

RESEARCH ARTICLE

Open Access

# Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria sinensis*

Yanhong Xu<sup>1†</sup>, Zheng Zhang<sup>1,2†</sup>, Mengxi Wang<sup>3†</sup>, Jianhe Wei<sup>1,2\*</sup>, Hongjiang Chen<sup>2</sup>, Zhihui Gao<sup>1</sup>, Chun Sui<sup>1</sup>, Hongmei Luo<sup>1</sup>, Xingli Zhang<sup>1</sup>, Yun Yang<sup>2</sup>, Hui Meng<sup>2</sup> and Wenlan Li<sup>3</sup>

## Abstract

**Background:** Agarwood is an expensive resinous heartwood derived from *Aquilaria* plants that is widely used in traditional medicines, incense and perfume. Only wounded trees can produce agarwood, and the huge demand for the agarwood products has led all *Aquilaria* spp. being endangered and listed in the Appendix II of the CITES (<http://www.cites.org>). The major components of agarwood are sesquiterpenes and phenylethyl chromones. Owing to a lack of genomic information, the molecular basis of wound-induced sesquiterpenes biosynthesis and agarwood formation remains unknown.

**Results:** To identify the primary genes that maybe related to agarwood formation, we sequenced 2 cDNA libraries generated from healthy and wounded *A. sinensis* (Lour.) Gilg. A total of 89,137 unigenes with an average length of 678.65 bp were obtained, and they were annotated in detail at bioinformatics levels. Of those associated with agarwood formation, 30 putatively encoded enzymes in the sesquiterpene biosynthesis pathway, and a handful of transcription factors and protein kinases were related to wound signal transduction. Three full-length cDNAs of sesquiterpene synthases (*ASS1-3*) were cloned and expressed in *Escherichia coli*, and enzyme assays revealed that they are active enzymes, with the major products being  $\delta$ -guaiene. A methyl jasmonate (MJ) induction experiment revealed that the expression of *ASS* was significantly induced by MJ, and the production of sesquiterpenes was elevated accordingly. The expression of some transcription factors and protein kinases, especially *MYB4*, *WRKY4*, *MPKK2* and *MAPK2*, was also induced by MJ and coordinated with *ASS* expression, suggesting they maybe positive regulators of *ASS*.

**Conclusions:** This study provides extensive transcriptome information for *Aquilaria* spp. and valuable clues for elucidating the mechanism of wound-induced agarwood sesquiterpenes biosynthesis and their regulation.

**Keywords:** Agarwood, *Aquilaria sinensis*, GC-MS, Sesquiterpenes, Transcriptome, Wound signal transduction

## Background

Agarwood is widely used in traditional medicines as a digestive, sedative, and antiemetic drug and is also popular as incense and perfume in the Middle East, South Asia,

Japan, and China. Additionally, the agarwood sculpturing for interior decoration is another important aspect of its value which generates a lot of income in Asia. In the international market, high-quality agarwood is more costly than gold. Agarwood is a dark, resinous, non-timber wood that forms in the stem, branch, or root of *Aquilaria* and *Gyrinops* trees after they are wounded and infected by a fungus (under natural conditions, the wounds can be caused by wind, lightning strikes, the gnawing of ants or insects, or microorganism invasion), but these natural processes always develop very slowly over decades. Owing to the economic value and great

\* Correspondence: [wjianh@263.net](mailto:wjianh@263.net)

†Equal contributors

<sup>1</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 151, Malianwa North Road, Haidian District, Beijing 100193, China

<sup>2</sup>Hainan Branch Institute of Medicinal Plant (Hainan Provincial Key Laboratory of Resources Conservation and Development of Southern Medicine), Chinese Academy of Medical Sciences & Peking Union Medical College, Wanning 571533, China

Full list of author information is available at the end of the article

demand of agarwood, natural *Aquilaria* forests have been destroyed in almost all countries in which agarwood is commercially exploited. For the protection of wild *Aquilaria* resources and their sustainable use, all *Aquilaria* spp. have been listed in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (<http://www.cites.org> [accessed 7 August 2012]), and *Aquilaria* cultivation has been drawn much attention in countries such as China, India, Vietnam, Indonesia, Malaysia, and Thailand. However, the current methods used by farmers, including partial trunk pruning, burn-chisel-drill, and fungi inoculation, all require considerable time to produce agarwood and result in a product with very low yield and quality.

*Aquilaria sinensis* (Lour.) Gilg is one of the most important plant resources for producing agarwood in China as well as the only certified source for agarwood listed in *China Pharmacopoeia* [1]. More than 20 million planted *A. sinensis* trees are estimated to be scattered throughout Hainan, Guangdong, and Yunnan Provinces, and more than one-fourth of them are older than 5 years and have become available for agarwood induction. Uncovering the mechanism of agarwood formation in wounded trees is crucial to establishing an efficient induction method for agarwood.

Studies have shown that sesquiterpenes and phenylethyl chromone derivatives are the main compounds in agarwood [2-7]. Thus, understanding the biosynthesis and regulation of sesquiterpenes and chromone in *Aquilaria* spp. is critically important in determining the mechanism of agarwood formation. However, to date, the biosynthesis pathway of chromone derivatives remains almost unknown. The terpenoid metabolism pathway is, by comparison, very clear. The biosynthesis of sesquiterpenes can reportedly occur via the mevalonic acid (MVA) [8,9] and 1-deoxy-D-xylulose-5-phosphate (DXP) [10] pathways (Additional file 1: Figure S1), in which sesquiterpene synthases are enzymes used in the final step to form sesquiterpenes. Currently, many studies on the functional identification and regulation of these enzymes have been reported. In *Gossypium arboreum*, GaWRKY1 positively regulates the expression of (+)- $\delta$ -cadinene synthase, which catalyzes the branch point leading to biosynthesis of sesquiterpene gossypol [11]. AaWRKY1 regulates the expression of amorpho-4,11-diene synthase (ADS), which is the key artemisinin biosynthesis enzyme in *Artemisia annua*. Overexpression of AaWRKY1 in tobacco or in *A. annua* can significantly activate ADS promoter activity, and artemisinin production increases accordingly [12]. The latest report on *Arabidopsis* has shown that transcription factor (TF) MYC2 interacts with DELLA proteins, positively regulating the expression of sesquiterpene synthases TPS21 and

TPS11 in the jasmonic acid (JA) and gibberellic acid (GA) signaling pathways [13]. However, studies on agarwood sesquiterpenes in *Aquilaria* spp. are still extremely limited, and with the exception of five  $\delta$ -guaiene synthase genes (*AcC1*, *AcC2*, *AcC3*, *AcC4*, and *AcL154*), which were cloned from cell cultures of *A. crassna* [5], most of the genes remain to be cloned. Kumeta & Ito [5] have also found that methyl jasmonate (MJ) treatment increases the expression of these genes and induces the production of sesquiterpene  $\delta$ -guaiene. Their report was the first and remains the only molecular report of sesquiterpene biosynthesis and regulation in *Aquilaria* spp..

Agarwood is produced after an *Aquilaria* tree is wounded [14-17]. We have observed that in the early period of agarwood formation, the resin in agarwood exists in 2 forms: tylose and gel (our unpublished data), and the sesquiterpenes in agarwood have antimicrobial and anti-disease activity [6]. Based on previous studies, we have hypothesized that agarwood is the product of a plant defense response [18]: when an *Aquilaria* tree is wounded, damage signals are induced and transmitted, activating the defense response, after which defensive substances such as sesquiterpenes and phenylethyl chromone derivatives are produced. When these products combine with the wood tissue to avoid damage expansion, agarwood forms. The key to testing this hypothesis is to reveal the mechanism of wound-induced biosynthesis of agarwood sesquiterpenes, but this goal is hindered by the very limited amount of genomic information available for *Aquilaria*. The present study made use of 454 sequencing technology in order to (i) provide detailed insight into the transcriptomes of healthy and wounded *A. sinensis*, (ii) discover candidate transcripts with significant homology to important enzymes in the sesquiterpene biosynthesis pathway and verify their function, and (iii) identify candidate regulators that might be related to wound signal transduction and regulation of agarwood formation.

## Results

### Transcriptome sequence assembly and annotation

Two cDNA libraries from the stems of the healthy and wounded *A. sinensis* (designated H and W) were constructed using Switching Mechanism At 5' end of RNA Transcript (SMART) technology. The 2 libraries were sequenced in one run using 454 GS-FLX pyrosequencing technology. A total of 562,060 (H) and 565,008 (W) raw reads were generated (Table 1). After the removal of sequences shorter than 100 bp and empty vectors, 1,073,469 high-quality reads [NCBI GEO, accession no. GSE42246], including 530,438 (94.37% of the raw reads) from the H library and 543,031 (96.11% of the raw reads) from the W library, were obtained. After *de novo* assembly of these reads, 67,972 unigenes from

**Table 1 Characteristics of the reads from the healthy and the wounded *A. sinensis***

Numbers	Healthy	Wounded
Total reads	562,060	565,008
Vector reads <sup>a</sup>	2,698	2,095
Low quality reads <sup>b</sup>	28,924	19,882
High quality reads <sup>c</sup>	530,438	543,031
Contigs <sup>d</sup>	66,207	71,657
Singlets <sup>e</sup>	1,765	2,482
Unigenes <sup>f</sup>	67,972	74,139

<sup>a</sup> Vector reads: The reads sequences are all vector sequences.

<sup>b</sup> Low quality reads: After removing empty reads and vector reads, the length of reads is less than the cutoff length (100 bp).

<sup>c</sup> High quality reads: After removing empty reads and vector reads, the length of reads is more than the cutoff length (100 bp).

<sup>d</sup> Contigs: reads number  $\geq 2$ .

<sup>e</sup> Singlets: reads = 1.

<sup>f</sup> Unigenes = Contigs + Singlets.

the H library and 74,139 unigenes from the W library were generated, respectively (Table 1). Combining the reads from the 2 libraries, we obtained 89,137 unigenes with an average length of 678.65 bp. More than 70% of these unigenes were between 200 bp and 600 bp long (Figure 1), and almost 17% contained more than 10 reads (Additional file 2: Table S1).

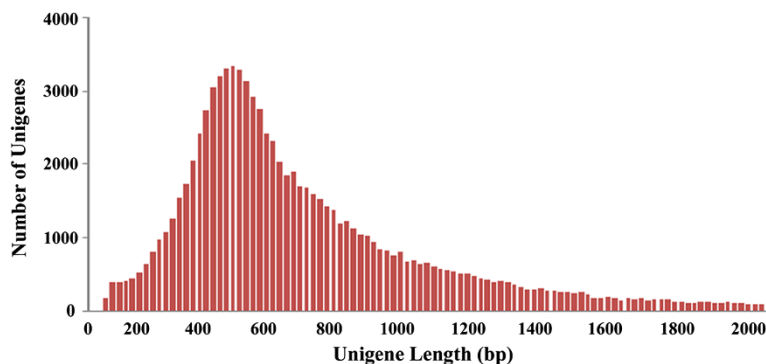
The BLAST annotation of the 89,137 unigenes indicated that 22,051 (24.7%) shared homology with the sequences in the NCBI nucleotide database, and 38,159 (42.8%) had matches to known protein sequences in the NCBI non-redundant database. To obtain more information, we also carried out the BLASTX search against the SWISS-PROT protein database, and 17.1% of the unigenes were thus annotated (Additional file 3: Table S2), indicating that the information about the genomes and transcriptomes of *Aquilaria* spp. was very limited. Thus, our 454 database is important, and the unigene sequences obtained in this study may represent most of the novel genes of *Aquilaria* spp..

Based on an InterPro analysis, 48,530 (54%) unigenes were annotated with at least one GO term, and the distribution of gene functions based on the cellular component, molecular function, and biological process categories was very similar in the 2 libraries (Additional file 4: Figure S2). KEGG annotation showed that, in the metabolism process, although the W library had a total number of unigenes smaller than that of the H library, it had a slightly higher number of unigenes related to carbohydrate metabolism and metabolism of terpenoids and polyketides (Figure 2). This result indicated that the pruning wound in *A. sinensis* had activated carbohydrate, terpenoid, and polyketides metabolism, which may be related to agarwood formation.

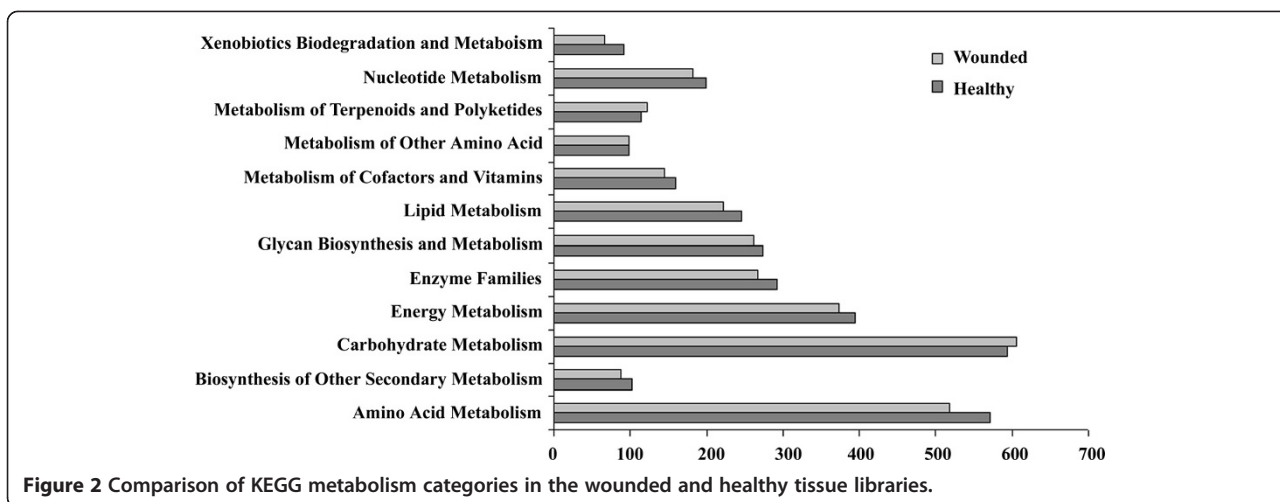
To investigate the differences between the 2 libraries, IDEG6 [19] was used to identify differentially expressed genes (DEGs). A total of 47,169 DEGs were obtained, including 14,988 and 21,165 expressed uniquely in the H and W libraries, respectively; 11,016 were expressed in both libraries but at different levels (Figure 3). KEGG classification showed that the number of “up” unigenes was far greater than the number of “down” unigenes (the relative abundance of the W library vs. the H library:  $>1$  marked “up”,  $<1$  marked “down”; Figure 4); for metabolism and genetic information processing, 814 (82.5%) and 652 (83.8%) of the DEGs were “up”, respectively. The percentage and distribution of DEGs annotated with GO terms are shown in Additional file 5: Figure S3.

#### Enzyme genes involved in sesquiterpenes biosynthesis

The biosynthesis of sesquiterpenes occurs via the MVA [8,9] and DXP [10] pathways, and 4 key enzymes are commonly involved: 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR), the first key enzyme in the MVP pathway; DXP synthase (DXPS), the first rate-limiting enzyme in the DXP pathway; farnesyl diphosphate (FPP) synthase (FPS); and sesquiterpenes synthases. In our 454 unigene database (H + W), 30



**Figure 1** Length distribution for *Aquilaria sinensis* unigenes. Over 70% of these unigenes are between 200 bp and 600 bp long.



unigenes were annotated as being related to the sesquiterpene biosynthesis pathway, and 22 of them putatively encode these enzymes, including 3 HMGRs, 12 DXPSs, 2 FPSs, and 5 sesquiterpene synthases (Table 2). Unigenes putatively encoding other enzymes in the sesquiterpene biosynthesis pathway were also found, including acetyl-coenzyme A acyltransferase (ACTC), mevalonate kinase (MK), phosphomevalonate kinase (PMK), and 2-C-methyl-D-erythritol-2, 4-cyclodiphosphate synthase (MCS) (see Table 2).

FPS is a branch point enzyme in terpenoid biosynthesis [20-22]. Its product, FPP, is the common precursor of sesquiterpenes, steroids, and farnesylated proteins. We found two unigenes that putatively encode FPS. One is a full-length cDNA (named *AsFPS*) that has an open reading frame of 1029 nucleotides and encodes a protein of 342 amino acids. The phenetic relationship between *AsFPS* and the FPSs characterized in other species is depicted in Additional file 6: Figure S4. *AsFPS* had the highest homology with the FPS sequence from *Aquilaria microcarpa* (GenBank ID: ADH95185; 97%

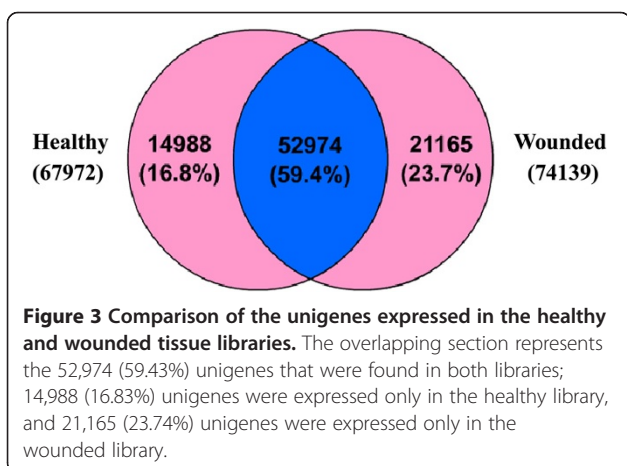
identity, with a difference of 5 amino acids), suggesting that *AsFPS* belongs to the FPS family.

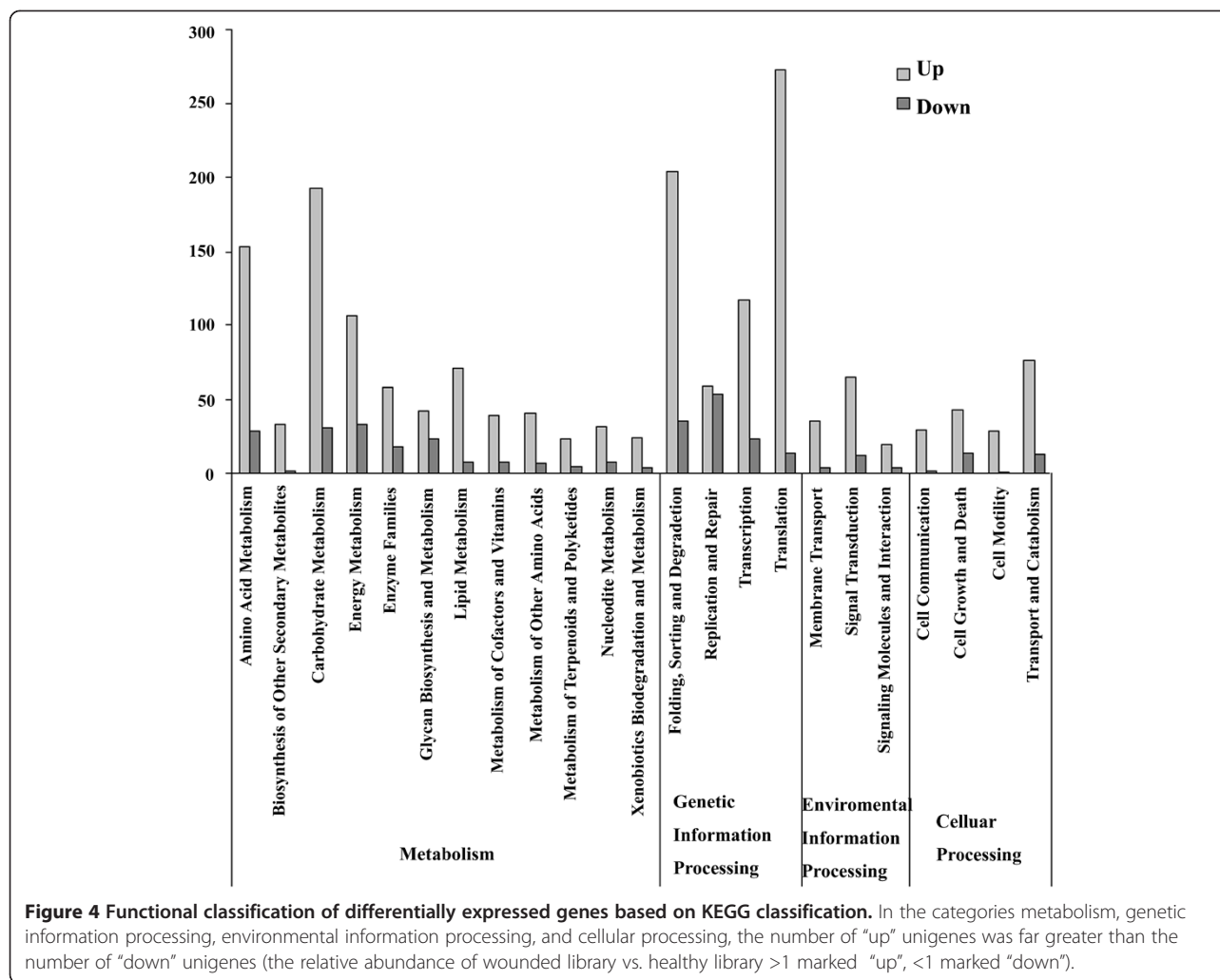
With the exception of one FPS, the full-length open-reading frames (ORFs) of other unigenes were unavailable. Sesquiterpene synthases are the key enzymes that catalyze FPP, leading to sesquiterpenes biosynthesis. Therefore, they were the focus of this study.

#### Cloning and functional identification of *A. quilaria* sesquiterpene synthase genes (*ASSs*)

Based on the degenerated primers, 3 full-length cDNAs of sesquiterpenes synthases (*ASS1*, *ASS2*, and *ASS3*) were cloned and deposited in the NCBI database under the accession numbers JQ712682, JQ712683, and JQ712684. They have ORFs of 1644 nucleotides, which encode almost the same protein of 547 amino acids. Their nucleotide and protein sequences have more than 92% identity with one another. The alignment of their nucleotide and deduced amino acid sequences are shown in Additional file 7: Figures S5 and Additional file 8: Figures S6, respectively. These proteins contain 2 motifs (the RRx<sub>8</sub>W motif at the N-terminus and the DDxxD motif known to be a divalent metal ion substrate-binding site) that are functionally important and highly conserved in all terpene synthase proteins. The phenetic relationship between the 3 putative proteins and other characterized sesquiterpenes synthases is shown in Additional file 9: Figure S7. *ASS1-3* were grouped with ACY38194, ACY38195, ACY38196, ACY38197 [5], AEG77018, AEG77021, and AEG77019 [23], which were cloned from *A. crassna*.

To confirm that *ASS1-3* encode active *ASSs*, we cloned them into a pET-28a vector and heterologously expressed them in *E. coli*. The fusion proteins were detected in soluble fractions (data not shown). The enzyme assays were performed using the geranyl



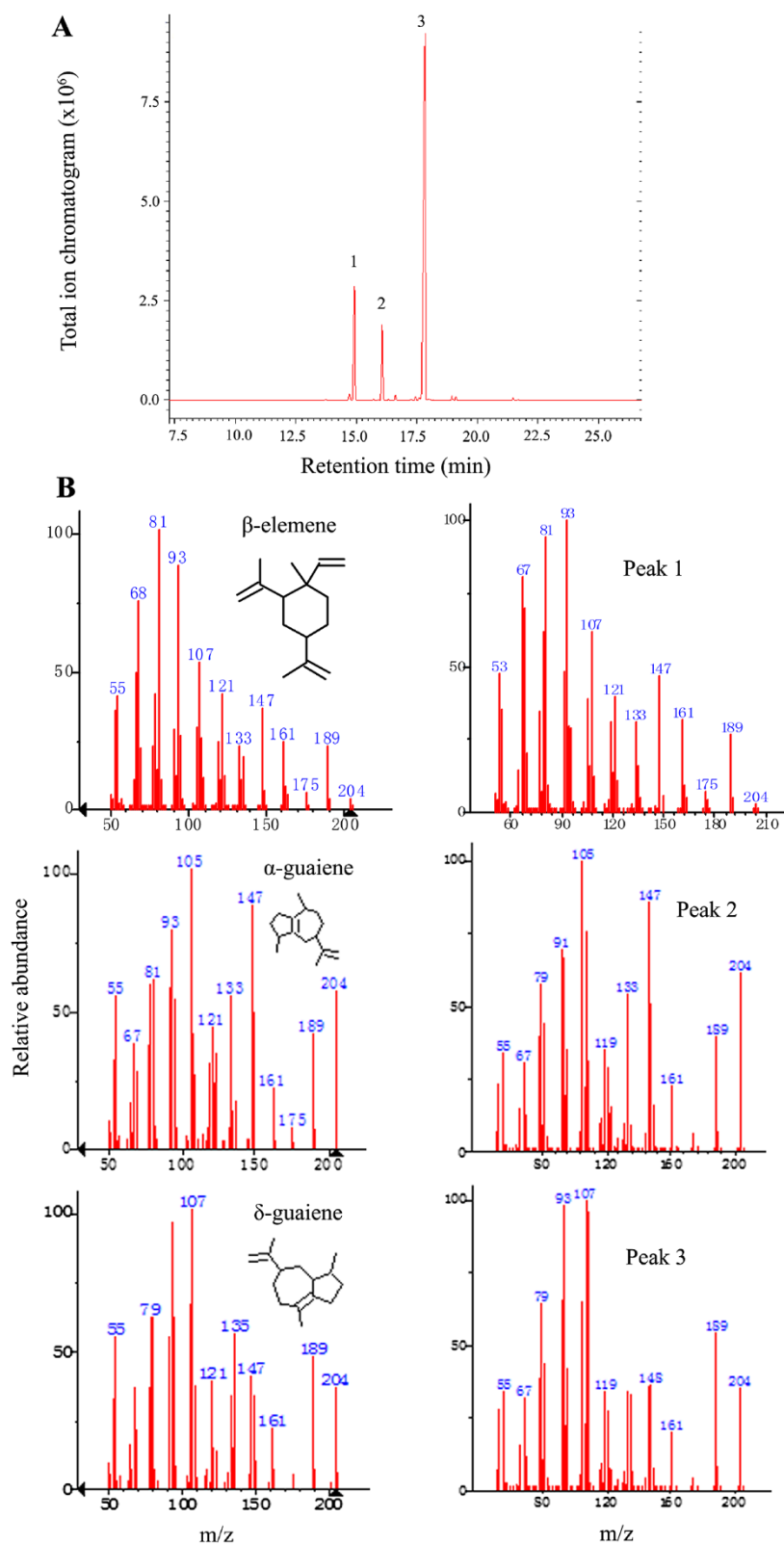


diphosphate (C10), FPP (C15), and geranylgeranyl diphosphate (C20) substrates, and the reaction products were analyzed using GC-MS. The result showed that ASS1-3 did not accept geranyl diphosphate or geranylgeranyl diphosphate as a substrate,

converting only FPP to terpene products. The 3 enzymes yielded the same compounds; the major product was identified as  $\delta$ -guaiene (74.2%), and the minor products as  $\beta$ -elemene (16.3%) and  $\alpha$ -guaiene (9.5%) (Figure 5).

**Table 2 Number of unigenes and 454 annotated reads involved in sesquiterpenes biosynthesis**

Enzyme Code (EC)	Abbreviated enzyme name	Enzyme name	Numbers of unigenes	Numbers of 454 reads	
				Healthy	Wounded
2.3.1.16	ACTC	acetyl-CoA acyltransferase	2	19	0
1.1.1.34	HMGR	HMG CoA reductase	3	151	21
2.2.1.7	DXPS	1-deoxy-D-xylulose-5- phosphate synthase	12	35	37
1.1.1.267	DXPR	1-deoxy-D-xylulose-5- phosphate reductoisomerase	1	3	2
2.7.1.36	MK	mevalonate kinase	1	48	1
2.7.4.2	PMK	phosphomevalonate kinase	3	9	12
4.6.1.12	MCS	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	1	11	11
2.5.1.10	FPS	farnesyl diphosphate synthase	2	78	5
	SS	sesquiterpene synthase	5	383	10



**Figure 5 Gas chromatography-mass spectroscopy profiles.** (A) Total ion chromatogram of the products formed by sesquiterpene synthases (ASSs) with farnesyl diphosphate as a substrate. (B) Mass spectra of the sesquiterpenes and their authentic standards. Peak 1 is  $\beta$ -elemene, peak 2 is  $\alpha$ -guaiene, and peak 3 is  $\delta$ -guaiene.

### Expression of ASSs induced by MJ and identification of sesquiterpenes produced in MJ-treated calluses

We analyzed the expression profile of ASS using RT-PCR. Two pools of cDNAs, one derived from the wounded stems of the 3-year-old *A. sinensis* that comprised the 454 library and another from *A. sinensis* calluses treated with MJ, were used for the analysis. MJ is an essential signaling molecule that modulates plant secondary metabolism [24-26] and is capable of inducing the production of sesquiterpenes in *Aquilaria* spp. [5,27,28]. Therefore, MJ-treated calluses were also used to investigate the expression of ASS. The results demonstrated that *ASS1* and *ASS2* were both induced by either mechanical wounding or MJ treatment (Figure 6A, 6B). The expression of *ASS1* was upregulated significantly—approximately 800 times in response to mechanical wounding (see Figure 6A) and 1000 times in response to MJ treatment (see Figure 6B).

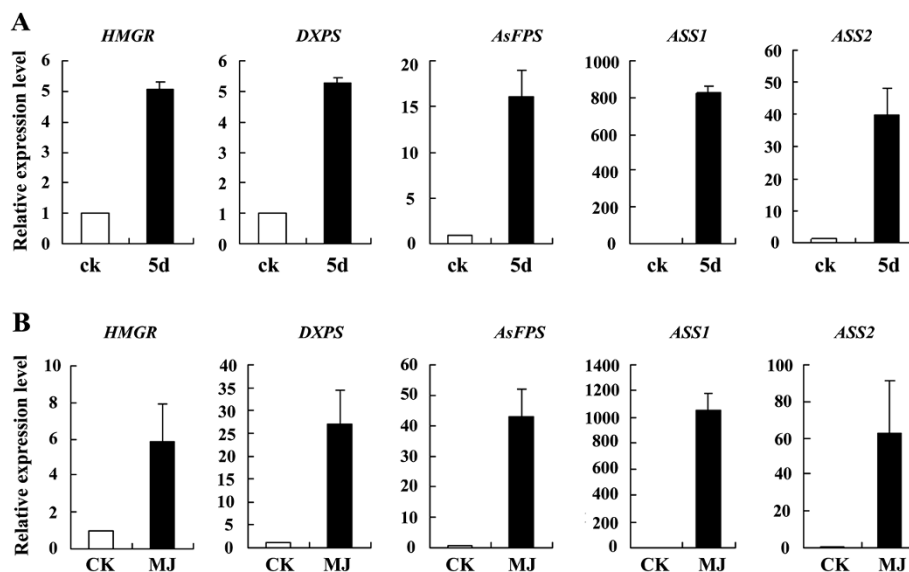
Furthermore, we tested the expression of other enzyme genes in sesquiterpene biosynthesis, including *HMGR*, *DXPS*, and *FPS*. They were also induced by mechanical wounding and MJ, but the upregulation rate was significantly weaker than that associated with *ASS1* and *ASS2*, fully demonstrating that *ASS1* is a typical stress-inducible gene. This result is consistent with those of solid-phase micro-extraction GC-MS analysis, showing that the production of sesquiterpenes increases in response to MJ treatment. In the untreated healthy callus

samples, only 2 sesquiterpenes ( $\delta$ -guaiene and  $\alpha$ -guaiene) were detected, with total peak area of 11.47. In the MJ-treated callus samples, 4 sesquiterpenes were detected, with total peak area of 850.6—approximately 74 times higher (Table 3). Among the 4 sesquiterpenes, the content of  $\delta$ -guaiene and  $\alpha$ -guaiene was more than 88%, an increase of more than 67-fold relative to the healthy control. The results demonstrated that inducible expression of ASS is responsible for the formation of agarwood sesquiterpenes in *A. sinensis*.

### Candidate regulators response to wound signals

TFs are critical regulators of gene expression and environmental stress responses [26,29]. In our 454 unigene database (H + W), a total of 4,786 unigenes represented homologs of various TF families (Additional file 10: Table S3): 86 were annotated as belonging to the APETLA2 (AP2)/ethylene-responsive-element-binding family [30,31], 66 to the bZIP family [32], 108 to the MYB family [33-35], 150 to the MYB-related family, 169 to the bHLH family [35], and 99 to the WRKY family [36,37]. All of these TFs are reportedly stress-related [36-39], and some homologous genes have been identified as positive or negative regulators in the biosynthesis of secondary metabolites in other plants [11,40-42].

In defensive response and wound signal transduction, calcium signaling and mitogen-activated protein kinase (MAPK) cascades play important roles [25,43,44]. In our



**Figure 6** Expression of 4 synthase genes in response to mechanical wounding and methyl jasmonate (MJ) treatment. (A) In the stem of 3-year-old *A. sinensis* trees, mechanical wounding increased the mRNA levels of 4 genes that encode enzymes in the sesquiterpene biosynthesis pathway. Each value is the mean  $\pm$  SE of 3 independent biological determinations. CK: healthy 3-year-old control trees; 5d: 3-year-old trees pruned after 5 days. (B) In *A. sinensis* calli, treatment with 100  $\mu$ M MJ increased the mRNA levels of 4 genes that encode enzymes in the sesquiterpene biosynthesis pathway. *A. sinensis* calli were transferred to MJ-containing medium and cultivated for 5 days before being sampled for analysis. Each value is the mean  $\pm$  SE of 3 independent biological determinations, and the data shown are relative ratios. The genes were named according to their functional annotation. CK: healthy calli controls; MJ: calli treated with MJ for 5 days.

**Table 3 Kinds and relative contents of sesquiterpenes from the untreated healthy calli and the MJ-treated calli**

T <sub>R</sub>	Sesquiterpene materials	RI	Peaks areas (x10 <sup>8</sup> )	
			Healthy	MJ-treated
12.808	β-humulene	1212.378	-	7.94
13.008	α-guaiene	1217.037	5.52	248.6
13.483	α-humulene	1228.101	-	86.96
14.461	δ-guaiene	1250.883	5.95	507.1

T<sub>R</sub>: retention time.  
 RI: retention indices.  
 '-': not containing the substance.

study, KEGG classification annotated 25 unigenes as being related to calcium signaling pathways and 41 to MAPK signaling pathways (Additional file 11: Table S4).

#### RT-PCR analysis of candidate regulators response to wound signals

To confirm which of the candidate regulators mentioned above are related to wound signals, we performed RT-PCR analysis to test the effects of MJ on the expression of these genes. Overall, 30 expressed regulators were selected (Additional file 12: Table S5), including TFs *MYB* and *WRKY*, protein kinases *CDPK* and *MAPK*, and some regulators related to signal molecules MJ, ethylene, and hydrogen peroxide. As shown in Figure 7, the expression of most of these genes was elevated in response to MJ treatment, and only 4 TFs, *HD-ZIP*, *WRKY23*, *WRKY60*, and *WRKY32*, were downregulated by MJ. Among the upregulated genes, 17 were increased more than 10-fold. They were TFs *WRKY4* and *MYB4*, protein kinases *CDPK1*, *CDPK2*, *CDPK5*, *CDPK6*, *MPKK2*, *MAPK2*, *MPK3*, *MAPK5*, and JA-, ethylene-, or hydrogen peroxide-related regulators *JAip*, *RBF*, *EIN3*, *ERF115*, *noxB*, *AOSC*, *ACO3*. The expression of 5 annotated unigenes, *MPKK2*, *MAPK2* [45], *MYB4* [46], *WRKY4*, and *noxB*, the superoxide-generating NADPH oxidase [47,48], was coordinate with *ASS1*, showing elevation of 126.9-, 63.9-, 80.9-, 23.7-, and 95.2-fold, respectively. The expression profiles of these wound-signaling-related genes are essentially consistent with the idea that the expression of *ASS1* is induced by wounding but regulated through a complex underlying mechanism.

## Discussion

### *Aquilaria sinensis* transcriptome sequencing

Although *Aquilaria* is the most important plant source of agarwood and all *Aquilaria* spp. are listed in Appendix II of the Convention on International Trade in Endangered Species of Wild Fauna and Flora [49], little progress to date has been made in deciphering the mechanism of agarwood formation and genomic and genetic information is lacking. The NCBI EST database contained only a few hundred *Aquilaria* sequences

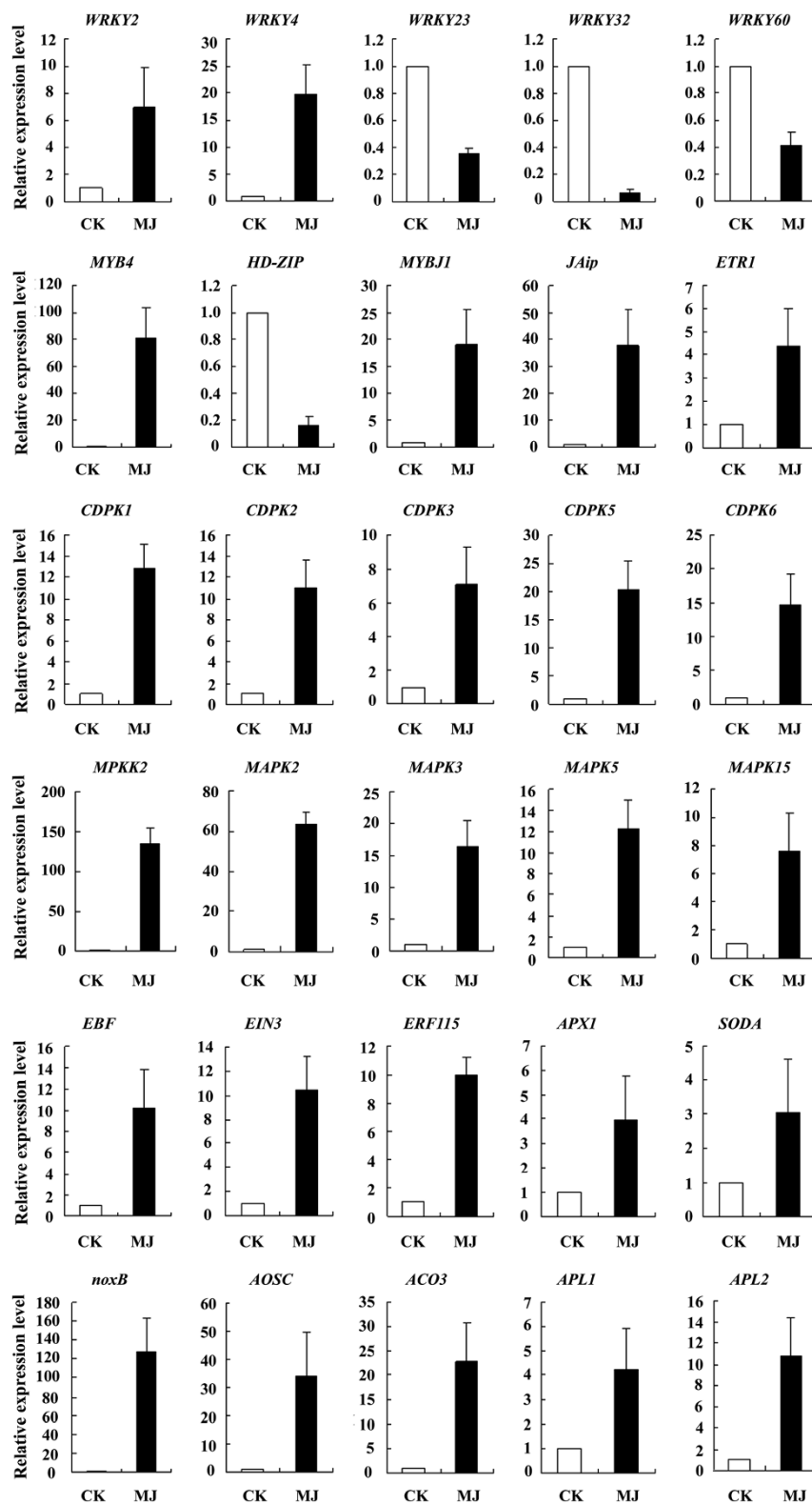
before this study. The 454 pyrosequencing reported herein represents the first concrete effort to generate cDNA resources for *Aquilaria* spp. We obtained 89,137 unigenes from the H and W *A. sinensis* cDNA libraries (Table 1) and annotated them in detail at bioinformatics levels (see Additional file 3: Table S2, Additional file 4: Figure S2 and Additional file 5: Figure S3). We estimated the genome size of *A. sinensis* preliminarily at approximately 2 G using flow cytometry. In the absence of a haploid, genome sequencing is difficult to carry out with current sequencing technology. In the absence of complete genome sequences, transcriptome mining provides the sole option for the subsequent annotation of sequences. The data reported herein may be useful for further research on the functional genomics of *Aquilaria* spp.. Importantly, by comparing the unigenes in the W and H libraries, we obtained a large number of DEGs (see Figure 3). These data suggested that the transcriptome of *A. sinensis* changed significantly in response to mechanical damage and also provided resources for screening genes related to wound signaling and metabolism. KEGG annotations of these DEGs showed that most of the unigenes were related to metabolism processes, especially carbohydrate metabolism, and had higher abundance in the W library than in the H library (see Figure 4). The number of carbohydrate metabolism-related unigenes was slightly greater in the W library, indicating that starch and sugar metabolism in *A. sinensis* may be stimulated by wounding (see Figure 2).

The synthesis of sesquiterpenes starts with acetyl-coenzyme A (the MVA pathway) [8] or glyceraldehyde-3-phosphate and pyruvate (the DXP pathway) [10] from glycolysis. Therefore, agarwood formation may have some association with starch and sugar metabolism and is essentially consistent with our electron microscopy findings that a high number of starch granules exist in the healthy wood of *A. sinensis*, but when the wood is wounded, resin in the agarwood accumulates while the starch granules degrade (our unpublished data).

### ASSs are inducible enzymes that may be responsible for the formation of agarwood sesquiterpenes

Considering that sesquiterpenes are the most important components of agarwood [2,6,7], we tried to identify the genes that encode the enzymes involved in sesquiterpene biosynthesis, and 30 such unigenes were found in our database, including 4 categories of critical genes (*HMGR*, *DXPS*, *FPS*, and *ASS*) and other genes encoding synthases in sesquiterpene metabolism in *Aquilaria* spp. (Table 2). These data will be valuable for further studies of the terpene biosynthesis pathway. Using a homology-based PCR approach, we obtained 3 full-length cDNAs of ASSs (*ASS1-3*) with very similar amino acid sequences. Phenetic analysis showed that the evolutionary





**Figure 7 Altered expression of a set of regulatory genes in response to MJ treatment.** Treatment with 100  $\mu$ M MJ altered the mRNA levels of a set of regulatory genes that may be involved in wound signal transduction. Each value is the mean  $\pm$  SE of 3 independent biological determinations, and the data shown are relative ratios. The genes were named according to their functional annotation. CK: healthy calli controls; MJ: calli treated with MJ for 5 days.

distances among the 3 ASSs were quite small compared to those of sesquiterpene synthases from other plant species, except *A. crassna* (see Additional file 9: Figure S7). Heterologous expression in *E. coli* and enzyme assays showed that ASS1-3 had ASS activity, and the major product was revealed to be  $\delta$ -guaiene (see Figure 5). Expressions analysis with real-time PCR showed that *ASS1* was barely detectable in either the healthy tissues or the calluses of *A. sinensis*, whereas in the wounded tissues or MJ-treated calluses, *ASS1* expression was increased almost 1000-fold (see Figure 6). Furthermore, GC-MS analysis showed that the major product was  $\delta$ -guaiene, essentially consistent with its enzyme reaction, and the sesquiterpene content was increased 74-fold (Table 3), clearly demonstrating that *ASS1* expression was wound-induced and is responsible for the formation of agarwood sesquiterpenes in *A. sinensis*.

Previous reports have shown that MJ can induced 3 sesquiterpenes ( $\alpha$ -guaiene,  $\alpha$ -humulene, and  $\delta$ -guaiene) in cultured cells of *A. crassna* [5,27,28] and be produced by the enzymatic reaction of cloned enzymes. Of the 5 clones (AcC1, AcC2, AcC3, AcC4, and AcL154), AcC1 and AcL154 have no TPS activity that might indicate pseudogenes [5]. The 3 clones (*ASS1-3*) in this study have high homology with AcC4 and AcC3 (see Additional file 9: Figure S7) and catalyzed the same major product. Thus, the present experiments allow us to identify *ASS1-3* as active enzymes of *A. sinensis*, and their inducible expression is responsible for the formation of agarwood sesquiterpenes.

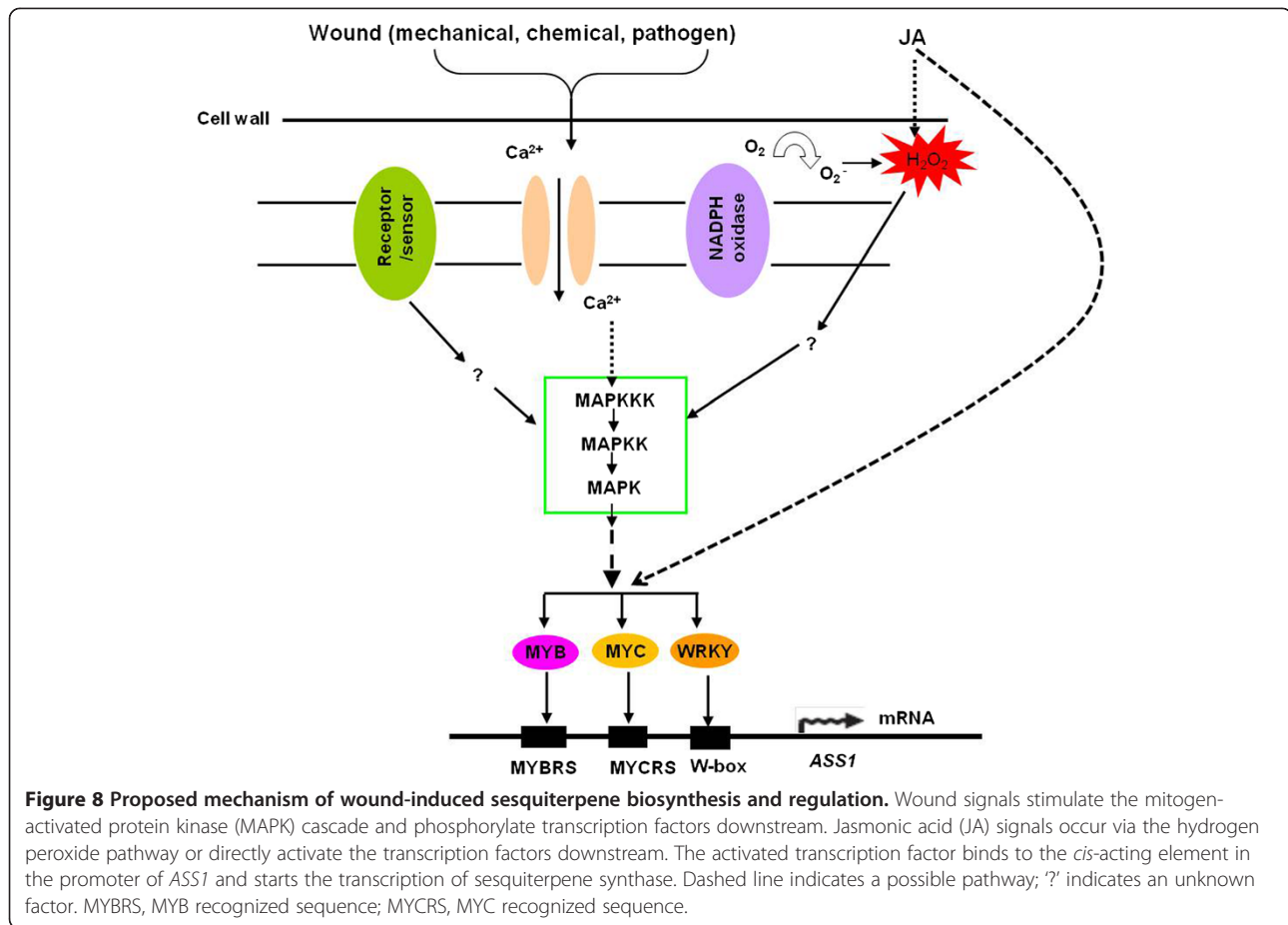
#### Wounding induces ASSs expression via a complex signaling pathway

Research has shown that only after wounding (including mechanical wounding, chemical induction, and fungal infection, among others) can agarwood be induced in the healthy wood of *Aquilaria* spp. [6,14-17]. It is, therefore, reasonable to suppose that the early period of agarwood formation is a process of wound signaling transduction before the strong expression of ASSs. Previous studies in other plant species have shown that TFs play important roles in signal transduction and gene expression. Some TFs, such as WRKY, MYB, MYC, bZIP, ERF, and EIN3 are involved in plant defense and stress responses [34,36,39]. Our 454 unigene database contained 4,786 putative TFs. Some homologous genes of these TFs, such as *AP2*, *WRKY*, and *MYC*, have been identified as positive regulators of sesquiterpene synthases in other plant species [11-13,42]. In *Catharanthus roseus*, the TF ORCA3 regulates the biosynthesis of terpenoid indole alkaloids and plays an important role in JA response [42]; in cotton, the TF GaWRKY1 positively regulates the expression of the sesquiterpene synthase (+)- $\delta$ -cadinene

synthase gene and the synthesis of gossypol sesquiterpene [11]. Similarly, in *A. annua*, AaWRKY1 and 2 JA-responsive AP2/ethylene response factor (ERF) TFs, ERF1 and ERF2, are involved in regulating the expression of the sesquiterpene synthase ADS and the production of artemisinin [12,49]; in *Arabidopsis*, MYC2 interacts with DELLA proteins and positively regulates the expression of TPS21 and TPS11 [13].

Reversible protein phosphorylation has been proven to be an important regulatory mechanism in the wound signaling pathway, and studies on *Arabidopsis*, tomato, and tobacco have indicated that the MAPK cascade is a widespread pathway [44]. In addition, calcium and calcium-binding proteins may play some roles in the regulation of early wound responses by elevating intracellular calcium levels and changing the phosphorylation patterns of proteins [50]. In this study, 41 unigenes were annotated as related to the MAPK pathway and 25 to calcium signaling, which may play roles in the wound response of agarwood formation.

In the present experiment, 30 DEGs related to wound signal were selected to investigate whether they are co-expression with *ASS1* in response to MJ. The results showed that 26 were upregulated and 4 were downregulated. Of these, 5 unigenes, including the TFs *MYB4* and *WRKY4*, the protein kinases *MAPK2* and *MPKK2*, and the NADPH oxidase *noxB*, were significantly upregulated by MJ. The downregulated genes were *WRKY23*, *WRKY32*, *WRKY60*, and *HD-ZIP* (see Figure 7). The *ASS1* promoter has been cloned and scanned with PLACE (<http://www.dna.affrc.go.jp/htdocs/PLACE/signalscan.html>). It contains some *cis*-acting elements that could be combined specifically by certain TFs, including WRKY, MYB, and MYC (data to be published elsewhere). Therefore, we speculate that TFs *MYB4* and *WRKY4* may positively regulate the expression of sesquiterpene synthase, and the downregulated genes may negatively regulate its expression. The generation of hydrogen peroxide has been confirmed to occur both locally and systematically, and its main source is the plasma-membrane-bound NADPH oxidase. In the present study, NADPH oxidase *noxB* was significantly upregulated by MJ, which is consistent with the theory that MJ induces the production of hydrogen peroxide [51-54]. MAPK signal transduction pathways are widespread mechanisms in eukaryotic cells that couple environmental responses and transcriptional regulation [55,56]. Two coupled components of the MAPK cascade, *MAPK2* and *MAPKK2*, were significantly upregulated by MJ, demonstrating they may be involved in MJ-mediated signaling. Calcium is a prevalent messenger molecule in higher plant cells that can mediate plant responses to external signals and regulate a variety of physiological processes. In our experiments, 4 CDPKs (*CDPK1*,



*CDPK2*, *CDPK5*, *CDPK6*) were upregulated more than 10-fold (see Figure 7).

Based on these results, we proposed a model to explain the mechanism by which *ASS1* is modulated at the transcriptional level (Figure 8). In this model, the MAPK cascade is activated by wound signals and phosphorylates downstream TFs such as MYB or WRKY. JA signaling may activate the TFs downstream via the hydrogen peroxide or JA-JAZ-TF pathways [57]. Activated TFs bind to the *cis*-acting element in the promoter of *ASS1* and initiate *ASS1* transcription. *ASS2* is also inducible and may be influenced by the same regulatory mechanisms as those affecting *ASS1*. The results reported herein enhance understanding of wound-induced sesquiterpene synthase expression and agarwood formation. Future studies will focus on verifying this proposed model based on the promoter of *ASS1* that we have cloned. We will identify which TFs interact with the *ASS1* promoter, confirm whether the MAPK cascade is involved in the wound signaling pathway, and determine the roles of them in *ASS1* expression. Clarifying the wound-induced sesquiterpene biosynthesis pathway will help to develop strategies for controlling agarwood production and for the long-term protection of *Aquilaria* spp..

## Conclusions

A comprehensive transcriptome analysis was conducted for healthy and wounded stems of *A. sinensis* in this study. A total of 89,137 unigenes with an average length of 678.65 bp were obtained, and they were annotated in detail at the bioinformatics levels. These data should be useful for further research on the functional genomics of *A. sinensis* and other *Aquilaria* spp.. Transcripts related to sesquiterpene metabolism and regulation were highlighted. Three full-length cDNAs of sesquiterpene synthase (*ASS1-3*) were cloned and identified. Through MJ-inducible expression pattern experiments and GC-MS analysis, we found that their expression was significantly induced by MJ, and sesquiterpenes content was elevated accordingly. The expression of a number of important regulatory genes was detected using RT-PCR and indicated that some of them may participate in one or more processes of wound signal transduction, transcriptional regulation of *ASS*, and biosynthesis of agarwood sesquiterpenes. Based on the analysis of *ASS1* promoter, we confirmed preliminarily that TFs *MYB4* and *WRKY4* may be key positive regulators of *ASS1* expression. These results provide a better understanding of sesquiterpene biosynthesis and regulation in *Aquilaria*

spp., and provide valuable clues and a basis for elucidating the molecular mechanism of agarwood formation.

## Methods

### Plant materials for cDNA library construction

Three-year-old *A. sinensis* trees were planted in Hainan Province of China. The stems were pruned with a knife, and samples were taken from 6 cm below the pruned site at various times after pruning: 2 h, 4 h, 6 h, 8 h, 12 h, 24 h, 3 d, and 5 d. All of the stems from the wounded plants were mixed in equal proportion to generate material for the wounded library, and the stems cut from healthy trees were used to generate material for the healthy library.

### Plant materials for real-time PCR (RT-PCR) and gas chromatography–mass spectrometry (GC-MS) analysis

Calluses originating from the stems of the healthy *A. sinensis* plants were used to investigate the effect of MJ on candidate genes expression and sesquiterpenes production. Calluses were transferred to 100  $\mu$ M MJ-containing medium and allowed to grow for 5 days before being sampled for analysis.

### Library construction and 454 sequencing

Total RNA was isolated from the wounded and healthy materials in the libraries using a TRIzol kit (Invitrogen, USA) and purified to enrich messenger RNA (mRNA) and exclude transfer RNA (tRNA) and ribosomal RNA (rRNA) using an mRNA purification kit (Promega, USA) according to manufacturer instructions. The mRNA was reverse-transcribed with PowerScript™ II (Takara, Japan) using SMART IV™ oligonucleotide PCR primers (5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3') and a CDS III/3' PCR primer (5'-ATTCTAGAGCCGAGCGGCCGACATG -d[T]30 N-1 N-3'). Long-distance PCR for double-stranded cDNA amplification was performed with LA Taq enzyme (Takara, Japan) for 25 cycles (95°C for 30 s, 68°C for 8 min) according to the SMART™ cDNA Library Construction Kit user manual. Finally, the double-stranded cDNA was purified using a DNA purification kit (Qiagen, Germany) to generate high-quality cDNA. Approximately 10  $\mu$ g cDNA was sheared and used for 454 sequencing. The cDNA sample was end-repaired and ligated to adapters. Streptavidin bead enrichment, DNA denaturation, and emulsion PCR were performed according to a procedure described by Margulies [58]. Each portion of the preparations from the healthy and wounded libraries were loaded on a full plate and sequenced in one run.

### 454 data processing and assembly

A Perl script was used to remove vector sequences and Poly(A)/(T) tails from the original sequences. Reads shorter than 100 bp were deemed low quality reads and

removed before assembly. The remaining high-quality reads were assembled using MIRA [59] to construct unique consensus unigenes, with the parameter --job = de novo, est, normal, 454 -SK: mnr = yes -SK: rt = 2454\_SETTINGS -LR: mxti = no, and then to remove the redundancy of unigenes, Ucluster [60] (version 2.0) was used to cluster the unigenes, with the parameter --tryRevC -R 50 -M 40 -E 2, finally the clusters of unigenes were assembled by Cap3 [61], with parameter -o 100 -p 95.

### Functional annotation and classification

#### NT and NR

The unigenes were searched against the NCBI's (<http://www.ncbi.nlm.nih.gov>) Non-redundant nucleotide database (NT) using the BLASTN program and against the Non-redundant protein database (NR) using the BLASTX program. For both searches, an *E* value  $\leq 1e-5$  was used.

#### SWISS-PROT

The unigenes were searched against the SWISS-PROT database (<http://www.expasy.ch/sprot>) using the BLASTX program with an *E* value  $\leq 1e-10$ .

#### KEGG and KEGG Orthology

Unigenes were compared with the Kyoto Encyclopedia of Genes and Genomes database (KEGG, release 50; <http://www.genome.jp/kegg/>) [62] using the BLASTX program with an *E* value  $\leq 1e-10$ . A Perl script was used to retrieve KEGG Orthology (KO) information from the BLAST results and then to establish pathway associations between the unigenes and KEGG.

#### Interpro and Gene Ontology

InterProScan [63] release 16.0 was used to annotate the InterPro domains of the unigenes, and then the functional assignments were mapped onto the Gene Ontology (GO) [64]. The WEGO tool [65] was used to assign GO classifications and draw GO trees.

#### Differentially expressed unigenes detection

The IDEG6 web tool [19] was used to identify unigenes that showed statistically significant differences in relative abundance (as reflected by the total count of individual sequence reads) between the two libraries. Comparisons between the wounded and healthy samples were performed using the general Chi2' method, which has been proven to be the most efficient test. The unigenes with  $P \leq 0.01$  were deemed to be significantly different between the two samples. This is similar to the credibility intervals approach that was employed earlier for the analysis of SAGE data [66].

### **Transcription Factors annotation**

Assembled transcripts were searched against all the transcription factor protein sequences of Plant transcription factor database [67] (PInTFDB, version 3.0) using BLASTX with an E-value cut-off of  $\leq 1e-10$ , and positives identity  $\geq 50\%$ .

### **Phenetic analysis**

The translated amino acid sequences and the putative amino acid sequences derived from the NCBI database search were aligned using the CLUSTAL W program. Evolutionary distances were calculated using the Poisson correction method, and a neighbor-joining (NJ) tree was constructed with MEGA4 [68]. Bootstrap values were obtained after 1,000 replications.

### **Quantitative real-time PCR (RT-PCR) analysis**

To investigate the expression of the mRNA of candidate genes, RT-PCR was performed using the Bio-Rad IQ-5 Real-Time System essentially according to the manufacturer's instructions. Total RNA was isolated from three-year-old *A. sinensis* stems and calluses using a Total RNA Purification Kit (LC Science, USA) with an on-column DNA digestion according to the manufacturer's instructions. The RNA samples were then reverse transcribed using the Superscript II RT kit (Invitrogen, USA). The amplification of the *AcHistone* gene [5] was used as an internal control. The efficiency of annealing of the oligonucleotide PCR primer sequences (listed in Additional file 13: Table S6) was evaluated using the Primer 5.0 program. The cDNA was amplified in a 10  $\mu$ L volume using SYBR Premix Ex Taq (TaKaRa, Japan) and a DNA Engine Opticon 2 thermal cycler with the following program: one cycle of 95°C for 10 s and 40 cycles of 94°C for 10 s and 60°C for 20 s. The amplification of the target genes was monitored every cycle by SYBR-green fluorescence. The Ct (threshold cycle), defined as the PCR cycle at which a statistically significant increase of reporter fluorescence was first detected, was used as a measure for the starting copy numbers of the target gene. Relative quantization of the target gene expression level was performed using the comparative Ct method. Three technical replicates were performed for each experiment. For all the RT-PCR analyses, the assays were repeated three times along with at least three independent repetitions of the biological experiments, and the means of the three biological experiments were calculated to estimate gene expression.

### **GC-MS analysis and identification of sesquiterpene components in *A. sinensis* calluses**

GC-MS analysis was performed with a Varian 450 GC (USA) equipped with a VF-5MS capillary column (30 m  $\times$  0.25 mm internal diameter; film thickness,

0.25  $\mu$ m), and a Varian 300 mass spectrometer with an ion-trap detector in full-scan mode under election impact ionization (70 eV). The carrier gas was helium, and the flow rate was 1 mL/min. The injections were performed in splitless mode at 250°C. The healthy and MJ-treated calluses were powdered and took the same weight (0.5 g) to place into the 15 mL sample bottle, and balanced for 30 min in 60°C water. The fused silica fiber of the Solid Phase Micro Extraction (SPME) was introduced into the headspace above the sample, adsorbing the volatile components for 30 min. The adsorbed components were followed by a thermal desorption process by introducing the SPME fiber into the injection port of a gas chromatography. The program was immediately started, and the fiber was removed after 10 min. The condition for GC-MS and identification of sesquiterpenes compounds were performed according to previously described procedures [6]. The two fibers were injected separately and ran in the same program, and the relative components were obtained by peak areas normalization without applying correction factors [6]. We didn't use the internal standard, for it is difficult to choose one suiting for both of the samples. Peaks areas can reflex the content of concerning compound in the two different samples in some extent, as the weight of the two samples and the analytical methods were identical.

### **Heterologous expression of *ASS* genes in *Escherichia coli* and enzyme assays**

The full-length sequence of *ASS1* was cloned into the XhoI (5' end)/BamHI (3' end) sites of a pET-28a vector (Novagen) and expressed in *E. coli* as a His-tag fusion protein using the following primers: forward primer, 5'-CCGCTCGAGATGTCTTCGGCAA~~AACTAGGT~~-3' and reverse primer, 5'-CGGGATCCGATTTCAATAGCATGACGCAAC-3' (XhoI and BamHI sites are indicated with underlining). To ensure that no errors were introduced by the PCR, the construct was checked with sequencing. Fusion proteins in the *E. coli* BL21 (DE3) cells were induced through treatment with 1 mM isopropylthio- $\beta$ -galactoside (IPTG) for 12 h at 30°C and 200 rpm. The protein purification was performed using a Profinia™ Protein Purification System (Bio-Rad, USA). Bacterial lysates and purified enzyme samples were examined on SDS-PAGE gels to visualize expected proteins. Protein concentrations were determined with the Bradford assay using bovine serum albumin as the standard. The enzyme reaction was performed according to the method described by Kumeta & Ito (2010), with minor modification. The reaction was performed in a 4-mL vial with a solid-top polypropylene cap using 50 mL of crude protein extract in a final volume of 200 mL containing Tris-HCl buffer (25 mM, pH 7.0) supplemented with 10% glycerol,

100 mM MgSO<sub>4</sub>, 5 mM DTT, and 46 mM FPP. After incubation at 30°C for 4 h, a SPME fiber was inserted into the headspace of the vial to collect volatiles for 30 min and then transferred to the injection port of a GC-MS system for desorption. The GC-MS system was described above and the operating parameters were essentially according to Kumeta & Ito [5]. The identification of sesquiterpenes was based on a comparison of retention times and mass spectra with authentic standards.

## Additional files

**Additional file 1: Figure S1.** Sesquiterpene biosynthesis pathway in plants.

**Additional file 2: Table S1.** Distribution of the number of reads in the *Aquilaria sinensis* unigenes.

**Additional file 3: Table S2.** Summary of annotation of the *A. sinensis* 454 assembled unigenes.

**Additional file 4: Figure S2.** Comparison of healthy and wounded tissue libraries based on GO terms.

**Additional file 5: Figure S3.** Functional classification of differentially expressed genes based on GO terms.

**Additional file 6: Figure S4.** Phenetic analysis of the proteins translated from the full-length AsFPS sequence from *Aquilaria sinensis* and some characterized farnesyl diphosphate synthase sequences from other plant species. Bootstrap values after 1000 replications are shown on the branches.

**Additional file 7: Figure S5.** Alignment of nucleotide sequences for 3 sesquiterpene synthase (ASS) genes.

**Additional file 8: Figure S6.** Alignment of deduced amino acid sequences for 3 ASSs.

**Additional file 9: Figure S7.** Phenetic analysis of 3 full-length sesquiterpene synthase sequences from *A. sinensis* (ASS1-3) and some characterized sesquiterpene synthase sequences from other plant species. Bootstrap values after 1000 replications are shown on the branches.

**Additional file 10: Table S3.** Major transcription factors identified from *A. sinensis*.

**Additional file 11: Table S4.** Unigenes annotated as being related to the calcium signaling and mitogen-activated protein kinase signaling pathways.

**Additional file 12: Table S5.** Candidate regulators used for RT-PCR analysis.

**Additional file 13: Table S6.** Primers used in this study.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

YHX participated in the design of the study, analyzed the data, performed the real-time PCR experiment and drafted the manuscript. ZZ contributed to the sample collection, participated in the design of the study and analysis of the data. MXW performed the cloning and identification of three sesquiterpene synthase genes. JHW initiated the project, helped conceive the study, and revised the manuscript. HJC participated in the GC-MS analysis. ZHG, HML, XLZ and WLL participated in the discussion of the result and helped to revise the manuscript. CS helped to analyze the result and with the phenetic analysis. YY and HM helped to the samples collection. All authors read and approved the final manuscript.

## Acknowledgements

The authors gratefully acknowledge Haibo Sun and Tao Feng for their assistance with bioinformatics. This study was supported by the National Natural Science Foundation of China (No. 31100220, No. 81173481, No. 31000136 and No. 81001607), the Program for New Century Excellent Talents

in University Funded by the Ministry of Education of China (No. 2008), National Key Technology R&D Program (No. 2011BAI01B07), Program of Modernization for Traditional Chinese Medicine of Hainan Provincial of China (No. 2012ZY002, 2010ZY001), and the Key Project in the Science & Technology Program of Hainan Provincial (No. ZDXM20120033, ZDXM20100006).

## Author details

<sup>1</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 151, Malianwa North Road, Haidian District, Beijing 100193, China. <sup>2</sup>Hainan Branch Institute of Medicinal Plant (Hainan Provincial Key Laboratory of Resources Conservation and Development of Southern Medicine), Chinese Academy of Medical Sciences & Peking Union Medical College, Wanning 571533, China. <sup>3</sup>Research Center on Life Sciences and Environmental Sciences, Harbin University of Commerce, Harbin 150076, China.

Received: 26 November 2012 Accepted: 22 March 2013

Published: 8 April 2013

## References

1. China pharmacopoeia Committee: *The Pharmacopoeia of People's Republic of China(I)*. Beijing: Chemical Industry Press; 2010.
2. Hashimoto K, Nakahara S, Inoue T, Sumida Y, Takahashi M: **A new chromone from agarwood and pyrolysis products of chromone derivatives.** *Chem Pharm Bull* 1985, **33**(11):5088–5091.
3. Ishihara M, Tsuneya T, Uneyama K: **Fragrant sesquiterpenes from agarwood.** *Phytochemistry* 1993, **33**:1147–1155.
4. Yagura T, Ito M, Kiuchi F, Honda G, Shimada Y: **Four new 2-(2-phenylethyl) chromone derivatives from withered wood of *Aquilaria sinensis*.** *Chem Pharm Bull(Tokyo)* 2003, **51**:560–564.
5. Kumeta Y, Ito M: **Characterization of  $\delta$ -guaiene synthases from cultured cells of *Aquilaria*, responsible for the formation of the sesquiterpenes in agarwood.** *Plant Physiol* 2010, **154**:1998–2007.
6. Chen HQ, Yang Y, Xue J, Wei JH, Zhang Z, Chen HJ: **Comparison of compositions and antimicrobial activities of essential oils from chemically stimulated agarwood, wild agarwood and healthy *Aquilaria sinensis* (Lour.) Gilg trees.** *Molecules* 2011, **16**:4884–4896.
7. Chen HQ, Wei JH, Yang JS, Zhang Z, Yang Y: **Chemical constituents of agarwood originating from the endemic genus *Aquilaria* plants.** *Chem Biodivers* 2012, **9**(2):236–250.
8. Gardner RG, Hampton RY: **A highly conserved signal controls degradation of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase in eukaryotes.** *J Biol Chem* 1999, **274**(44):31671–31678.
9. Shang CH, Zhu F, Li N, OU-Yang X, Shi L, Zhao MW, Li YX: **Cloning and characterization of a gene encoding HMG-CoA reductase from *Ganoderma lucidum* and its functional identification in yeast.** *Biosci Biotech Bioch* 2008, **72**(5):1333–1339.
10. Rohmer M: **The discovery of a mevalonate independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants.** *Nat Prod Rep* 1999, **16**:565–574.
11. Xu YH, Wang JW, Wang S, Wang JY, Chen XY: **Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)- $\delta$ -cadinene synthase-A.** *Plant Physiol* 2004, **135**:507–515.
12. Ma DM, Pu GB, Lei CY, Ma LQ, Wang HH, Guo YW, Chen JL, Du ZG, Wang H, Li GF, Ye HC, Liu BY: **Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the Amorpha-4,11-diene synthase gene, a key gene of artemisinin biosynthesis.** *Plant Cell Physiol* 2009, **50**(12):2146–2161.
13. Hong GJ, Xue XY, Mao YB, Wang LJ, Chen XY: **Arabidopsis MYC2 interacts with DELLA proteins in regulating sesquiterpene synthase gene expression.** *Plant Cell* 2012, **24**:2635–2648.
14. Ng LT, Chang YS, Kadir AA: **A review on agar (gaharu) producing *Aquilaria* species.** *J Trop Forest Products* 1997, **2**(2):272–285.
15. Itoh T, Tabata Y, Widjaja E, Mulyaningsih T, Parman, Wiriadinata H, Mandang YI: **Structure and artificial induction of aloe wood. The Fifth Pacific Regional Wood Anatomy Conference. Abstracts of Papers and Posters.** *IAWA Journal* 2002, **23**(4):466–467.
16. Pojanagaroon S, Kaewrak C: **Mechanical methods to stimulate aloes wood formation in *Aquilaria crassna* Pierre ex H.Lec. (Kritsana) trees.** *ISHS Acta Horticulturae* 2005, **676**:161–166.

17. Persoon GA: Growing 'the wood of the gods': agarwood production in Southeast Asia. *Smallholder Tree Growing for Rural Development and Environmental Service. Advance in Agroforestry* 2008, **5**:245–262.
18. Zhang Z, Yang Y, Wei JH, Meng H, Sui C, Chen HQ: Advances in studies on mechanism of agarwood formation in *Aquilaria sinensis* and its hypothesis of agarwood formation induced by defense response. *Chinese Traditional and Herbal Drugs* 2010, **41**(1):156–159.
19. Romualdi C, Bortoluzzi S, DAlessi F, Danieli GA: IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol Genomics* 2003, **12**(2):159–162.
20. Matsushita Y, Kang WK, Charlwood BV: Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from *Artemisia annua*. *Gene* 1996, **172**:207–209.
21. Delourme D, Lacroute F, Karst F: Cloning of an *Arabidopsis thaliana* cDNA coding for farnesyl diphosphate synthase by functional complementation in yeast. *Plant Mol Biol* 1994, **26**:1867–1873.
22. Cao XY, Yin T, Miao Q, Li HG, Ju XY, Sun Y, Jiang JH: Molecular characterization and expression analysis of a gene encoding for farnesyl diphosphate synthase from *Euphorbia pekinensis* Rupr. *Mol Biol Rep* 2012, **39**(2):1487–1492.
23. Kumeta Y, Ito M: Genomic organization of  $\delta$ -guaiene synthase genes in *Aquilaria crassna* and its possible use for the identification of *Aquilaria* species. *J Nat Med* 2011, **65**(3):508–513.
24. Raymond P, Farmer EE: Jasmonate and salicylate as global signals for defense gene expression. *Curr Opin Plant Biol* 1998, **1**:404–411.
25. Leon J, Rojo E, Sánchez-Serrano JJ: Wound signalling in plants. *J Exp Bot* 2001, **52**:1–9.
26. Zhao J, Davis LC, Verpoorte R: Elicitor signal transduction leading to production of plant secondary metabolites. *Biotechnol Adv* 2005, **23**:283–333.
27. Ito M, Okimoto K, Yagura T, Honda G: Induction of sesquiterpenoid production by methyl-jasmonate in *Aquilaria sinensis* cell suspension culture. *J Essent Oil Res* 2005, **17**(2):175–180.
28. Okudera Y, Ito M: Production of agarwood fragrant constituents in *Aquilaria calli* and cell suspension cultures. *Plant Biotechnol* 2009, **26**(3):307–315.
29. Hussain SS, Kayani MA, Amjad M: Transcription factors as tools to engineer enhanced drought stress tolerance in plants. *Biotechnol Progr* 2011, **27**(2):297–306.
30. Riechmann JL, Meyerowitz EM: The AP2/EREBP family of plant transcription factors. *Biol Chem* 1998, **379**(6):633–646.
31. Shigyo M, Hasebe M, Ito M: Molecular evolution of the AP2 subfamily. *Gene* 2006, **366**:256–265.
32. Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F: bZIP transcription factors in *Arabidops*. *Trends Plant Sci* 2002, **7**(3):106–111.
33. Chen YH, Yang XY, He K, Liu MH, Li JG, Gao ZF, Lin ZQ, Zhang YF, Wang XX, Qiu XM, Shen YP, Zhang L, Deng XH, Luo JC, Deng XW, Chen ZL, Gu HY, Qiu LJ: The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* 2006, **60**:107–124.
34. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L: MYB transcription factors in *Arabidopsis*. *Trends Plant Sci* 2010, **15**(10):573–581.
35. Feller A, Machemer K, Braun EL, Grotewold E: Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J* 2011, **66**(1):94–116.
36. Pandey SP, Somssich IE: The role of WRKY transcription factors in plant immunity. *Plant Physiol* 2009, **150**:1648–1655.
37. Rushton PJ, Somssich IE, Ringler P, She QJ: WRKY transcription factors. *Trends Plant Sci* 2010, **15**(5):247–258.
38. Schwechheimer C, Zourelidou M, Bevan MW: Plant transcription factor studies. *Annu Rev Plant Physiol Plant Mol Biol* 1998, **49**:127–150.
39. Singh KB, Foley RC, Oñate-Sánchez L: Transcription factors in plant defense and stress responses. *Curr Opin Plant Biol* 2002, **5**:430–436.
40. Bradley JM, Davies KM, Doles SC, Bloor SJ, Lewis DH: The maize regulatory gene up-regulates the flavonoid biosynthetic pathway of petunia. *Plant J* 1998, **13**:381–392.
41. Grotewold E, Chamberlain M, Snook M, Siame B, Butler L, Swenson J, Maddock S, Claire GS, Bowen B: Engineering secondary metabolism in maize cells by ectopic expression of transcription factors. *Plant Cell* 1998, **10**:721–740.
42. Van der Fits L, Memelink J: ORCA3, a jasmonate responsive transcriptional regulator of plant primary and secondary metabolism. *Science* 2000, **289**:295–297.
43. Jonak C, Ligterink W, Hirt H: MAP kinases in plant signal transduction. *Cell Mol Life Sci* 1999, **55**:204–213.
44. Pitzschke A, Schikora A, Hirt H: MAPK cascade signalling networks in plant defence. *Curr Opin Plant Biol* 2009, **12**:1–6.
45. Teige M, Scheik E, Eulgem T, Doczi R, Ichimura K, Shinozaki K, Dang JL, Hirt H: The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Mol Cell* 2004, **15**:141–152.
46. Pasquali G, Biricolti S, Locatelli F, Baldoni E, Mattana M: Osmby4 expression improves adaptive responses to drought and cold stress in transgenic apples. *Plant Cell Rep* 2008, **27**:677–1686.
47. Babior BM: NADPH oxidase. *Curr Opin Immunol* 2004, **16**:42–47.
48. Mittler R, Vanderauwera S, Gollery M, Breusegem FV: Reactive oxygen gene network of plants. *Trends Plant Sci* 2004, **9**(10):490–498.
49. CITES: *Amendments to Appendix I and II of CITES*. Bangkok, Thailand: Proceedings of Thirteenth Meeting of the Conference of the Parties; 2004:2–14.
50. Rojo E, Titarenko E, Leon J, Berger S, Vancanney G, Sánchez-Serrano JJ: Reversible protein phosphorylation regulates JA-dependent and -independent wound signal transduction pathway in *Arabidopsis thaliana*. *Plant J* 1998, **13**:153–165.
51. Orozco-Cárdenas ML, Ryan CA: Hydrogen peroxide is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway. *Proc Natl Acad Sci USA* 1999, **96**:6553–6557.
52. Orozco-Cárdenas ML, Narváez-Vásquez J, Ryan CA: Hydrogen peroxide acts as a second messenger for the induction of defense genes in tomato plants in response to wounding, systemin, and methyl jasmonate. *Plant Cell* 2001, **13**:179–191.
53. Hung KT, Kao CH: Nitric oxide acts as an antioxidant and delays methyl jasmonate -induced senescence of rice leaves. *J Plant Physiol* 2004a, **161**:43–52.
54. Hung KT, Hsu YT, Kao CH: Hydrogen peroxide is involved in methyl jasmonate-induced senescence of rice leaves. *Physiol Plantarum* 2006, **127**:293–303.
55. Nakagami H, Pitzschke A, Hirt H: Emerging MAP kinase pathways in plant stress signaling. *Trends in Plant Science* 2005, **10**(7):339–346.
56. Colcombet J, Hirt H: Arabidopsis MAPKs: a complex signalling network involved in multiple biological processes. *Biochem J* 2008, **413**:217–226.
57. Kazan K, Manners JM: Jasmonate Signaling: Toward an Integrated View. *Plant Physiol* 2008, **146**:1459–1468.
58. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, et al: Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005, **437**:376–80.
59. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, Suhai S: Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 2004, **14**:1147–1159.
60. Trivedi N, Bischof J, Davis S, Pedretti K, Scheetz TE, Braun TA, Roberts CA, Robinson NL, Sheffield VC, Soares MB, Casavant TL: Parallel creation of non-redundant gene indices from partial mRNA transcript. *Fut Generation Comput Syst* 2002, **18**:863–870.
61. Huang X, Madan A: CAP3: a DNA sequence assembly program. *Genome Res* 1999, **9**:868–77.
62. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M: From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 2006, **34**:D354–D357.
63. Zdobnov EM, Apweiler R: InterProScan an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 2001, **17**(9):847–848.
64. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, Richter J, Rubin GM, Blake JA, Bult C, Dolan M, Drabkin H, Eppig JT, Hill DP, Ni L, Ringwald M, Balakrishnan R, Cherry JM, Christie KR, Costanzo MC, Dwight SS, Engel S, Fisk DG, Hirschman JE, Hong EL, Nash RS, et al: The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 2004, **32**:D258–D261.
65. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang L: WEGO: a web tool for plotting GO annotations. *Acids Research* 2006, **34**:293–297.

66. Vencio RZ, Brentani H, Pereira CA: **Using credibility intervals instead of hypothesis tests in SAGE analysis.** *Bioinformatics* 2003, **19**:2461–2464.
67. Perez-Rodriguez P, Riano-Pachon DM, Correa LGG, Rensing SA, Kersten B, Mueller-Roeber B: **PlnTFDB: updated content and new features of the plant transcription factor database.** *Nucleic Acids Res* 2010, **38**:D822–D827.
68. Sui C, Zhang J, Wei JH, Chen SH, Li Y, Xu JS, Jin Y, Xie CX, Gao ZH, Chen HJ, Yang CM, Zhang Z, Xu YH: **Transcriptome analysis of *Bupleurum chinese* focusing on genes involved in the biosynthesis of saikosaponins.** *BMC Genomics* 2011, **12**:539.

doi:10.1186/1471-2164-14-227

**Cite this article as:** Xu et al.: Identification of genes related to agarwood formation: transcriptome analysis of healthy and wounded tissues of *Aquilaria sinensis*. *BMC Genomics* 2013 **14**:227.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

