reductase; MVA pathway, cytosolic mevalonate pathway; AACT, acetyl-CoA acetyltransferase; HMGS, hydroxy methylglutaryl-CoA synthase; HMGR, hydroxy methylglutaryl-CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; MDD, diphosphomevalonate decarboxylase; IDI, isopentenyl-diphosphate Delta-isomerase; GPPS, geranyl diphosphate synthase; FPPS, farnesyl pyrophosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; P450, cytochrome P450; TPS, terpenoid synthase; diTPS, diterpene synthase; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, gene ontology; qRT-PCR, quantitative real-time PCR; DEGs, differentially expressed genes; Ent-CDS, ent-copalyl diphosphate synthase; Ent-KO, ent-kaurene oxidase; CDS, copalyl diphosphate synthase; LDS, (13E)-Labda-7, 13-dien-15-ol synthase; DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate; LSDs, least significance differences; FPKM, Fragments per kilobase of transcript per million mapped reads; DB, diterpenoid biosynthesis; TBB, terpenoid backbone biosynthesis.

* Corresponding author at: Key Laboratory of Plant Resources Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Institute of Agro-bioengineering, Guizhou University, Guiyang, 550025, China.
E-mail address: xpwensc@hotmail.com (X. Wen).

https://doi.org/10.1016/j.indcrop.2021.113827
Received 8 March 2021; Received in revised form 18 June 2021; Accepted 12 July 2021
Available online 17 July 2021
0926-6690/© 2021 Published by Elsevier B.V.

were positive correlation with oleoresin yield in the new needle, whereas *FPPS* (Farnesyl pyrophosphate synthase) was a negative correlation with that in the old needle; furthermore, *HMGS* (Hydroxy methylglutaryl-CoA synthase), *IDI* and *CDS* (Copalyl diphosphate synthase) were positively correlated with oleoresin synthesis in the trunk xylem. Moreover, the Pearson correlation analysis was also carried out to determine the potential relationships between the 13 genes of the TBB pathways and *ent-CDS*, *CDS*, and *LDS* ((13E)-labda-7 13-dien-15-ol synthase) from DB pathways within three tissues of the 60-year-old masson pine. Furthermore, the diterpenoid and diterpene resin acids (DRAs) demonstrated highly positive correlations with oleoresin yield, and DRAs were detected as the most abundant diterpenoids in oleoresin. Promisingly, *HMGR*, *MCS*, *HDS*, *IDI*, *ent-CDS,* and *CDS* may be used as the key target genes of the molecular-assisted selection for high-oleoresin-yielding germplasm in this species.

1. Introduction

Conifer oleoresin is a promising natural resource and rich in various terpenes compounds that are used as raw material in medicine, cosmetics and food industry (Kelkar et al., 2006; Neis et al., 2019), as well as is also an advanced liquid biofuel (Rodrigues-Corrêa et al., 2012a). These terpenes mainly consisting of volatile monoterpenoid, sesquiterpenoid and nonvolatile diterpenoid components. Resin ducts are a complex network that produces, store and translocate the oleoresin in pine trees, and stored oleoresin will physically flow out when the resin ducts are mechanical damaged (Vázquez-González et al., 2020). Conifers have a complex suite of resistance mechanisms that defend against insects and pathogens, including mechanical and chemical mechanisms that can be present constitutively or induced upon attack (Keeling and Bohlmann, 2006; Huang et al., 2020). Once exposed to pressure, the trees may proactively generate a rational defence strategy through distributing storage reasonably to the growth and emersion of constitutive metabolites, while also retaining storage to induced metabolites when required (Züst and Agrawal, 2017). The above features that oleoresin plays in conifer defence against insects and disease may be considered as a significant origin for the universal distribution and evolutionary success of this species (Phillips et al., 2006). Previous reports were concentrated mainly in induced regulation of oleoresin by wounding or chemical stimulants (Moreira et al., 2014; Huang et al., 2020; Ott et al., 2021). Although several research pieces associated with constitutive defence have been performed in coniferous trees, little information is still insufficient to explain the mechanisms.

Some researchers put forward that the oleoresin yield is an advanced heritable trait, such as several investigations in loblolly pine, maritime pine (*Pinus pinaster*), and slash pine (*Pinus elliottii*) have illustrated that genetic factors mightily affect oleoresin yield (Lombardero et al., 2000; Tadesse et al., 2002; Roberds et al., 2003). Extensive genetic gains can be acquired from the selection of high oleoresin yielders through conventional breeding methods. For example, Westbrook et al. (2013) and Zeng et al. (2013) simultaneously proposed that oleoresin yield was heritable and could be raised in loblolly pine (*P. taeda*) by a selection. Processing a long gradual development and got prosper, Allen et al. (2015) actualized PCR markers of terpene synthase genes for assisted selection of high resin yielders in *Pinus roxburghii* Sarg. High-yield oleoresin germplasms have been bred successfully at the provenance and family level in pine trees, but little achievements were obtained at the molecular level (Liu et al., 2015; Junkes et al., 2019). Meanwhile, the key candidate genes involved in potential oleoresin yield are chiefly revealed by transcriptome analyses (Wen et al., 2018). However, another research also demonstrated that pine oleoresin productivity was positively or negatively correlated with tree morphological factors, for instance, height, diameter, tracheid length, and branching angle (Knebel et al., 2008). And the oleoresin yield was also related to tree dendrometry, stand density, and tapping-induced activities (Rodríguez-García et al., 2014). Although the selection of high-yield trees has been widely used, research and application are relatively backward so far in pine production.

Oleoresin is a complicated admixture mainly composed of terpenes, and some researchers have proposed that the diterpenoids occupied a vital position in the yield and quality of oleoresin (Liu et al., 2015a). For example, especially the abietane-type diterpene resin acids (DRAs), labdane-type DRAs, and other pimarane-type DRAs were considerable ingredients of the oleoresin defence system in species of fir (*Abies*), pine (*Pinus*) and spruce (*Picea*) (Whitehill et al., 2019). Diterpenoids are derived from terpenes and are also known as isoprenoids, and are synthesized from C5 isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) by the MEP and MVA pathways localized to the plasmids and cytoplasm, respectively (Tholl, 2015). The diversity of diterpenoids were generated by diterpene synthases (diTPSs), which catalyze the biosynthesis of the central precursor geranylgeranyl diphosphate (GGPP); the GGPP structure was cyclized and rearranged to produce a series of scaffolds with different stereochemistry (Peters, 2010). Recently, it was proposed that diTPSs shared a typical γβα structure domain and were defined as three pivotal diTPS types: monofunctional class-I diTPSs, monofunctional class-II diTPSs, as well as bifunctional class-I/II diTPSs (Zerbe and Bohlmann, 2015). Ent-copalyl-PP was proposed as a possible precursor to the gibberellin family, and ent-copalyl diphosphate synthase was reported to be involved in the first step of gibberellin cyclization. Copalyl-PP was presented as a possible precursor to the DRAs family, and miltiradiene synthase/ copalyl diphosphate synthase might be an essential regulatory enzyme of the DRAs biosynthesis. Some reports demonstrated that the expression level of terpenoid biosynthesis-related genes probably parallels terpenoid productivity, for example, the 12-Deoxyphorbol-13-phenylacetate in *Euphorbia resinifera* (Zhang et al., 2019). Several studies isolated the downstream *diTPS* genes from pines, while the upstream TB gene information is also relatively straightforward. And it remains fuzzy whether there are some correlations between oleoresin and gene expression patterns in the terpenoids metabolic pathway so far.

Masson pine (*Pinus massoniana* L.) is the primary native pine species tapped in China; it occupies a vast scale in oleoresin total yields of the country annually, and the proportion even reached 90 % in a period (Liu et al., 2015a). Remarkable genetic variations of oleoresin yield have been found at different masson pine families, ranging from 14.12–50.55 g per day (Liu et al., 2020). With the development of molecular biology, genetic identity and functional genomics had been exploited for molecular assistance in breeding work (Burdon and Wilcox, 2007). To date, there have been no reports involved in the complete genome sequences to study the DRAs biosynthetic pathway of masson pine. Only fractional germplasms have been explored by Next-generation sequencing technologies, such as RNA-Seq analysis of gene expression, which has contribute significantly in the achievements of basic plant science (Liu et al., 2015; Mao et al., 2019). Also, the downstream information about diterpenoid biosynthesis still is very insufficient in masson pine so far. The application of single-molecule long-read isoform sequencing (SMRT-Seq) technology by Pacific Biosciences (PacBio) has enabled researchers to gain long-read or full-length transcriptomes with complete coding sequences (Hoang et al., 2017). The scheme of combining Illumina RNA-Seq and SMRT-Seq has been used to produce more comprehensive knowledge at the transcriptional level. In the future, the global knowledge of the diterpenoid biosynthesis pathway in masson pine will be fully characterized.

The present study reveals the essential genes of a diterpenoid

biosynthetic pathway in a high oleoresin yielder and unravels the correlation between the expression levels and oleoresin yield in masson pine. The transcriptome data of high- and low-oleoresin yield samples of masson pine were conducted using the SMRT-Seq and Illumina RNA-Seq technology, and the differentially expressed genes (DEGs) were identified using enrichment analyses in needle and trunk xylem. And the expression status of the 13 DEGs and three types of *DXS* genes were validated using quantitative real-time PCR (qRT-PCR) in both 12-year-old and 1-year-old high and low oleoresin-yield clones. Subsequently, we performed the Pearson correlation analysis in 60-year-old trees with different oleoresin productivities, and the expression levels of the 16 genes varied among the new needle, old needle, and trunk xylem. Moreover, the Pearson correlation analysis was also performed to ascertain the potential relationships between the 13 upstream genes of the TBB pathways and three downstream genes from DB pathways within the three tissues in 60-year-old masson pine. Parallel to this, the oleoresin compositions of 60-year-old germplasms were committed to metabolite profile *via* GC–MS analysis. We found a correlation between oleoresin yield and gene expression level, which would supply reference of masson pine high-oleoresin-yielding germplasm for molecular-assisted selection.

2. Materials and methods

2.1. Plant material

The 12-year-old clones were cultivated in a seed orchard of the Nanning Forestry Research Institute, Guangxi, China. According to the yield of oleoresin, all samples were divided into two groups. One group was a high-oleoresin clone (three trees as three biological replicated trees) with oleoresin yields more than 15.0 $g{\cdot}d^{-1}$ cm^{-1}, another was a low yielder (another three trees as three biological replicated trees) with oleoresin yields less than 5 $g{\cdot}d^{-1}$ cm^{-1}. A total of six trees were gathered in the same environmental and management conditions for further investigation. The trunk xylem of six trees was collected in August 2018 for PacBio SMRT transcriptome sequencing. And 12 samples were used for Illumina RNA sequencing, *i.e.,* two clones (high-yield and low-yield) × two tissues (trunk xylem and needle) × three replicated trees. The xylem samples were harvested from the trunk one meter above the ground. An equal volume of needles was obtained from four orientations at the same height in each tree.

The high and low oleoresin-yield clones of 1-year-old seedlings were cultivated from the seeds of 12-year-old high and low oleoresin-yield mature masson pine, respectively. When the seeds germinate and grow to a height of 10 cm, the seedlings were grown under natural conditions at the Guizhou University. The stem and needle were collected for qRT-PCR in August 2020, and three biological replicates were included in the qRT-PCR assays.

The 60-year-old masson pine was grown outdoor and had no shade with the same environmental conditions of the Guizhou University, Guizhou, China. A total of nine trees have the same appearance, for example, tree height and breast diameter. The oleoresin yield was gain using the bark streak method of wounding for oleoresin tapping from the trunk one meter above the ground (Rodrigues-Corrêa et al., 2012b). The new needle (current year), old needle, and trunk xylem (from the trunk one meter above the ground) were collected for qRT-PCR in October 2020. All the above samples were placed into liquid nitrogen immediately and then stored at -80 °C until further processing. An opening schematic with a diagrammatic depiction of sample collection details is shown in Fig. 1.

2.2. Total RNA preparation and cDNA synthesis

The total RNA was extracted from the needle, stem, and trunk xylem

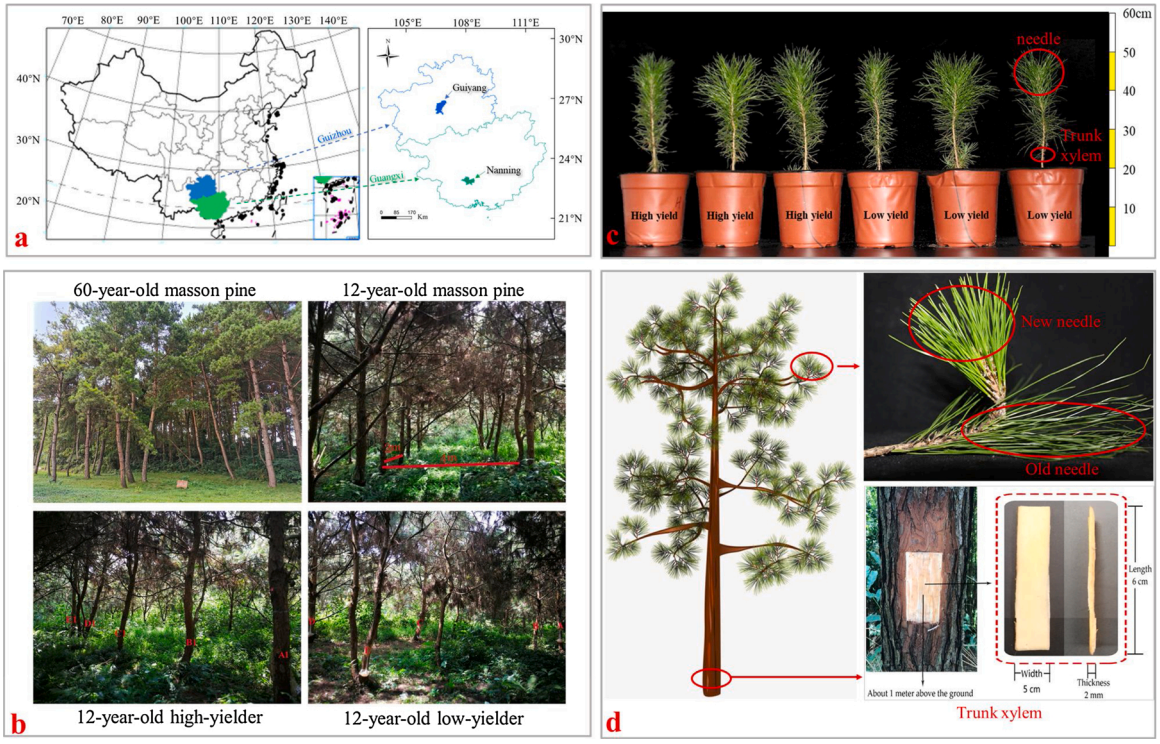

Fig. 1. Scheme of plant sample collection for the manuscript. (a) Geographical distribution of all masson pine in this study; (b) The morphological character of 12-year-old (from Guangxi Province) and 60-year-old (from Guizhou Province) mature masson pine; (c) The high and low oleoresin-yield clones of 1-year-old seedlings and the location of sampling tissues. (d) The location of sampling tissues in two types of mature masson pine. The 1-year-old seedlings were cultivated from the seeds of 12-year-old high and low oleoresin-yield mature masson pine. The red circle indicates the location of sampling tissues. After removing the bark and phloem at the sample trees about one meter from the ground, the trunk xylem was obtained that length was 60 mm, the width was 50 mm, and 2 mm of thickness.

by polysaccharide polyphenol plant total RNA extraction kit (SENO, Zhangjiajie, China). The microplate spectrophotometer (Thermo Fisher Scientific, USA) was used to detect RNA purity, concentration, and integrity. According to the manual, the cDNA was first synthesized using a PrimeScript TMRT reagent kit with gDNA Eraser (TaKaRa, Dalian, China).

2.3. RNA sequencing, assembly, and annotation

The total cDNA library was isolated and constructed on an Illumina HiSeq Xten platform (Illumina Inc., CA, USA) at Biomarker Technologies Corporation (Beijing, China) according to the method described by Barrero et al. (2011). After removing raw reads that included adaptor sequences, empty reads, or low-quality sequences, the clean reads were mapped to the non-redundant SMRT reference through RSEM software and STAR software (Li and Dewey, 2014). Gene function classifications with gene ontology (GO) annotations of all the unigenes were determined by Blast2GO software (Conesa and Götz, 2008) and with WEGO software to display GO functional classification (Ye et al., 2018). The relative expression levels were normalized by FPKMs (Fragments per kilobase of transcript per million mapped reads). The differential expression based on the read counts values among sample groups was analyzed by DESeq2 (Anders and Huber, 2010). Both the absolute values of the FDR (false discovery rate) <0.01 and $|\log2 \text{ (fold change)}| \geq 1$ were chosen as thresholds to detect differential significance. The statistical analysis and data visualization of DEGs were implemented using ggplot2 R Package (Yu et al., 2012).

2.4. Quantitative real-time PCR (qRT-PCR) assay

Unigenes related to terpenoid metabolic pathways were selected for validation by qRT-PCR. The Maxima® SYBR Green/ROX qPCR Master Mix (2X) (Thermo Fisher Scientific, USA) was used to detect the target sequence on a CFX96 Real-Time PCR Manager 3.1 (BioRad, United States). Each PCR mixture (10 μL) contained 1 μL of synthesized cDNA, 5 μL of SYBR Green Real-time PCR Master Mix, 0.3 μL of each primer (10 mM), and 3.4 μL of ddH_2O. Specific primers used in qRT-PCR were designed using Primer Premier 5 according to the open reading frame (ORF) sequence of target unigenes. The *UBI* gene was selected as the reference gene. The information of primer sequences was listed in Supplementary Table S1. Water was used to replace the cDNA template as a negative control in the same qRT-PCR reaction for each primer pair. Three technical and three biological replicates with similar results were performed in the qRT-PCR assays. The CT values of the 16 genes in a needle and trunk xylem were used to evaluate the FPKM results of transcriptome in 12-year-old clones. The qPCR experiment data was analyzed using the $2^{-\Delta CT}$ and $2^{-\Delta\Delta CT}$ methods (Livak and Schmittgen, 2001).

2.5. Oleoresin terpenes extraction and GC–MS analysis

Oleoresin terpenes were obtained from the trunk xylem in 60 years old masson pines. The oleoresin from the trunk xylem was gained by cutting the tree without the use of chemical stimulants. On each selected tree, one wound was made that length was 80 mm and width was 20 mm, 5 mm of deep fresh secondary xylem tissues adjoining the cambium layer were wounded at the sample trees after removing the bark and phloem. Each wound was made at the same height (breast height); on the same level, the right end of the wound facing the sun has a slight upwards slope (30°). Just below the damage, a tightly fitting end of tin foil (100 × 100 mm) was pasted in the tree bark, the other end of the tin foil was rolled into a tube with a diameter of 20 mm, and the tube was inserted into the collection tube and left from 12 h to fill with oleoresin. Collection tubes were hermetically closed and refrigerated until required for analysis. Samples were taken on August 30, 2020.

For qualitative analysis, the terpenoids in xylem from 1 g of oleoresin were extracted using 0.5 mL of ethanol for 12 h, and 50 μL of tetramethylammonium hydroxide was added as methylation reagent. The heptadecanoic acid was used as internal standards of diterpenes, and the isobutylbenzene was used as internal standards of monoterpene and sesquiterpene as previously described Livak et al. (2001). Quantitative analyses of terpenoid extractions were determined on an Agilent 7890B gas chromatograph equipped with an Agilent 7000D mass spectrometer (Agilent J&W Scientific, USA) and an HP-5MS column (60 m × 0.25 mm ×0.25 μm; Agilent Technologies, USA). Helium was used as a carrier gas at 1.0 mL min^{-1}. The oven program temperature was started at 50 °C for 2 min, increased to 80 °C at a rate of 3 °C min^{-1} and held for 4 min, followed by an increase at a rate of 5 °C min^{-1} to 180 °C (6 min hold), and then at a rate of 10 °C min^{-1} (6 min hold) reached to 230 °C, a final step to 250 °C at a rate of 20 °C min^{-1} (6 min hold). MS conditions were performed in EI mode, and electron energy was 70 eV, ion source was 230 °C; GC–MS interface zone was 250 °C, and a scan range of 50–500 mass units. Terpenoids were validated by comparison of mass spectra with corresponding compounds in the NIST Mass Spectral Library, and specific standard substances from Sigma-Aldrich were used as references of the retention time.

2.6. Statistical analysis

The statistical analyses were carried out using ANOVA and Pearson's simple correlation using SPSS software (version 9.2). Least significance differences (LSDs) were calculated at $P < 0.05$ and $P < 0.01$.

3. Results

3.1. Functional annotation and classification of transcriptomes in trunk xylem and needle

Sequencing of the cDNA libraries from the trunk xylem of the high oleoresin yielder on the PacBio Sequel platform generated 25.70 Gb of nucleotides, the low-yield clone generated 23 Gb of nucleotides. Summary of SMRT full-length transcriptome and RNA-Seq data have been published by Mei et al. (2020), including clean reads, mapped reads, number of full-length reads, full-length percentage, quality (%), *etc.* Based on GO annotation, a total of 20,560 and 17,915 unigenes in trunk xylem and needle were classified into 48 subcategories belonging to three main categories (Fig. 2, Supplementary Table S2). In the biological process category, "metabolic process" (7336/ 6362 Note: the former was in trunk xylem, the latter was in the needle, the following same format was the same commentary) and "cellular process" (6361/ 5681) were included a significantly larger number of unigenes compared with other subcategories, and were followed by "single-organism process" (5210/ 4465), "response to stimulus" (2719/ 2343), "biological regulation" (2252/ 1938), and "localization" (1701/ 1477), and so on. The subcategories "cell part" (5328/ 4955), "cell" (5328/ 4955) and "organelle" (3920/ 3760) in the cellular component category as well as "catalytic activity" (7029/ 5800) and "binding" (6883/ 5776) in the molecular function category contained the largest number of unigenes. However, few genes were clustered to subcategories of "nutrient reservoir activity" (1/ 0), "metallochaperone activity" (3/ 2), "extracellular region part" (4/ 4), and "cell killing" (4/ 3).

A total of 8361/ 7477 unigenes were successfully assigned to 125 KEGG pathways in trunk xylem and needle, respectively (Fig. 3, Supplementary Table S3). Among them, the four most frequently represented pathways that more than 270 unigenes, including "Spliceosome" (388/ 328), "Protein processing in endoplasmic reticulum" (381/ 323), "Carbon metabolism" (304/ 309), "Biosynthesis of amino acids" (296/ 277). Some crucial pathways involved in terpenoids synthesis were also identified, including "Terpenoid backbone biosynthesis" (43/ 52) and "diterpenoids biosynthesis" (22/ 13). Moreover, 121/ 93 unigenes were clustered into the "Flavonoid biosynthesis" pathway, and 129/ 76 unigenes were assigned to the "Phenylpropanoid biosynthesis" pathway.

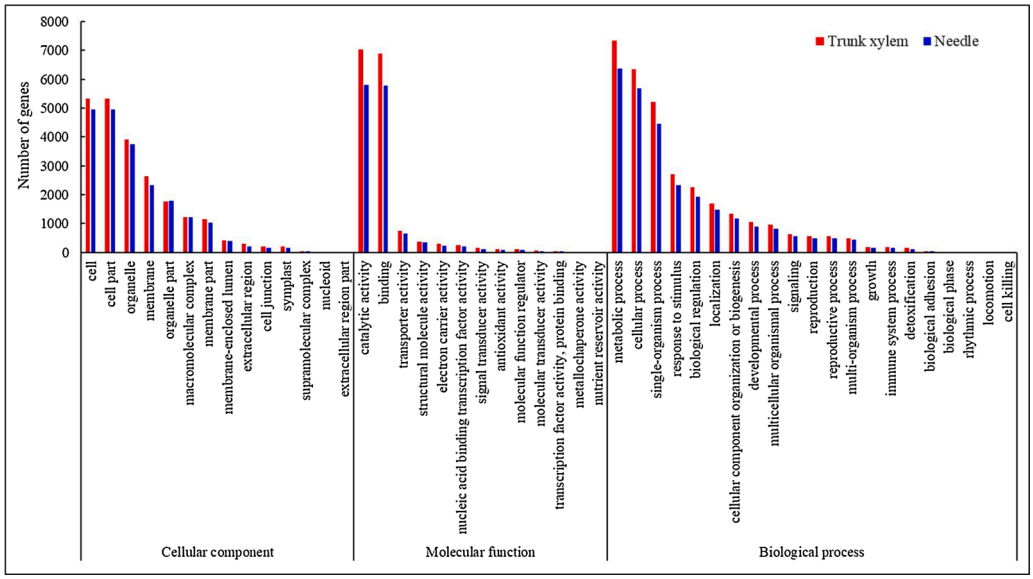

Fig. 2. Gene ontology classifications of unigenes in trunk xylem and needle. The results were summarized in three main categories: biological process, cellular component, and molecular function. Red histogram represented the trunk xylem, and blue histogram represented the needle.

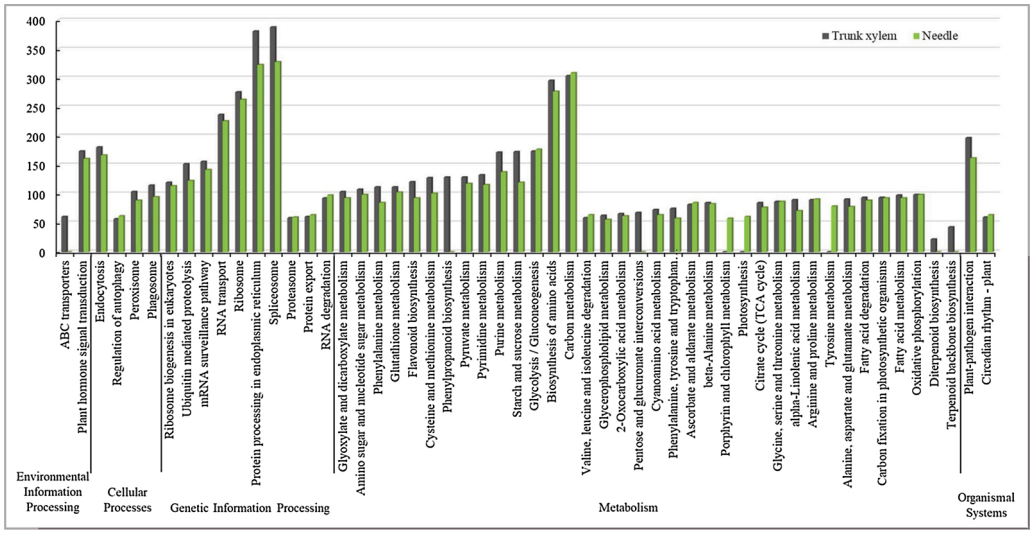

Fig. 3. KEGG classifications of unigenes in trunk xylem and needle. Gray histogram represented the trunk xylem, and green histogram represented the needle. The y-axis represented the number of unigenes.

Compared with the needle, the "ABC transporters", "Phenylpropanoid biosynthesis", "Pentose and glucuronate interconversions", "Diterpenoid biosynthesis" and "Terpenoid backbone biosynthesis" were only found in the trunk xylem. While the "Porphyrin and chlorophyll metabolism", "Photosynthesis" and "Tyrosine metabolism" were only discovered in the needle (Fig. 3).

3.2. KEGG enrichment pathways of DEGs between high and low oleoresin-yield masson pine

KEGG enrichment pathways showed that 1909 DEGs in trunk xylem were enriched in 20 KEGG pathways, the followed pathways were significantly enriched, including "ABC transporters", "Spliceosome", "Phenylpropanoid biosynthesis", "Glutathione metabolism", "Plant-pathogen interaction", "Flavonoid biosynthesis", "Cyanoamino acid metabolism", and "Diterpenoid biosynthesis" (Fig. 4a). In contrast, a total of 1994 DEGs in the needle were enriched in 20 KEGG pathways, of which the "Photosynthesis-antenna proteins", "Photosynthesis", "Regulation of autophagy", "Glutathione metabolism", and "Circadian rhythm-plant" were significantly enriched pathways (Fig. 4b). Additionally, to clearly present the number of differential expression genes, the Venn diagrams were implemented. Compared with low-yielder, a total of 11,198 up-regulated and 10,757 down-regulated DEGs in the high-yielder xylem; 9353 up-regulated and 8562 down-regulated DEGs in the high-yielder needle (Fig. 4c).

3.3. Candidate DEGs of the diterpenoid biosynthesis pathway in 12-year-old masson pines

An annotation found the characterization of the potential candidate genes involved in diterpenoid biosynthesis in the masson pine transcriptome. Totally, 172 unigenes were found in the Terpenoid Backbone Biosynthesis (TBB) pathway (KEGG map00900), the Diterpenoid Biosynthesis (DB) pathway (KEGG map00904) found 52 unigenes. Eighteen genes of the TBB pathway were matched with 1–16 unigenes (Mei et al., 2020). Here, 29 and 20 DEGs were identified involving the

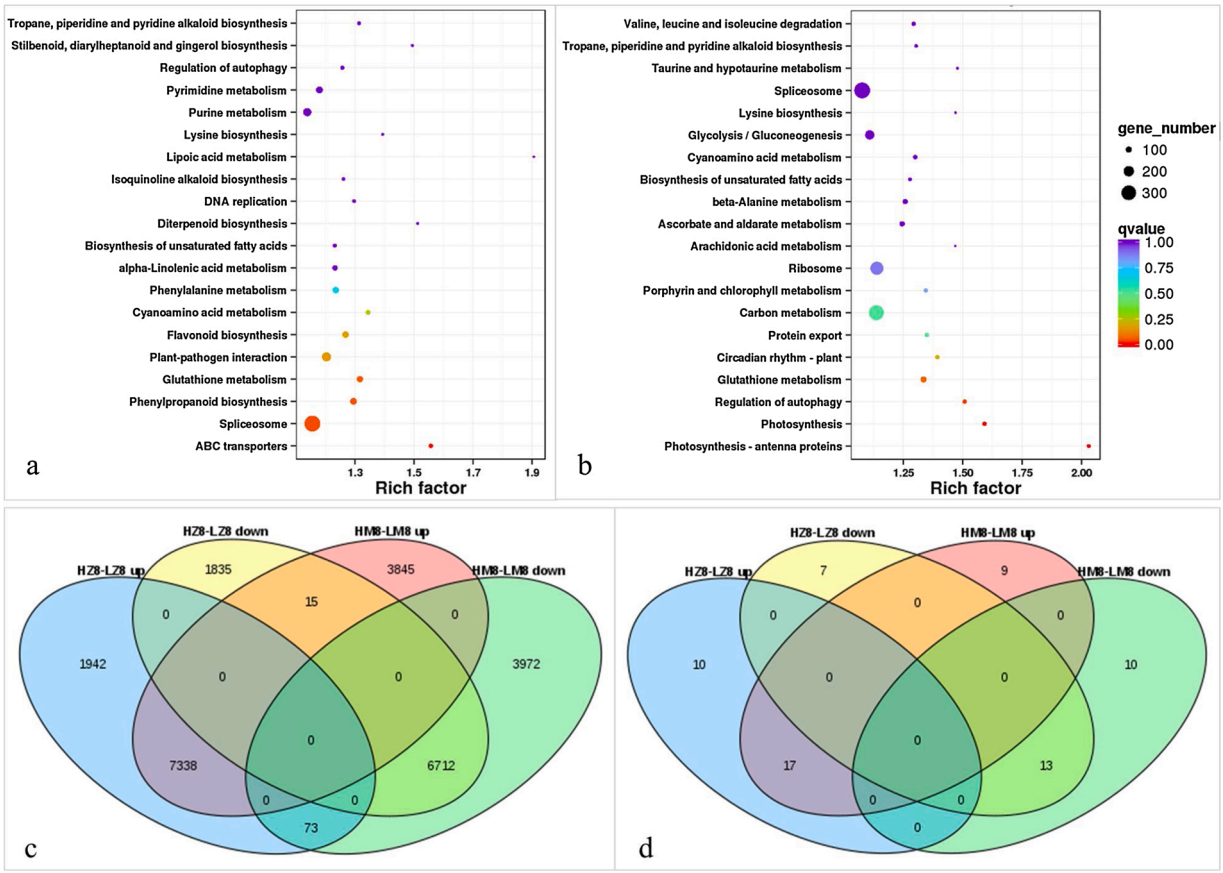

Fig. 4. KEGG enrichment pathways of DEGs and Venn analysis between high-yielder and low-yielder in the xylem and needle. (a) KEGG enrichment pathways in high yielder *vs.* low yielder trunk xylem; (b) KEGG enrichment pathways in high yielder *vs.* low yielder needles. (c) The total of differential expression genes between high-yielder and low-yielder in the xylem and needle by Venn diagram. (d) The differential expression TBB and DB genes between high-yielder and low-yielder in the xylem and needle by Venn diagram.

TBB pathway and the DB pathway in the trunk xylem, 35 and 12 DEGs in the TBB and the DB in the needle. Compared with low-yielder, a total of 26 up-regulated and 23 down-regulated DEGs in the high-yielder xylem; 27 up-regulated and 20 down-regulated DEGs in the high-yielder needle (Fig. 4d, Fig. 5).

Compared with the low oleoresin-yield masson pine, 13 up-regulated expression DEGs with complete ORF sequence were obtained by Venn analysis in trunk xylem and needle of high oleoresin-yield masson pine, simultaneously. As the first key enzyme in the MEP pathway, three types of *DXS* genes by KEGG annotation also were used to be regarded as the candidate genes at the subsequent qRT-PCR analysis. The expression levels of the diterpenoid biosynthesis genes were analyzed by the FPKM values of the transcriptome (Fig. 6). The results showed that the expression level of *AACT7*, which had a relatively higher expression level compared with other genes of the MVA pathway in the combination of HM-LM. The *MCS4*, *IDI7,* and *ent-CDS9* had somewhat higher expression levels than other genes in HZ-LZ. Conversely, the *PMK11* and *CDS* had relatively lower expression levels among the 16 tested genes in xylem and needle. Compared with the MVA genes, the expression levels of MEP biosynthesis gene were visibly higher, except for *AACT7*. In the trunk xylem, the expression level of *DXS6* is the highest, but the difference was not significant between the high- and low-yielder trunk xylem. In the needle, the *DXS3* had a relatively higher expression level than *DXS6* and *DXS8*.

3.4. Expression profile of target genes in 12-year-old masson pines

The accuracy of RNA-Seq was validated by qRT-PCR. High resemblances were observed, except for *DXS6* and *DXR2*, which had relatively lower expression levels in high-yielders than other genes (Fig. 7). Moreover, the expression of *MCS*, *HDS*, and *IDI*, which catalyzed the essential conversion step from IPP to DMAPP in the MVA and MEP pathways, was significantly higher than that of other MEP pathway genes in high-yield masson pine. Combined with the FPKM and qRT-PCR values, the results indicated that the MEP pathway genes had a higher expression content than other MVA pathway genes, except for three *DXS* genes. Additionally, the comparison with the upstream genes revealed that the expression levels of *ent-CDS9* involved in diterpenoid biosynthesis were visibly higher in the needle, while the expression level of *CDS1* was visibly lower in two tissues.

3.5. Expression profile of terpenoid biosynthesis-related genes in 1-year-old tree

In the present study, the expression profile of terpenoid biosynthesis-related genes was analyzed in the needle and trunk xylem of 1-year-old masson pine (Fig. 8). It was emphasized that all 1-year-old masson pine was cultivated from the seeds of 12-year-old. The results showed that most genes significantly different expression in the two tissues of high-yield-oleoresin masson pine compared with low-yield. However, it was distinct that the *HMGR7*, *HMGS7,* and *DXR2* were up-regulated in the low-yield needle compared with the 12-year-old masson pine. Additionally, the fold change of most genes from the MEP pathway, including *DXS3*, *DXS6*, *DXR2*, *MCS4*, *HDS4*, but not *DXS8, and PMK11*, were higher in high-yield trunk stems than in the MVA pathway. Notably, the *PMK11* of MVA pathway, *ent-CDS9* related to the gibberellin biosynthesis and *CDS1* of DRA biosynthesis, were exceptionally significantly up-regulated expression in high-yield masson pine compared with low-

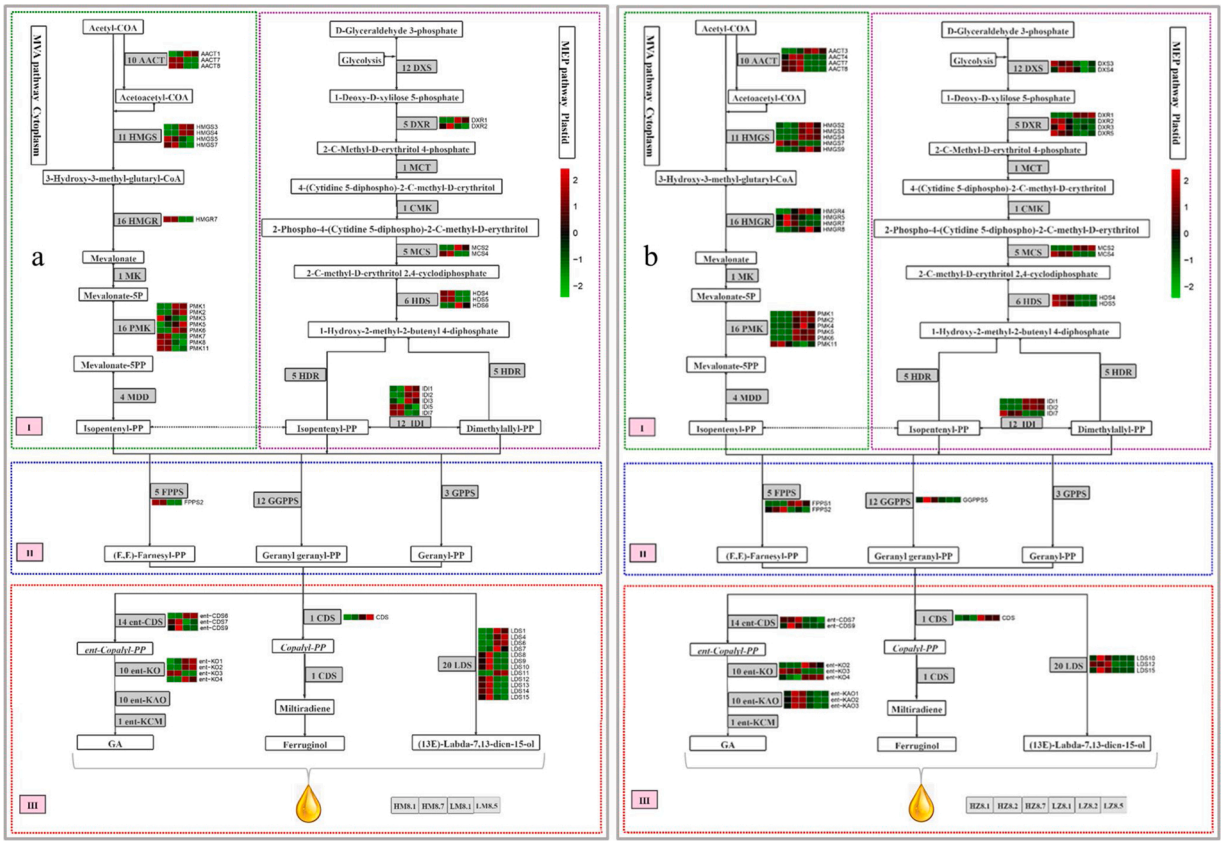

Fig. 5. The KEGG pathways of diterpenoid biosynthesis as well as the expression patterns of related DEGs in trunk xylem and needle. (a) KEGG pathways and expression patterns in the trunk xylem; (b) KEGG pathways and expression patterns in the needles. In the gray box of the KEGG pathway, for example, "10 AACT", the 10 represented the number of AACT transcripts. The expressions were showed by heatmap, estimated using FPKM value for each transcript. Red represented a high expression level, and green was a low expression level; the color gradually from green to red meant gene expression abundance from low to high. At the bottom of the figure, the gray frame represented the location of each sample in the heatmap. "H" was the high oleoresin yielder; "L" represented the low oleoresin yielder; "M" was trunk xylem, "Z" was needle; "8" was August. The "1", "2", and "7" meant three biological duplications in high yielder, the "1", "2", and "5" meant three biological duplications in low yielder. Additionally, DEGs represented the differentially expressed genes; MVA, mevalonic acid; MEP, methylerythritol 4-phosphate; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MCS, 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase; HDS, 1-hydroxy-2-methyl-2-(*E*)-butenyl4-diphosphate synthase; HDR, 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate reductase. AACT, acetyl-CoA acetyltransferase; HMGS, hydroxy methylglutaryl-CoA synthase; HMGR, hydroxy methylglutaryl-CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; MDD, diphosphomevalonate decarboxylase; IDI, isopentenyl-diphosphate Delta-isomerase; GPPS, geranyl diphosphate synthase; FPPS, farnesyl pyrophosphate synthase; GGPPS, geranylgeranyl diphosphate synthase; ent-CDS, ent-copalyl diphosphate synthase; ent-KO, ent-kaurene oxidase; CDS, copalyl diphosphate synthase; LDS, (13E)-Labda-7, 13-dien-15-ol synthase.

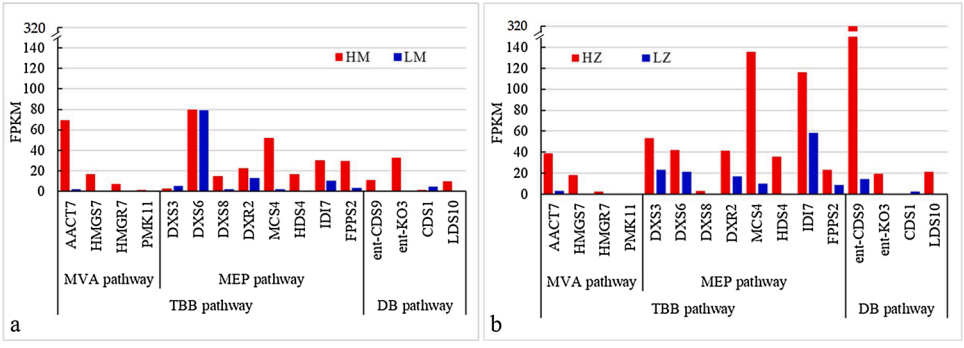

Fig. 6. The expression levels of target genes with fragments per kilobase of transcript per million mapped reads (FPKM) values in the trunk xylem and the needle of 12-year-old trees. (a) Gene expression levels in high yielder *vs.* low yielder trunk xylem, (b) Gene expression levels in high yielder *vs.* low yielder needles. "H" was the high oleoresin yielder; "L" represented the low oleoresin yielder; "M" represented the trunk xylem; "Z" represented the needle. The red histogram represented a high-yield oleoresin clone, and the blue histogram represented a low-yield oleoresin clone. MVA was the mevalonic acid pathway, MEP was the methylerythritol 4-phosphate pathway. TBB was the terpenoid backbone biosynthesis pathway, DB was the diterpenoid biosynthesis pathway. The full name of genes was listed in Fig. 5.

yield one.

3.6. Correlations between terpenoid gene-expression and oleoresin yield in 60-year-old tree

We obtained nine diverse oleoresin-yield germplasms in 60-year-old

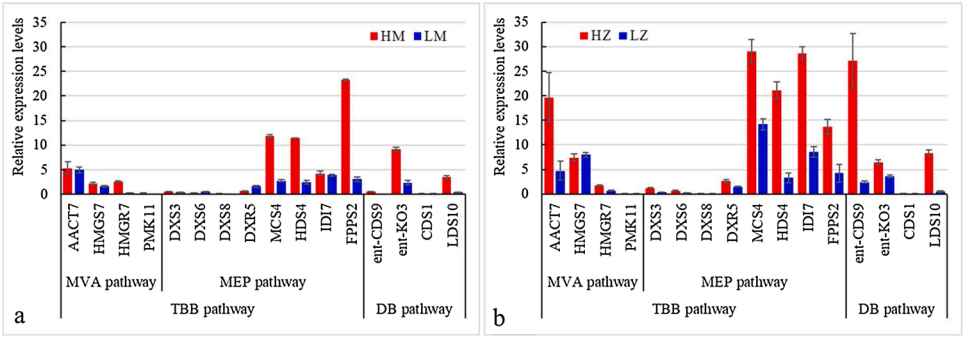

Fig. 7. The expression levels of target genes in the trunk xylem and the needle of 12-year-old masson pine detected by qRT-PCR. (a) Gene expression levels in high yielder *vs.* low yielder trunk xylem, (b) Gene expression levels in high yielder *vs.* low yielder needles. "H" was the high oleoresin yielder; "L" represented the low oleoresin yielder; "M" was the trunk xylem; "Z" was the needle. MVA was the mevalonic acid pathway, MEP was the methylerythritol 4-phosphate pathway. TBB was the terpenoid backbone biosynthesis pathway, DB was the diterpenoid biosynthesis pathway. The red histogram meant a high-yield oleoresin clone, the blue histogram represented a low-yield one. Error bars refer to the standard deviation (SD) of three biological repeats. The full name of genes was listed in Fig. 5.

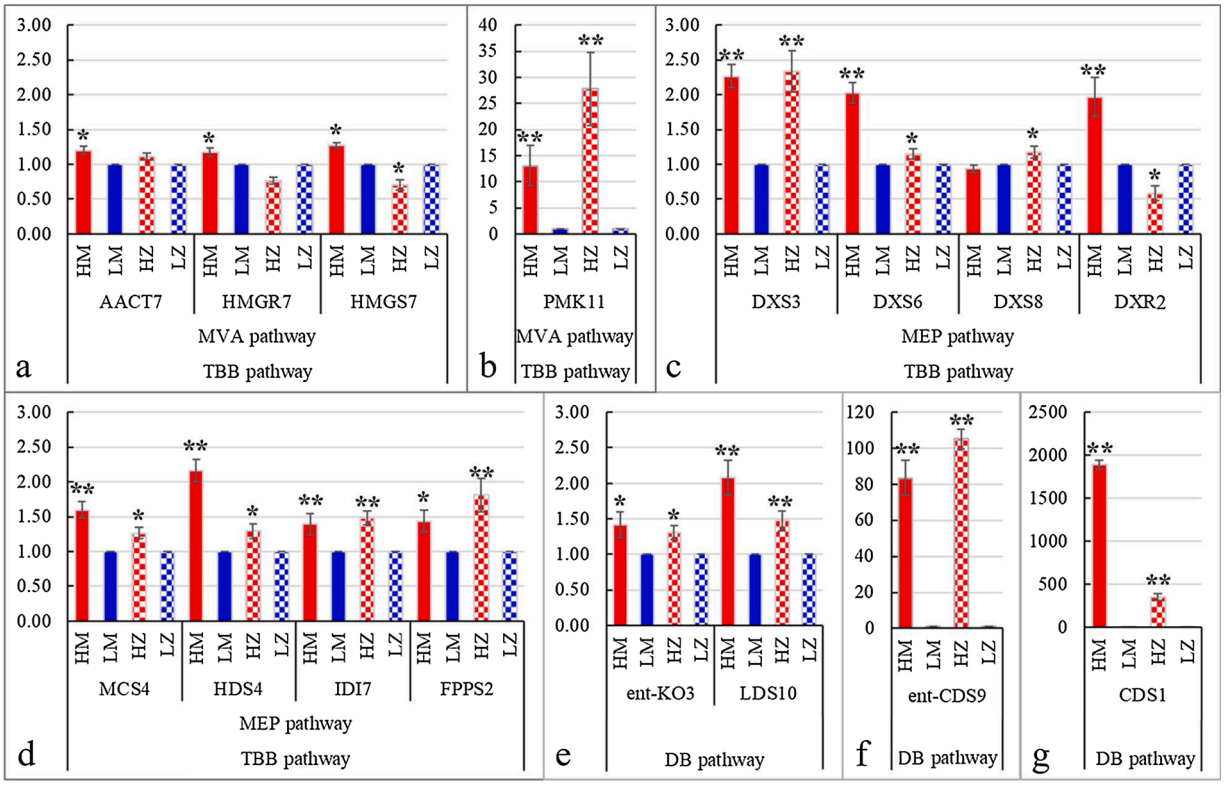

Fig. 8. The fold change of target genes in 1-year-old high-yield compared with the low-yield masson pine detected by qRT-PCR. (a) and (b) indicates that four genes belong to the MVA pathway of the TBB pathway. (c) and (d) indicated that the eight genes originate from the MEP pathway of the TBB pathway. (e), (f) and (g) indicated that the four genes belong to the DB pathway. "H" was the high oleoresin yielder; "L" represented the low oleoresin yielder; "M" was the trunk xylem; "Z" was the needle. The red histogram represented a high-yield oleoresin clone, and the blue histogram was a low-yield oleoresin clone. MVA was the mevalonic acid pathway, MEP was the methylerythritol 4-phosphate pathway. TBB was the terpenoid backbone biosynthesis pathway, DB was the diterpenoid biosynthesis pathway. Error bars refer to the standard deviation (SD) of three biological repeats. According to two-tailed analysis, symbol ** and * showed statistically significant differences between high and low yielder samples at the levels of $p < 0.01$ and $p < 0.05$. The full name of genes was listed in Fig. 5.

masson pine, and we also investigated the expression patterns of candidate genes in new foliage, old foliage, and trunk xylem (Fig. 9). The results found that all 15 genes were expressed in the three tissues, exclude the expression of *ent-CDS9* in the trunk xylem. Four genes from the MEP pathway, including *DXS3*, *DXS6*, *DXS8,* and *DXR2*, were at a lower expression in the trunk xylem than the needle. The *DXS3*, *DXS6,* and *DXS8* had identical expression patterns in new foliage. Subsequently, *MCS4*, *HDS4*, and *IDI7* of the MEP pathway had a higher expression than other genes of TBB, whereas the expression levels of genes involved in the MVA pathway were relatively lower. Compared with the young needle, the *MCS4* and *HDS4* had analogous expression patterns between the old needle and trunk xylem. Moreover, the *PMK11* and *CDS1* had similar expression patterns between the three tissues, respectively. The expression of *ent-CDS9* showed extremely significant tissue differences, and it had a higher expression in a new needle. Compared with the trunk xylem, the *ent-KO3* had similar expression patterns between the new and old needles.

We performed correlation analysis to determine the potential relationships between oleoresin yield and relative expression values of the TBB and DB pathways genes within the three tissues, as shown in Table 1. In new needle, *ent-CDS9* ($r = 0.9394$; $p < 0.01$) and *HMGR7* ($r = 0.9579$; $p < 0.01$) had obviously positive correlations with oleoresin yield, but *HMGS7* (r=-0.8142; $p < 0.01$) had significantly negative correlation with oleoresin yield; in old needle, only *FPPS2* (r=-0.832; p

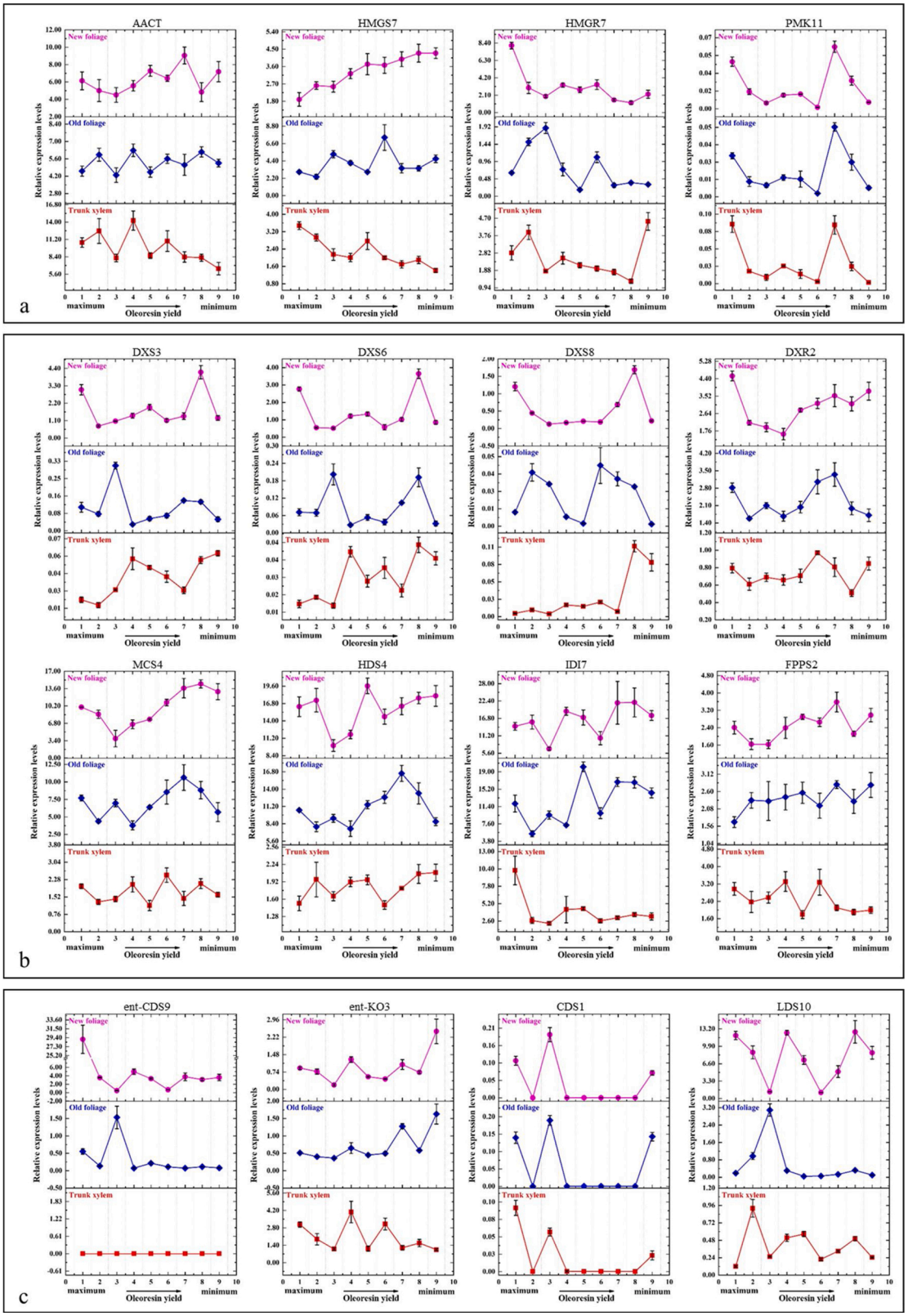

Fig. 9. The expression levels of target genes in 60-year-old masson pine with different yields. (a) The four genes belong to the MVA pathway of the TBB pathway; (b) The eight genes belong to the MEP pathway of the TBB pathway; (c) The four genes from the DB pathway. The number of 1 to 9 on the X-axis represented nine trees with different yields from high to low. Error bars refer to the standard deviation (SD) of three biological repeats. MVA was the mevalonic acid pathway, MEP was the methylerythritol 4-phosphate pathway. TBB was the terpenoid backbone biosynthesis pathway, DB was the diterpenoid biosynthesis pathway. The full name of genes was listed in Fig. 5.

Table 1
Pearson correlation analysis between the expression trends of 16 genes and oleoresin yield.

Tissues		*MCS4*	*PMK11*	*FPPS2*	*ent-KO3*	*ent-CDS9*	*AACT7*	*DXR2*	*HDS4*
New foliage	Pearson correlation	−0.2155	0.3475	−0.2025	−0.1724	0.9394**	−0.1491	0.4352	−0.0465
	Sig. (2-tailed)	0.5775	0.3595	0.6014	0.6574	0.0002	0.7019	0.2417	0.9054
Old foliage	Pearson correlation	−0.024	0.171	−.832**	−0.366	0.279	−0.372	0.25	−0.179
	Sig. (2-tailed)	0.95	0.659	0.005	0.333	0.467	0.325	0.516	0.646
Trunk xylem	Pearson correlation	0.149	0.56	0.422	0.42	–	0.299	0.126	−0.585
	Sig. (2-tailed)	0.702	0.117	0.258	0.26	–	0.434	0.747	0.098
Tissues		*HMGR7*	*LDS10*	*DXS3*	*DXS8*	*HMGS7*	*DXS6*	*CDS1*	*IDI7*
New foliage	Pearson correlation	0.9579**	0.2763	0.281	0.2999	−0.8142**	0.3319	0.3987	−0.296
	Sig. (2-tailed)	0	0.4716	0.4639	0.433	0.0075	0.3829	0.2878	0.4392
Old foliage	Pearson correlation	0.114	−0.015	0.035	−0.197	−0.262	−0.113	0.383	−0.203
	Sig. (2-tailed)	0.771	0.97	0.928	0.612	0.495	0.772	0.309	0.601
Trunk xylem	Pearson correlation	0.108	−0.308	−0.553	−0.437	.821**	−0.545	.806**	.902**
	Sig. (2-tailed)	0.783	0.42	0.122	0.239	0.007	0.129	0.009	0.001

Note: Symbol ** and * showed significant correlations at the levels of $p < 0.05$ and $p < 0.01$, respectively, according to two-tailed analysis. The full name of genes was listed in Fig. 5.

< 0.01) had remarkably negative correlation with oleoresin yield. Additionally, *HMGS7* ($r = 0.821$; $p < 0.01$), *CDS1* ($r = 0.806$; $p < 0.01$) and *IDI7* ($r = 0.902$; $p < 0.01$) had highly positive correlation with oleoresin yield in trunk xylem.

We also performed correlation analysis to detect the potential relationships between downstream genes and upstream genes within the three tissues, as shown in Supplementary Table S4, S5, S6. Remarkably, *AACT7* ($r = 0.9634$; $p < 0.01$) had obviously positive correlation with *FPPS2* in young foliage; *HMGR7* ($r = 0.9150$; $p < 0.01$) had highly positive correlation with *ent-CDS9*. In old foliage, *HMGR7* ($r = 0.820$; $p < 0.01$) and *ent-CDS9* ($r = 0.900$; $p < 0.01$) had extremely positive correlations with *LDS10; ent-CDS9* ($r = 0.744$; $p < 0.05$) had significant correlations with *CDS1*. In trunk xylem, *MCS4* ($r = 0.707$; $p < 0.05$), *AACT7* ($r = 0.840$; $p < 0.01$) and *FPPS2* ($r = 0.888$; $p < 0.01$) had clearly correlations with *ent-KO3; IDI7* ($r = 0.685$; $p < 0.05$) had distinctly correlations with *CDS1*.

3.7. Correlations between DRAs and oleoresin yield in 60-year-old tree

The pimaric acid, sandaracopimaric acid, isopimaric acid, palustric acid, levopimaric acid, dehydroabetic acid, abietate acid, and neoabietic acid were the primary components of DRAs in the present study. Similar to the trend of oleoresin yield in nine trees, the accumulation of oleoresin and diterpenoid were highest in No.1 (54.00 g and 414.09 mg g^{-1}

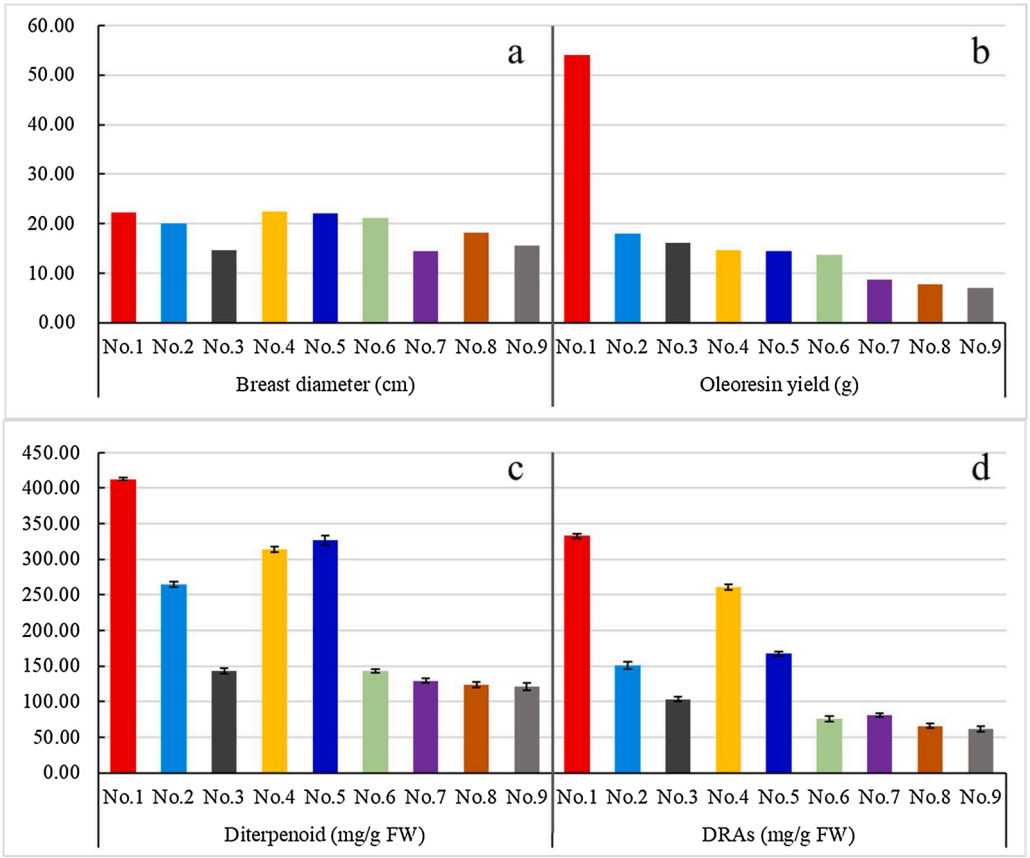

Fig. 10. Content of oleoresin compounds was detected through GC–MS among different-yield of masson pine. (a) The change of breast diameter in nine trees. (b) The oleoresin yield of nine trees within 12 h. (c) The range of diterpenoid content in just 12 h. (d) The relative content of diterpene resin acids (DRAs) in just 12 h. "FW" indicated the fresh weight. Error bars refer to the standard deviation (SD) of three biological repeats.

fresh weight (FW)), the lowest abundance of oleoresin and diterpenoid were observed in No.9 (6.99 g and 121.57 mg g^{-1} FW). Furthermore, DRAs concentrations ranging from 61.50 mg g^{-1} FW in the lowest-yield sample (No.9) to 333.04 mg g^{-1} FW in the highest-yield sample (No.1). The results of correlation showed that the diterpenoid ($r = 0.752$; $p < 0.05$) and DRAs ($r = 0.819$; $p < 0.01$) had significant positive correlations with oleoresin yield. In contrast, the monoterpenes and sesquiterpenes had no significant correlation with oleoresin yield. Additionally, DRAs ($r = 0.942$, $p < 0.01$) had obvious positive correlations with diterpenoid ((Fig. 10, Table 2).

4. Discussion

The yield of oleoresin is a highly heritable trait, and substantial genetic gains can be obtained from the selection of high oleoresin yielders by conventional breeding styles (Allen et al., 2015). However, the process of breeding for high yielders only through the traditional breeding method was so slow. The molecular-assisted selection for the high-oleoresin yield germplasm will give a great future. High-throughput screening (HTS) technology offers new opportunities for discovering and studying the biosynthesis mechanisms of oleoresin compounds in plants. Subsequently, the Illumina RNA-Seq and SMRT-Seq were successfully applied in many *Pinus* species, which will offer a scientific basis for molecular breeding (Niu et al., 2013; Zimin et al., 2014). In the present study, some essential pathways involved in terpenoid biosynthesis were identified, including "Terpenoid backbone biosynthesis" (43 unigenes were obtained in trunk xylem / 52 unigenes in needle) and "Diterpenoids biosynthesis" (22 unigenes in trunk xylem / 13 unigenes in needle). Previous studies illustrated that the diterpenoid was one of the obviously different components between high and low yielder, mostly was the DRAs (Liu et al., 2015). The current results were consistent with the findings of previous studies at the transcriptional level. Furthermore, a total of 13 DEGs including four genes (*AACT*, *HMGS*, *HMGR*, *PMK*) in the MVA pathway, four genes (*DXR*, *MCS*, *HDS, IDI*) from MEP pathway, *FPPS*, and four genes (*ent-CDS, CDS, LDS* and *ent-KO3*) from down-stream diterpenoid biosynthesis pathway were identified as the candidate genes for screening higher oleoresin yield of masson pine. The ultimate product was IPP, which could be isomerized to DMAPP by the effection of *IDI* (Pankratov et al., 2016). The function and role of ent-CDS9 and ent-KO3 are related to the gibberellin biosynthesis. While the CDS1 is a rate-limiting enzyme of DRAs biosynthesis, and LDS is the crucial enzyme of (13E)-Labda-7, 13-dien-15-ol biosynthesis (Alicandri et al., 2020). In short, we obtained several key genes which regulating the diterpenoid metabolic process in high-yield oleoresin masson pine.

Previous reports illustrated that the MVA pathway manufactures cytosolic IPP to synthesize brassinosteroids, sesquiterpenes, sterols, and triterpenes in plants (Sauret-Güeto et al., 2006). In contrast, IPP and DMAPP for the photosynthesis-related isoprenoids biosynthesis were synthesized simultaneously in the MEP pathway, such as monoterpenoids and diterpenoids (Rodríguez-Concepción et al., 2004). Combined with the FPKM and qRT-PCR values in 12-year-old masson pine, the MEP pathway genes had higher expression contents than other MVA pathway genes, including *DXS3*, *DXS6*, *DXR2*, *MCS4*, *HDS4*, but not *DXS8* and *PMK11*. Similarly, the expression patterns in 12-year-old masson pine, especially *MCS4* and *HDS4*, also had higher expression than other genes of precursor biosynthesis in 1-year-old and 60-year-old germplasm. The previous study illustrated that the expression levels of some genes from the MEP pathway, for example, *MCS* and *HDS*, were predominantly up-regulated under high light (Liang et al., 2016). Conversely, when exposed to light, the expression of the MVA pathway genes was down-regulated, and the yield of sterols was reduced (Rodríguez-Concepción, 2006); while up-regulates the expression of MEP pathway genes as well as genes in carotenoid and chlorophyll biosynthesis (Cordoba et al., 2009; Mongélard et al., 2011). Therefore, the MEP pathway plays a more critical role than the MVA pathway in regulating diterpenoid biosynthesis.

Oleoresin productivities can be forecasted using the essential object gene related to the terpenoid biosynthesis as an index; molecular-assisted breeding can provide a valuable backup to conventional direct measurement methods (Chen et al., 2018). Presently, many researchers have been appraised the potential relationships between the transcript expression profiles of secondary metabolites-biosynthesis genes and its accumulation, for example, tanshinone (Yang et al., 2019). However, scarcely efforts about the relations between the gene expression involved in oleoresin biosynthesis pathways and accumulation have been available so far. In this regard, we tried to set up a connection between the expression level of diterpenoids biosynthesis genes and oleoresin yield at a molecular biological level, and the potential relationships between downstream genes and upstream genes of precursor biosynthesis within the three tissues also were performed. We found a significant positive correlation between the expression levels of *ent-CDS9 and HMGR7* and oleoresin yield in the young needle, and *HMGR7* had an extremely significant positive correlation with *ent-CDS9*. In the old needle, only *FPPS2* had a remarkably negative correlation with oleoresin yield, the expressions of *HMGR7* and *ent-CDS9* had extremely substantial positive correlations with *LDS10*, *ent-CDS9* had obviously correlations with *CDS1*. In trunk xylem, *HMGS7*, *CDS1,* and *IDI7* had significant positive correlations with oleoresin yield; *MCS4* and *AACT7* had substantial correlations with *ent-KO3, IDI7* had a significant correlation with *CDS1*. Amazingly, no significant relation between *DXR* or *DXS* expression level and oleoresin yield was found in the present study. DXR and DXS were the critical enzymes regulating metabolic flux in the MEP pathway, some reports showed that the expression levels of the two genes were positive related to the accumulation of isoprenoid (Muñoz-Bertomeu et al., 2006; Kim et al., 2009). Some studies also showed that the regulatory role of DXS and DXR presents species-specific in plants. For instance, the DXR enzyme does not play a critical status or presents a deficiency of direct correlation with carotenoid accumulation in *Lycopersicon esculentum* (Rodríguez-Concepción et al., 2001) or the stock of volatile terpenoids in *Antirrhinum majus* (Dudareva et al., 2005).

Some genes annotated in the diterpenoid biosynthesis pathway were related to gibberellins (GAs) biosynthesis in our research. Especially, the ent-CDS was the critical rate-limiting terpene synthesis enzyme that controlling metabolic flux in the gibberellin biosynthesis (Zi et al., 2014; Bathe and Tissier, 2019). Remarkably, the expression level of *ent-CDS9* in high yielder was significantly higher than in low yielder, and an apparent positive correlation was determined between *ent-CDS9* expression and the oleoresin yield in the young needle. The *ent-CDS9* had strong tissue specificity, and it had a higher expression level in young tissue. The result was consistent with the function and role of GAs. Namely, GAs are conserved diterpenoid compositions, exist in all higher plants as a class of hormones that function in many developmental processes, such as germination, cell elongation and differentiation, and flower and fruit development (Sun and Kamiya, 1994). Additionally, we also obtained the ent-CDS9 protein by Prokaryotic expression, and *in vitro* enzyme activity test showed that the ent-CDS9 protein was active and could react with the substrate (GGPP, GPP,

Table 2
Pearson correlation analysis between the diterpenoid or DRAs content and oleoresin yield.

Types		Oleoresin yield	Diterpenoid
Diterpenoid	Pearson Correlation	0.752*	1
	Sig. (2-tailed)	0.019	
DRAs	Pearson Correlation	0.819**	0.942**
	Sig. (2-tailed)	0.007	0.000

Note: Symbol ** and * showed significant correlations at the levels of $p < 0.05$ and $p < 0.01$, respectively, according to two-tailed analysis. The DRAs represented diterpene resin acids.

FPP) to form some terpenoid products (unpublished data). *Ent-CDS9* had significant positive correlations with *HMGR7* that an upstream pathway gene in the young needle, and *HMGR7* also had a significant positive correlation with oleoresin yield. The conversion of 3-Hydroxy-3-methyl-glutaryl-coenzyme A to mevalonate was catalyzed by HMGR, a key rate-limiting enzyme related to the isoprenoid biosynthesis in the MVA pathway (Enfissi et al., 2005). Research showed a marked increase of monoterpenoids such as 1,8-cineole and camphor by overexpressing *HMGR* in *Lavandula latifolia* (Mendoza-Poudereux et al., 2015). Similarly, overexpression of *HMGR* in *Salvia miltiorrhiza* also accelerated the accumulation of tanshinones, a kind of diterpenoid compound that is supposed to be originated from the MEP pathway (Kai et al., 2011). A reasonable commentary is that the cross-talk between the MVA and MEP pathways has been a reality. For example, reports showed that more IPP synthesized in the plastid is shifted to the cytoplasm to action as a precursor to the MVA pathway (Gutensohn et al., 2013), and the reverse held true as well (Mendoza-Poudereux et al., 2015). Our results also suggested that although GAs shared the upstream biosynthesis pathway, and there was a significant correlation between the *ent-CDS9* gene and the key genes related to another diterpenoid biosynthesis (for example, DRAs), such as *CDS1* and *LDS10*. Nevertheless, the correlation between GAs and DRAs should be further investigated to acquire a better commentary of the metabolic control in different coniferous tissues.

Previous research illustrated that DRAs were detected as the most abundant diterpenoids in oleoresin, and it was a complex mixture by mankind naming (Wang et al., 2013; Whitehill et al., 2019). Generally, the DRAs were mainly composed of abietate acid, dehydroabetic acid, isopimaric acid, levopimaric acid, neoabietic acid, palustric acid, pimaric acid, and sandaracopimaric acid (Hall et al., 2013; Ma et al., 2019; Chen et al., 2021). According to the location in the KEGG pathway, CDS1 is the key rate-limiting enzyme of DRAs biosynthesis. The current results indicated a significant correlation between the *CDS1* expression and the oleoresin yields only in the trunk xylem. In masson pine, the trunk xylem is the central tissue involved in the oleoresin storage, and the needle produces metabolites more relevant to photosynthesis. The upstream gene, *IDI7,* had a significant correlation with *CDS1*. As mentioned previously, diterpenoids are mainly synthesized *via* the MEP pathway in plastids. IDI is used to guarantee an optimal IPP/DMAPP proportion that the muster of the C5 units producing terpenoid precursors (Alicandri et al., 2020). Shortage of *IDI1* decreased the accumulation of carotenoids in cotyledons, flowers, and fruits, but not in ripe foliages in tomato (Pankratov et al., 2016). The diterpenoids and DRAs showed significant positive correlations with oleoresin yield by GC–MS, but monoterpenoids and sesquiterpenoids had no significant correlation with oleoresin yield. However, the expression level of *CDS1* in the 12-year-old high yielder was lower than in the low yielder trees, the gene expression trend in the high yielder of 1-year-old and 60-year--old masson pine was higher. Previously, Liu et al. (2015) reported that the high oleoresin yielder had a more accumulation of sesquiterpene and fewer diterpene content than in low yielder. The oleoresin biosynthesis is a too complicated process affected by some exogenous and endogenous factors, such as different plant tissues, tree age, environmental perturbations, *etc*. Therefore, performing molecular-assisted selection for screening the high-oleoresin yield germplasm, suggesting that the above factors should be considered in the future, as well as further study about these key genes should be paid attention to gain a better interpretation of the metabolic control at gene and protein levels.

5. Conclusions

The combination of transcriptome and metabolism analysis illustrated that the diterpenoid and diterpenoid biosynthesis pathways were the key differences between the high- and low-yield oleoresin germplasm. MEP pathway plays a more pivotal role than the MVA pathway in oleoresin biosynthesis, particularly *MCS* and *HDS*, demonstrating significantly high expression levels in the high-yield ones. Promisingly, oleoresin yield may be possibly evaluated using the expression levels of target genes *HMGR* and *ent-CDS* in the young needle as the candidate indicators instead of only a conventional direct measurement. In the trunk xylem, the *IDI* and *CDS* may also be as the indexes. The above four genes may reliably facilitate the selection for high oleoresin productivity germplasms of masson pine.

Funding

This work was supported by the National Key Research and Development Plan (grant numbers 2017YFD0600303); the Innovation Talent Program of Guizhou Province, P. R. China (grant numbers 2016-4010); and the Postgraduate Research Project of Guizhou Province (grant numbers Qian JiaoHe Postgraduate Research Project Zi YJSCXJH [2018]048]).

CRediT authorship contribution statement

Lina Mei: Methodology, Data curation, Writing - original draft. **Youjin Yan:** Software, Visualization. **Zhengchun Li:** Validation, Investigation. **Jiaxin Ran:** Writing - review & editing. **Luonan Shen:** Resources, Investigation. **Rongju Wu:** Resources, Investigation. **Qiandong Hou:** Writing - review & editing. **Tianjiao Shen:** Validation. **Xiaopeng Wen:** Conceptualization, Writing - review & editing. **Zhangqi Yang:** Resources. **Yuanheng Feng:** Resources.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

Thanks to Prof. Xinlong Dai (College of Tea, Guizhou University) for help in the GC–MS detection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.indcrop.2021.113827.

References

Alicandri, E., Paolacci, A.R., Osadolor, S., Sorgonà, A., Badiani, M., Ciaffi, M., 2020. On the evolution and functional diversity of terpene synthases in the *Pinus* Species: A Review. J. Mol. Evol. 88, 253–283. https://doi.org/10.1007/s00239-020-09930-8.

Allen, S., Ginwal, H.S., Barthwal, S., 2015. Development of PCR based markers in Terpene synthase genes for marker assisted selection of high resin yielders in *Pinus roxburghii* Sarg. Silvae Genet. 64, 211–220. https://doi.org/10.1515/sg-2015-0020.

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11 (10), R106. https://doi.org/10.1186/gb-2010-11-10-r106.

Barrero, R.A., Chapman, B., Yang, Y., Moolhuijzen, P., Keeble-Gagnère, G., Zhang, N., Tang, Q., Bellgard, M.I., Qiu, D., 2011. De novo assembly of *Euphorbia fischeriana* root transcriptome identifies prostratin pathway related genes. BMC Genomics 12, 600. https://doi.org/10.1186/1471-2164-12-600.

Bathe, U., Tissier, A., 2019. Cytochrome P450 enzymes: a driving force of plant diterpene diversity. Phytochemistry 161, 149–162. https://doi.org/10.1016/j.phytochem.2018.12.003.

Burdon, R.D., Wilcox, P.L., 2007. Population management: potential impacts of advances in genomics. New For. 34, 187–206. https://doi.org/10.1007/s11056-007-9047-6.

Chen, B., Xiao, Y., Li, J., Liu, H., Chen, H., Jia, J., Chao, N., Gai, Y., Jiang, X., 2018. Cloning and characterization of geranylgeranyl diphosphate synthetase from *Pinus massoniana* and its correlation with resin productivity. J. For. Res. 29, 311–320. https://doi.org/10.1007/s11676-017-0443-2.

Chen, R., Huang, K., Pan, S., Xu, T., Tan, J., Hao, D., 2021. Jasmonate induced terpene-based defense in *Pinus massoniana* depresses Monochamus alternatus adult feeding. Pest Manag. Sci. 77, 731–740. https://doi.org/10.1002/ps.6068.

Conesa, A., Götz, S., 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int. J. Plant Genomics, 619832. https://doi.org/10.1155/2008/619832.

Cordoba, E., Salmi, M., León, P., 2009. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. J. Experiment. Botany 60, 2933–2943. https://doi.org/10.1093/jxb/erp190.

Dudareva, N., Andersson, S., Orlova, I., Gatto, N., Reichelt, M., Rhodes, D., Boland, W., Gershenzon, J., 2005. The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. Proc. Natl. Acad. Sci. U.S.A. 102, 933–938. https://doi.org/10.1073/pnas.0407360102.

Enfissi, E.M.A., Fraser, P.D., Lois, L.M., Boronat, A., Schuch, W., Bramley, P.M., 2005. Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. Plant Biotechnol. J. 3, 17–27. https://doi.org/10.1111/j.1467-7652.2004.00091.x.

Gutensohn, M., Orlova, I., Nguyen, T.T.H., Davidovich-Rikanati, R., Ferruzzi, M.G., Sitrit, Y., Lewinsohn, E., Pichersky, E., Dudareva, N., 2013. Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. Plant J. 75, 351–363. https://doi.org/10.1111/tpj.12212.

Hall, D.E., Zerbe, P., Jancsik, S., Quesada, A.L., Dullat, H., Madilao, L.L., Yuen, M., Bohlmann, J., 2013. Evolution of conifer diterpene synthases: diterpene resin acid biosynthesis in lodgepole pine and jack pine involves monofunctional and bifunctional diterpene synthases. Plant Physiol. 161, 600–616. https://doi.org/10.1104/pp.112.208546.

Hoang, N.V., Furtado, A., Mason, P.J., Marquardt, A., Kasirajan, L., Thirugnanasambandam, P.P., Botha, F.C., Henry, R.J., 2017. A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing. BMC Genomics 18, 395. https://doi.org/10.1186/s12864-017-3757-8.

Huang, J., Rücker, A., Schmidt, A., Gleixner, G., Gershenzon, J., Trumbore, S., Hartmann, H., 2020. Production of constitutive and induced secondary metabolites is coordinated with growth and storage in Norway spruce saplings. Tree Physiol. 40 https://doi.org/10.1093/treephys/tpaa040.

Junkes, C.F.de O., Duz, J.V.V., Kerber, M.R., Wieczorek, J., Galvan, J.L., Fett, J.P., Fett-Neto, A.G., 2019. Resinosis of young slash pine (*Pinus elliottii* Engelm.) as a tool for resin stimulant paste development and high yield individual selection. Ind. Crops Prod. 135, 179–187. https://doi.org/10.1016/j.indcrop.2019.04.048.

Kai, G., Xu, H., Zhou, C., Liao, P., Xiao, J., Luo, X., You, L., Zhang, L., 2011. Metabolic engineering tanshinone biosynthetic pathway in *Salvia miltiorrhiza* hairy root cultures. Metab. Eng. 13, 319–327. https://doi.org/10.1016/j.ymben.2011.02.003.

Keeling, C.I., Bohlmann, J., 2006. Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. New Phytol. 170, 657–675. https://doi.org/10.1111/j.1469-8137.2006.01716.x.

Kelkar, V.M., Geils, B.W., Becker, D.R., Overby, S.T., Neary, D.G., 2006. How to recover more value from small pine trees: essential oils and resins. Biomass Bioenergy 30, 316–320. https://doi.org/10.1016/j.biombioe.2005.07.009.

Kim, Y.B., Kim, S.M., Kang, M.K., Kuzuyama, T., Lee, J.K., Park, S.C., Shin, S.C., Kim, S. U., 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, 737–749. https://doi.org/10.1093/treephys/tpp002.

Knebel, L., Robison, D.J., Wentworth, T.R., Klepzig, K.D., 2008. Resin flow responses to fertilization, wounding and fungal inoculation in loblolly pine *(Pinus taeda*) in North Carolina. Tree Physiol. 28, 847–853. https://doi.org/10.1093/treephys/28.6.847.

Li, B., Dewey, C.N., 2014. RSEM: Accurate transcript quantification from RNA-seq data with or without a reference genome. In: Bioinformatics: The Impact of Accurate Quantification on Proteomic and Genetic Analysis and Research, 12, p. 323. https://doi.org/10.1201/b16589 (1).

Liang, C., Zhang, W., Zhang, X., Fan, X., Xu, D., Ye, N., Su, Z., Yu, J., Yang, Q., 2016. Isolation and expression analyses of methyl-d-erythritol 4-phosphate (MEP) pathway genes from *Haematococcus pluvialis*. J. Appl. Phycol. 28, 209–218. https://doi.org/10.1007/s10811-015-0604-7.

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, e0132624. https://doi.org/10.1371/journal.pone.0132624.

Liu, Q., Xie, Y., Liu, B., Huanhuanyin, Zhou, Z., Feng, Z., Chen, Y., 2020. A transcriptomic variation map provides insights into the genetic basis of *Pinus massoniana* Lamb. evolution and the association with oleoresin yield. BMC Plant Biol. 20 (1) https://doi.org/10.1186/s12870-020-02577-z.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25 (4), 402–408. https://doi.org/10.1006/meth.2001.1262.

Lombardero, M.J., Ayres, M.P., Lorio, P.L., Ruel, J.J., 2000. Environmental effects on constitutive and inducible resin defences of *Pinus taeda*. Ecol. Lett. 3, 329–339. https://doi.org/10.1046/j.1461-0248.2000.00163.x.

Ma, L.T., Lee, Y.R., Tsao, N.W., Wang, S.Y., Zerbe, P., Chu, F.H., 2019. Biochemical characterization of diterpene synthases of *Taiwania cryptomerioides* expands the known functional space of specialized diterpene metabolism in gymnosperms. Plant J. 100, 1254–1272. https://doi.org/10.1111/tpj.14513.

Mao, J., He, Z., Hao, J., Liu, T., Chen, J., Huang, S., 2019. Identification, expression, and phylogenetic analyses of terpenoid biosynthesis-related genes in secondary xylem of loblolly pine (*Pinus taeda* L.) based on transcriptome analyses. PeerJ 7, e6124. https://doi.org/10.7717/peerj.6124.

Mei, L., Li, Z., Yan, Y., Wen, Z., Wen, X., Yang, Z., Feng, Y., 2020. Identification and functional study of oleoresin terpenoid biosynthesis-related genes in masson pine (*Pinus massoniana* L.) based on transcriptome analysis. Tree Genet. Genomes 16, 53. https://doi.org/10.1007/s11295-020-01448-w.

Mendoza-Poudereux, I., Kutzner, E., Huber, C., Segura, J., Eisenreich, W., Arrillaga, I., 2015. Metabolic cross-talk between pathways of terpenoid backbone biosynthesis in *spike lavender*. Plant Physiol. Biochem. 95, 113–120. https://doi.org/10.1016/j.plaphy.2015.07.029.

Mongélard, G., Seemann, M., Boisson, A.M., Rohmer, M., Bligny, R., Rivasseau, C., 2011. Measurement of carbon flux through the MEP pathway for isoprenoid synthesis by 31P-NMR spectroscopy after specific inhibition of 2-C-methyl-d-erythritol 2,4-cyclodiphosphate reductase. Effect of light and temperature. Plant Cell Environ. 34, 1241–1248. https://doi.org/10.1111/j.1365-3040.2011.02322.x.

Moreira, X., Mooney, K.A., Rasmann, S., Petry, W.K., Carrillo-Gavilán, A., Zas, R., Sampedro, L., 2014. Trade-offs between constitutive and induced defences drive geographical and climatic clines in pine chemical defences. Ecol. Lett. 17, 537–546. https://doi.org/10.1111/ele.12253.

Muñoz-Bertomeu, J., Arrillaga, I., Ros, R., Segura, J., 2006. Up-regulation of 1-Deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic *spike lavender*. Plant Physiol. 142, 890–900. https://doi.org/10.1104/pp.106.086355.

Neis, F.A., de Costa, F., de Araújo, A.T., Fett,, J.P., Fett-Neto, A.G., 2019. Multiple industrial uses of non-wood pine products. Ind. Crops Prod. 130, 248–258. https://doi.org/10.1016/j.indcrop.2018.12.088.

Niu, S.H., Li, Z.X., Yuan, H.W., Chen, X.Y., Li, Y., Li, W., 2013. Transcriptome characterisation of *Pinus tabuliformis* and evolution of genes in the *Pinus* phylogeny. BMC Genomics 14 (1), 263. https://doi.org/10.1186/1471-2164-14-263.

Ott, D.S., Davis, T.S., Mercado, J.E., 2021. Interspecific variation in spruce constitutive and induced defenses in response to a bark beetle fungal symbiont provides insight into traits associated with resistance. Tree Physiol. 41 (7), 1109–1121. https://doi.org/10.1093/treephys/tpaa170.

Pankratov, I., McQuinn, R., Schwartz, J., Bar, E., Fei, Z., Lewinsohn, E., Zamir, D., Giovannoni, J.J., Hirschberg, J., 2016. Fruit carotenoid-deficient mutants in tomato reveal a function of the plastidial isopentenyl diphosphate isomerase (*IDI1*) in carotenoid biosynthesis. Plant J. 88, 82–94. https://doi.org/10.1111/tpj.13232.

Peters, R.J., 2010. Two rings in them all: the labdane-related diterpenoids. Nat. Prod. Rep. 27, 1521–1530. https://doi.org/10.1039/c0np00019a.

Phillips, M.A., Bohlmann, J., Gershenzon, J., 2006. Molecular regulation of induced terpenoid biosynthesis in conifers. Phytochem. Rev. 5, 179. https://doi.org/10.1007/s11101-006-0001-6.

Roberds, J.H., Strom, B.L., Hain, F.P., Gwaze, D.P., McKeand, S.E., Lott, L.H., 2003. Estimates of genetic parameters for oleoresin and growth traits in juvenile loblolly pine. Can. J. For. Res. 33, 2469–2476. https://doi.org/10.1139/x03-186.

Rodrigues-Corrêa, K.C.da S., de Lima, J.C., Fett-Neto, A.G., 2012. Pine oleoresin: tapping green chemicals, biofuels, food protection, and carbon sequestration from multipurpose trees. Food Energy Secur. 1 (2), 81–93. https://doi.org/10.1002/fes3.13.

Rodríguez-Concepción, M., 2006. Early steps in isoprenoid biosynthesis: multilevel regulation of the supply of common precursors in plant cells. Phytochem. Rev. 5, 1–15. https://doi.org/10.1007/s11101-005-3130-4.

Rodríguez-Concepción, M., Ahumada, I., Diez-Juez, E., Sauret-Güeto, S., María Lois, L., Gallego, F., Carretero-Paulet, L., Campos, N., Boronat, A., 2001. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. Plant J. 27, 213–222. https://doi.org/10.1046/j.1365-313X.2001.01089.x.

Rodríguez-Concepción, M., Forés, O., Martínez-García, J.F., González, V., Phillips, M.A., Ferrer, A., Boronat, A., 2004. Distinct light-mediated pathways regulate the biosynthesis and exchange of isoprenoid precursors during *Arabidopsis* seedling development. Plant Cell 16, 144–156. https://doi.org/10.1105/tpc.016204.

Rodríguez-García, A., López, R., Martín, J.A., Pinillos, F., Gil, L., 2014. Resin yield in *Pinus pinaster* is related to tree dendrometry, stand density and tapping-induced systemic changes in xylem anatomy. For. Ecol. Manage. 313, 47–54. https://doi.org/10.1016/j.foreco.2013.10.038.

Sauret-Güeto, S., Botella-Pavía, P., Flores-Pérez, U., Martínez-García, J.F., San Román, C., León, P., Boronat, A., Rodríguez-Concepción, M., 2006. Plastid cues posttranscriptionally regulate the accumulation of key enzymes of the methylerythritol phosphate pathway in *arabidopsis*. Plant Physiol. 141, 75–84. https://doi.org/10.1104/pp.106.079855.

Sun, Taiping, Kamiya, Y., 1994. The *Arabidopsis* GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. Plant Cell 6, 1509–1527. https://doi.org/10.1105/tpc.6.10.1509.

Tadesse, W., Nanos, N., Auñon, F.J., Alía, R., Gil, L., 2002. Evaluation of high resin yielders of *Pinus pinaster* Ait. For. Genet. 8, 4271–4278.

Tholl, D., 2015. Biosynthesis and biological functions of terpenoids in plants. Adv. Biochem. Eng. Biotechnol. 63–106. https://doi.org/10.1007/10_2014_295.

Vázquez-González, C., Zas, R., Erbilgin, N., Ferrenberg, S., Rozas, V., Sampedro, L., 2020. Resin ducts as resistance traits in conifers: linking dendrochronology and resin-based defences. Tree Physiol. Tpaa 064. https://doi.org/10.1093/treephys/tpaa064.

Wang, Y., Lim, L., Diguistini, S., Robertson, G., Bohlmann, J., Breuil, C., 2013. A specialized ABC efflux transporter GcABC-G1 confers monoterpene resistance to *Grosmannia clavigera*, a bark beetle-associated fungal pathogen of pine trees. New Phytol. 197, 886–898. https://doi.org/10.1111/nph.12063.

Wen, L., Shi, R., Wang, J., Zhao, Y., Zhang, H., Ling, X., Xiong, Z., 2018. Transcriptome analyses to reveal genes involved in terpene biosynthesis in resin producing pine tree *Pinus kesiya* var. langbianensis. BioResources 13, 1852–1871. https://doi.org/10.15376/biores.13.1.1852-1871.

Westbrook, J.W., Resende, M.F.R., Munoz, P., Walker, A.R., Wegrzyn, J.L., Nelson, C.D., Neale, D.B., Kirst, M., Huber, D.A., Gezan, S.A., Peter, G.F., Davis, J.M., 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, 89–100. https://doi.org/10.1111/nph.12240.

Whitehill, J.G.A., Yuen, M.M.S., Henderson, H., Madilao, L., Kshatriya, K., Bryan, J., Jaquish, B., Bohlmann, J., 2019. Functions of stone cells and oleoresin terpenes in

the conifer defense syndrome. New Phytol. 221, 1503–1517. https://doi.org/10.1111/nph.15477.
Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., Xu, H., Huang, X., Li, S., Zhou, A., Zhang, X., Bolund, L., Chen, Q., Wang, J., Yang, H., Fang, L., Shi, C., 2018. WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Res. 46 (W1), W71–W75. https://doi.org/10.1093/nar/gky400.
Yu, G., Wang, L.G., Han, Y., He, Q.Y., 2012. ClusterProfiler: an R package for comparing biological themes among gene clusters. Omics A J. Integr. Biol. 16, 284–287. https://doi.org/10.1089/omi.2011.0118.
Zeng, L.H., Zhang, Q., He, B.X., Lian, H.M., Cai, Y.L., Wang, Y.S., Luo, M., 2013. Age trends in genetic parameters for growth and resin-yielding capacity in masson pine. Silvae Genet. 62, 7–18. https://doi.org/10.1515/sg-2013-0002.
Zerbe, P., Bohlmann, J., 2015. Plant diterpene synthases: exploring modularity and metabolic diversity for bioengineering. Trends Biotechnol. 33 (7), 419–428. https://doi.org/10.1016/j.tibtech.2015.04.006.
Zhang, K., Fan, W., Huang, Z., Chen, D., Yao, Z., Li, Y., Yang, Y., Qiu, D., 2019. Transcriptome analysis identifies novel responses and potential regulatory genes involved in 12-deoxyphorbol-13-phenylacetate biosynthesis of *Euphorbia resinifera*. Ind. Crops Prod. 135, 138–145. https://doi.org/10.1016/j.indcrop.2019.04.030.
Zi, J., Mafu, S., Peters, R.J., 2014. To gibberellins and beyond! Surveying the evolution of (Di)terpenoid metabolism. Annu. Rev. Plant Biol. 65, 259–286. https://doi.org/10.1146/annurev-arplant-050213-035705.
Zimin, A., Stevens, K.A., Crepeau, M.W., Holtz-Morris, A., Koriabine, M., Marçais, G., Puiu, D., Roberts, M., Wegrzyn, J.L., de Jong, P.J., Neale, D.B., Salzberg, S.L., Yorke, J.A., Langley, C.H., 2014. Sequencing and assembly of the 22-Gb loblolly pine genome. Genetics 196 (3), 875–890. https://doi.org/10.1534/genetics.113.159715.
Züst, T., Agrawal, A.A., 2017. Trade-offs between plant growth and defense against insect herbivory: an emerging mechanistic synthesis. Annu. Rev. Plant Biol. 68, 513–534. https://doi.org/10.1146/annurev-arplant-042916-040856.