Characterization of terpene synthase gene family in Artemisia annua L.

Weizhang Jia¹,², Ying Wang³*, Xuesong Yu¹,², Ling Zhang¹,², Shunhui Liu¹,², Zebo Huang¹,²*

¹School of Bioscience and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou 510006, China
²Guangdong Provincial Key Laboratory of Biotechnology Candidate Drug Research, Guangzhou 510006, China
³College of Horticulture, South China Agricultural University, Guangzhou 510640, China

*Corresponding author: wangying615@gmail.com (YW); zhuang@whu.edu.cn (ZBH)

Abstract

Terpene synthases (TPSs) are responsible for the synthesis of the various terpene molecules which play important roles in the accumulation of secondary metabolites. However, systematic analysis of Artemisia annua TPS family has not been carried out, and little is known about the differences of transcript levels for TPS genes in A. annua (AaTPSs) and the regulation thereof. Though plant TPSs are generally divided into seven clades, a family of TPSs in A. annua was characterized with lineages having a majority of TPS in TPS-a and TPS-b clades. Semi-quantitative RT-PCR demonstrates that 17 AaTPS genes are functional or potentially functional and express in at least some tissues or organs of the plant. In trichomes, the transcript of majority of AaTPSs and artemisinin synthesis-related genes are express at a higher level than in leaves containing trichomes, except that AaBFS, AaCPS and AaSQS appear to be lower in trichomes but higher in leaves. AaECS is mainly express in trichomes, but although the highest expression level was detected in old leaf trichomes. The expression of AaSOS in leaves was slightly higher than that of trichomes, indicating that this gene is also express mainly in other parts of leaves instead of trichomes. Characteristics of sequence motifs, phylogenetic relationships and transcript levels of AaTPSs provide a foundation for future functional analysis. The present study contributes more data on the tissue-specific expression of AaTPSs in A. annua, and provides clues to dissect the concerted regulation for AaTPSs expression and the specific competitor for farnesyl diphasphate (FDP) in A. annua trichomes.

Keywords: Artemisia annua L, Terpene synthase, Transcript level, Artemisinin.

Abbreviations: ACT, Artemisinin-based combination therapy; ADS, amorpha-4,11-diene synthase; ALDH1, aldehyde dehydrogenase; BPS, β-pinene synthase; CPR, cytochrome P450 reductase; CPS, β-caryophyllene synthase; CYP71AV1, cytochrome P450 monooxygenase; DBR2, double bond reductase; DHAA, dihydroartemisinic acid; DMAPP, dimethylallyl diphasphate; ECS, 8-epicdrol synthase; EST, expressed sequence tags; FDP, farnesyl diphasphate; GAS, germacrene A synthase; LS, (R)-linalool synthase; MVA, mevalonic acid; SQS, squalene synthase; TSA, transcriptome shotgun assembly.

Introduction

Sweet wormwood (Artemisia annua L.) is an important Chinese traditional medicine resource, which has received increasing attention due to the fact that the plant produces sesquiterpene lactone endoperoxide, artemisinin, that has become the most important agent in the treatment of malaria, particularly in the form of artemisinin-based combination therapy (ACT) (Haynes, 2006). The biosynthesis pathway of artemisinin is almost completely elucidated in A. annua. It is a sesquiterpenoid synthesized in the glandular trichomes of A. annua which are 10-cell structures with apical cells, sub-apical cells and three pairs of secretory cells (Olsson et al., 2009). The first committed step in artemisinin biosynthesis is the cyclization of farnesyl diphasphate (FDP) to generate amorpha-4,11-diene, catalysed by amorpha-4,11-diene synthase (ADS) (Mercke et al., 2000). The next reaction is sequential oxidation of amorpha-4,11-diene to artemisinic acid via formation of artemisinic alcohol and artemisinic aldehyde by a cytochrome P450 monooxygenase (CYP71AV1), which is the final enzymatic intermediate and precursor of artemisinin B (Teoh et al., 2006). A double bond reductase (DBR2) and an aldehyde dehydrogenase (ALDH1) were also characterized, and indicated that they involved in artemisinin biosynthesis pathway (Zhang et al., 2008). They appears to function in the conversion of artemisinic aldehyde to its dihydro form and then to dihydroartemisinic acid (DHA), respectively. All of these genes of biosynthesis of artemisinin have been shown to be highly expressed in glandular trichomes, and metabolic engineering of the artemisinin-specific pathway could enhance artemisinin biosynthesis (Xiang et al., 2012). Artemisinin has traditionally been isolated from the leaf tissue of field-grown plants, but yield of the drug is characteristically low and vary widely among different cultivars or ecotypes of A. annua. Despite production through synthetic chemistry approaches appear promising (Paddon et al., 2013), breeding and genetic engineering of A. annua plants for increased artemisinin production still remain of great interests (Graham et al., 2010; Supaibulwattana et al., 2011; Zhang et al., 2013). So far, however, little work has been carried out on systematic analysis of A. annua terpene synthase gene family (AaTPSs), which involved in terpene biosynthesis and closely related to artemisinin biosynthesis.

Terpenoids are compounds that are derived from the 5-carbon isopentenyl diphasphate (IPP) and its highly electrophilic isomer dimethylallyl diphasphate (DMAPP). Two independent pathways in plants, the methylerthritol phosphate (MEP) and mevalonic acid (MVA) pathways, lead
to the synthesis of both IPP and DMAPP. The terpenes produced directly by TPS enzymes, such as monoterpenes, sesquiterpenes and diterpenes, and those TPS products obtained after modification reactions may serve various ecological and physiological roles (Heiling et al., 2010). A recent classification of plant TPS divides them into seven clades TPS-a, TPS-b, TPS-c, TPS-d, TPS-e/f, TPS-g, and TPS-h (Chen et al., 2011). The major components of the essential oil from *A. annua* are monoterpenes and sesquiterpenes produced in trichomes, which belong to TPS-a and TPS-b clades. So far three monoterpane syntheses (3R)-Linalool synthase (LS, QH1 and QH5) (Jia et al., 1999) and β-pinene synthase (BPS) (Lu et al., 2002) genes were cloned from the plant. Sesquiterpene synthases epi-cedrol synthase (ECS) (Mercke et al., 1999), amorpha-4,11-diene synthase (ADS) (Mercke et al., 2000), β-caryophyllene synthase (CPS) (Cai et al., 2002), β-farnesene synthase (BFS) (Picaud et al., 2005), germacrene A synthase (GAS) (Bertea et al., 2006) and other novel sesquiterpene synthases (Liu et al., 2002; Wang et al., 2009) genes have been characterized from *A. annua*. Moreover, other classes of monoterpenes and sesquiterpenes were also isolated from the plant (Brown, 2010). There has evidence that IPP for artemisinin production originates from both the cytosol and plastid arms of the terpenoid biosynthetic pathway (Schramek et al., 2010). Therefore, all these terpene synthases, especially sesquiterpene synthases, may compete for the same substrate, which is intermediate and precursor of artemisinin. Consequently, the activity of various terpene synthases may influence the yield of artemisinin in the plant. Considering the importance of TPS genes, there is a need to have a systematic analysis of *A. annua* TPS gene family to enrich the existing information helping to identify candidates for gene bioprospecting and downstream expression analysis, and provide a foundation for future functional analysis of *AaTPSs*. Moreover, squalene synthase (SQS) that condensates two molecules of FDP to squalene involves in the biosynthesis of sterols and triterpenes. A significant amount of SQS5 was expressed in the tissues with high artemisinin biosynthesis, which may affect the yield of artemisinin (QH1 and QH5), AaMTS1, AaMTS2, AaMTS3, AaADS, AaCPS, AaECS, AaBFS, AaGAS, AaAG24640, AaSTSI and AaSTS2 (Table S1 and Table S2).

Because of the lack of a completely sequenced and annotated *A. annua* genome, it is not currently possible to estimate the number of TPS genes in such genomes, but analysis from EST and TSA databases for *A. annua* suggests that 17 TPS genes are the expressed functional TPS genes. Phylogenetic analysis indicates TPS-a and TPS-b clades have substantially diverged from the other TPS clades (Fig. 3). Many of the genes in these two clades have been functionally characterized in model and non-model systems (Table S2). Cluster analysis of *A. annua* TPS protein sequences indicates that 11 belong to the TPS-a clade, representing potentially functional TPS genes (AaADS, AaCPS, AaECS, AaBFS, AaGAS, AaAG24640, Aa16267, Aa14765, Aa02039, AaSTSI and AaSTS2). The sesquiterpene synthase TPS-a clade appears to be highly divergent in all seed plants characterized to date, it can be further divided into two subgroups, TPS-a-1 and TPS-a-2. All characterized TPSs in TPS-b subfamily are either monoterpane synthases or isoprene synthases, *AaTPSs* genes are in clade TPS-b, including 6 potentially functional ones (AaBPS, AaLS (QH1 and QH5), AaMTS1, AaMTS2, AaMTS3). Among these 17 *AaTPS* proteins sequences, the RR(x)8W motif is absolutely conserved in all *AaTPS* of the TPS-b group that have putative N-terminal plastid transit peptide. Variations of the RR(x)8W motif are found in those *AaTPS* of the TPS-a group of sesquiterpene synthases. Conserved RxR and DDxxD motifs are separated by a short region of 34 amino acid residues. The RxR motif is highly conserved in all *AaTPS* of the *AaTPS* group. However, the modification reaction of the prenyl diphosphate substrate and stabilization of a reactive diphosphate anion and carbocation upon ionization. The C-terminal motif of NSE/DYTE is not completely conserved almost all *AaTPSs* of the TPS-a and TPS-b subfamilies.

**Expression analysis of *AaTPSs* genes in tissues**

RT-PCR analysis of different tissues of *A. annua* with specific primers for 17 potentially functional genes identified transcript levels for all of these genes (Fig. 4). Evidences for mRNA accumulation were obtained for all genes encoding *AaTPS* proteins. These monoterpane synthase *AaAPS*, *AaLS*, *AaMTS1* and *AaMTS2* genes were detected in all tissues but their levels were variable. However, the monoterpane synthases *AaMTS3* genes was detected only in flower bud. The transcripts of *AaADS*, *AaECS*, *AaBFS* and *AaSTSI* genes were found in all root, stem, young and old leaves and flower.
Fig 1. Tissues and isolation of trichomes from A. annua. A Root; B Stem; C Old leaf; D Young leaf; E Leaf primordial; F Flower bud; G and H Isolated trichomes.

Fig 2. Alignment of the conserved RR(x)8W, RxR, DDxDD and NSE/DTE motifs of AaTPS members. The RR(x)8W absolutely conserved in all AaTPSs of monoterpene synthases that have putative N-terminal plastid transit peptides, variations of the RR(x)8W motif are found in those AaTPSs of the group of sesquiterpene synthases. Two conserved motifs, RxR and DDxDD, are separated by a short region of 34 amino acids in AaTPSs. The conserved NSE/DTE motif was also found in AaTPSs. AaBPS (-)-beta-Pinene Synthase (accession no. Q94G53); AaLS (QH1 and QH5) (3R)-Linalool synthase (accession nos. Q9SPN0 and Q9SPN1); AaMTS1, AaMTS2 and AaMTS3 three novel monoterpene synthases; AaADS amorpha-4,11-diene synthase (accession no. CAB94691); AaCPS beta-caryophyllene synthase (accession no. AAL79181); AaECS 8-epicedrol synthase (accession no. AAF80333); AaBFS (E)-beta-farnesene synthase (accession no. AAX39387); AaGAS germacrene A synthase (accession no. ABE03980); Aa16267, Aa14765 and Aa02039 represent that the sequence were identified by Wang et al. (2009) and there is no corresponding protein ID on the GenBank record; AaAAG sesquiterpene synthase (accession no. AAG24640); AaSTS1 and AaSTS2 two novel sesquiterpene synthases. Amino acid residues are numbered relative to AaBPS, the highly conserved residues are shaded in dark grey, and other highly similar residues are shaded in light grey.

Fig 3. Phylogenetic tree of the AaTPSs family and representative terpene synthases of known protein. A Maximum likelihood tree of all 251 proteins was constructed based on analysis of amino acid sequences by the MEGA 5.05 program. The sequences of TPS used for analysis are derived from the GenBank and SwissPort (Table S2). Based on the phylogeny and function of known TPSs, subfamilies of TPSs include TPS-a, TPS-b, TPS-c, TPS-d, TPS-e/f and TPS-g.
bud. The transcripts of AaCPS, AaGAS and AaSTS2 genes were only found in young and old leaves and flower bud. The sesquiterpene synthase AaAAG24640 and Aa14765 genes are more highly expressed in A. annua flower bud than in root, but stem and old leaf tissues were undetected. The sesquiterpene synthase Aa02039 and Aa16267 genes were detected few in flower bud and no transcripts in other tissues. The expression level of AaSQS was slightly higher than that of almost all monoterpene and sesquiterpene synthases, except AaADS genes.

**Comparative transcript profiling of AaTPSs in tissues and trichomes**

In order to comparative transcript profiling of AaTPSs in tissues and their trichomes, younger and old leaves were selected and used to isolated trichomes in this study. According to the result of expression analysis of AaTPS genes in tissues, 11 AaTPS and AaSQS genes were studied in young and old leaves of A. annua (Fig. 5). Transcripts of all the genes studied could be detected at different transcript levels using qPCR. The expression levels in different tissues are all compared separately to old leaves, which were arbitrary chosen as reference tissue. Majority of AaTPS genes are more highly expressed in A. annua young leaf than in old leaf, but AaECS transcript was detected higher in old leaf. The expression of monoterpene synthase AaBPS, AaLS, AaMTS1 and AaMTS2 genes were monitored in our qPCR experiments. Among that AaBPS, AaMTS1 and AaMTS2 have higher expression in trichomes than in leaf tissue, the expression of AaLS is slightly higher expression in trichomes, but no significant difference compared with the leaves. The regulation of artemisinin biosynthesis by AaADS was reflected in the result that around 290- , 770- and 3-fold more expression was seen in highly biosynthetically active young leaves, young leaves trichomes and old leaf trichomes, respectively, than in old leaves. It is interesting to note that AaADS was enhanced about 3-fold in young and old leaf trichomes than in young and old leaves. This should be due to the gene specific expression in the trichomes. The AaECS, AaGAS and AaSTS2 are mainly expressed in the trichomes, which has a similar expression pattern with AaADS gene. However, three sesquiterpene synthase AaBPS, AaCPS and AaSTS1 genes are mainly expressed in leaf tissue, rather than the trichomes. The expression of AaSQS in leaves was slightly higher than that of trichomes, indicating that the gene should be also express mainly in other parts of leaf instead of trichomes.

**Expression analysis of artemisinin biosynthetic genes**

The RT-PCR and qRT-PCR data presented here clearly demonstrate that the expression levels of AaADS, AaCYP71AV1, AaCPR, AaDBR2 and AaALDH1 genes were very high in young leaves and young leaf trichomes compared to that in old leaves, approximately 290- and 770-fold more AaADS, 300- and 720-fold more AaCYP71AV1, 90- and 360-fold more AaDBR2, 2- and 3-fold more AaCPR and 140- and 450-fold more AaALDH1 in young leaves and young leaf trichomes, respectively (Fig. S1 and Fig. 6). These high expression levels indicate a high capacity to produce artemisinin precursors in these tissues carrying biosynthetically active trichomes.

**Discussion**

Trichomes of A. annua produce valuable compounds for pharmaceutical uses, most notably the anti-malarial artemisinin. When leaf size increases, there is a significantly decreased of trichomes density due to the increased expansion of leaf compared with the apical leaf primordia. The original numbers of trichomes seem to be spread out on a larger surface as the leaf expansion and maturation. Consistent with earlier study (Kjær et al., 2012), the density of trichomes also shows significantly higher on the abaxial side than the adaxial side in 6 months old plant. Moreover, the intensity of the fluorescence signal seems to be a process of gradual reduction from the young leaves to old leaves. These suggest that trichomes are actually gradual collapsed rather than just reduced in relative number per unit area as cells expand during leaf maturation.

Terpenes constitute a large class of compounds that serve multiple roles in the interactions between plants and their environment, such as attracting pollinators and defending the plant against pests (Falara et al., 2011). The plant TPS that yield the low molecular weight terpenoids (isoprene, monoterpenoid, sesquiterpene and diterpenoid) have a common phylogenetic origin and are members of a diverse TPS family that is divided into seven subfamilies. Two TPS clades were previously demonstrated to encode sesquiterpene synthases that catalyze the formation of mostly epi-cedrol, amorpha-4,11-diene, β-caryophyllene, β-farnesene and germacrene A, respectively (Mercke et al., 1999; 2000; Cai et al., 2002; Picaud et al., 2006; Wang et al., 2013), the tissue-specific expression of TPSs in one or two clades, with some plant lineages having a majority of their TPSs in one or two clades, indicating lineage specific expansion of specific type of gene. A family of TPS in A. annua was characterized with lineages having a majority of TPSs in TPS-a and TPS-b clades. The TPS-a clade has the largest number of potentially functional TPS genes (11 TPSs characterized in A. annua). Five of the TPS-a clade genes, AaECS, AaADS, AaBPS, AaBFS and AaGAS were previously demonstrated to encode sesquiterpene synthases that catalyze the formation of mostly epi-cedrol, amorpha-4,11-diene, β-caryophyllene, β-farnesene and germacrene A, respectively (Mercke et al., 1999; 2000; Cai et al., 2002; Picaud et al., 2005; Bertea et al., 2006). Two new sesquiterpene synthases (AaSTS1 and AaSTS2) and other members (AaAAG24640, Aa16267, Aa14765 and Aa02039) have been delineated into the TPS-a subfamily (Wang et al., 2009). Most enzymes previously characterized from angiosperm-specific subfamily TPS-a are monoterpene synthases, this phenomenon is consistent with functional A. annua TPS-b clade genes. Three of the TPS-b clade genes, AaLS (Q81 and Q85) and AaBPS, were previously demonstrated to encode monoterpene synthases that catalyze the formation of mostly (3R)-Linalool and β-pinene from GPP, respectively (Jia et al., 1999; Lu et al., 2002). Three new monoterpene synthases (AaMTS1, AaMTS2 and AaMTS3) have a absolutely conserved amino acid sequence with a characterized RR(x)W motif, and that contain a putative N-terminal plastid transit peptide. Together with 8 terpene synthases whose activities were previously reported (Jia et al., 1999; Lu et al., 2002; Mercke et al., 1999; 2000; Cai et al., 2002; Picaud et al., 2005; Bertea et al., 2006; Wang et al., 2013), the tissue-specific expression patterns of 17 AaTPSs in A. annua have now been characterized. Our RT-PCR results with gene-specific primers for all potentially functional AaTPS genes indicated the transcripts of 17 AaTPS genes were expressed in at least one or another of A. annua tissues or organs, including root, stem, young and old leaves and flower bud. Four AaTPS genes, AaADS, AaECS, AaBFS and AaSTS1, exhibited a broad expression profiles with transcripts detected in all tissues. On the other hand, there were several genes (Aa16267, Aa14765, Aa02039 and AaSTS3) that were expressed specifically in only one or just a few tissues. Overall, the tissues from which the largest number of AaTPS genes are expressed in root, leaf
Fig 4. Semi-quantitative RT-PCR analysis of AaTPSs from A. annua. Expression levels were measured through RT-PCR using the primers listed in Table S3. The accession numbers of AaTPSs have been shown in Fig. 2. AaSQS squalene synthase (accession no. AY445506); AaActin β-actin (accession no. EU531837). R Root; S Stem; OL Old leaf; YL Young leaf; FB Flower bud.

Fig 5. Comparative transcript profiling of AaTPSs and AaSQS in leaves and their trichomes by real-time quantitative in A. annua. Expression levels were measured through qPCR using the primers listed in Table S3. The accession numbers of AaTPSs, AaSQS and AaActin have been shown in Fig. 2 and Fig. 4. The expression ratios were calculated relative to the expression in old leaves. All graphs represent the values of two independent qPCR runs with cDNA prepared from different plants. Sample triplicates were used in all qPCR runs. Expression ratios are illustrated by box and whisker plots. Boxes above expression ratio 1 represent higher gene expression levels compared to old leaves. A Young leaf; B Young leaf trichomes; C Old leaf trichomes.
Fig 6. Comparative transcript profiling of artemisinin biosynthetic genes in leaves and their trichomes by real-time quantitative in A. annua. AaADS, AaEPSs, AaCPS, AaBPS, AaALDH1, AaTPSs, AaECS, AaSQS, AaCPR, AaDBR2, AaDBR1, AaACPS and AaSTSI gene are mainly expressed in leaf tissues, rather than the trichomes, the tissue-specific expression patterns are significantly different from other AaTPS genes. AaECS is mainly express in trichomes, but although the highest expression level was detected in old leaf trichomes. This finding is consistent with other research obtained using a high dihydroartemisinic acid chemotype (Graham et al., 2010; Olofsson et al., 2011). Apparently, the expression patterns of AaTPSs have a few differences to studies using transgenic plants expressing TPS promoter-GUS fusions (Wang et al., 2013; 2014), it is possible that one or more cis-acting regulatory elements, such as enhancer and silencer region, are removed from the cloned AaTPSs promoter region. The differential control of AN biosynthesis as it relates to developmental stage and trichomes maturation and collapse (Arsenault et al., 2010), our studies indicates that the transcripts levels of AaTPSs and artemisinin synthesis genes are more highly correlated with trichomes developmental phases, especially these genes mainly expressed in trichomes. The relative amounts of the TPS genes were calculated using the 2^−∆ΔCT method, AaADS is the dominating AaTPS gene in young leaf where the biosynthesis of artemisinin precursors take place. The expression of AaSQS in leaf was slightly higher than isolated trichomes, indicating that the gene should be also express mainly in other parts of leaf instead of trichomes. Earlier studies showed that down-regulation of AaSQS by hairpin-RNA-mediated gene silencing in A. annua resulted in a 3-fold increased artemisinin production (Zhang et al., 2009). However, the expression pattern of AaSQS suggests that this enzyme may be involved in substrate competition, but not necessarily play a key role in reducing artemisinin production.

Materials and Methods

Plant materials

A. annua was obtained from Youyang of Chongqing province, China. Plant was cultivated and maintained at medicinal botanical garden of Guangdong Pharmaceutical University. The plant materials of root, stem, young and old leaves (leaf node 1-5 and node 11-15 respectively, counting from the apical leaf primordia) were collected separately from approximately 6 months old plant in mid August. Then, plants were allowed to develop flower bud at end of august and harvested the early stage flower bud. All material was flash frozen in liquid nitrogen for storage at -80 °C preceding extraction and analysis. Trichomes were visualised on the adaxial surface of the terminal leaflets of leaves with an Olympus fluorescent microscope (BX51, green excitation 550 nm and red emission 620 nm). Images were recorded using Bioquant Osteo, 2009 V9.00.XP software (Bioquant Image Analysis Corp. Nashville, US).

cDNA preparation

The tissues of root, stem, young and old leaves, flower bud were used for total RNA extraction. The trichomes were...
isolated from the tissues of young and old leaves, respectively. Typical, the tissues were flash frozen in liquid nitrogen prior to abrasion with crushed dry ice for 30 sec in 50 mL plastic collection tubes. Tissues were kept frozen by frequent immersion in liquid nitrogen during this process. Isolated trichomes adhered to the walls of the collection tubes and fragmented leaf tissues were removed by inversion of the tube and gentle agitation. Trichomes isolations at a time were pooled by vortexing with a total of 2 mL Trizol Reagent (Invitrogen). The trichomes suspensions were filtered sequentially through 150 μm and 60 μm nylon sieves. Trichomes enrichment was routinely checked during preparation by microscopy. Trichomes preparation that suspended in Trizol Reagent was ground directly using a mortar and used for total RNA extraction. Total RNA was extracted from frozen ground root, stem, young and old leaves, flower bud and the ground trichomes preparation with Trizol Reagent according to the manufacturer’s instructions. The isolated total RNA was further purified using the RNeasy Mini Kit (Qiagen), treated with the DNase I (Fermentas) to remove genomic DNA contamination and used for first-strand cDNA synthesis using SuperScript II reverse transcriptase and Oligo (dT)18 primer (Invitrogen) according to the manufacturer’s protocol.

Identification of novel AaTPSs in A. annua

Using the A. annua nucleotide sequences that were made publicly available from the National Center for Biotechnology Information (NCBI) conducted the screening of TPS expressed transcripts with known TPS sequences from the non-redundant protein database. In particular, the known TPS sequences of A. annua were used as a query sequence for TBLASTN searches of Expressed Sequence Tags (EST) and Transcriptome Shotgun Assembly (TSA) network databases. These candidate TPS sequences were then used as seed queries to search against the NCBI EST and TSA databases. The ESTs in each group were used to construct contigs with CAP3 sequence assembly program where overlapping regions of EST sequences show at least 95% sequence identity. The resulting contigs and singlets were used as BLASTx queries against the non-redundant protein database to confirm the assembly quality and the relationships between these and the known AaTPSs. AaTPSs were further manually curated to accomplish a more accurate gene annotation considering available ESTs as well as TSA sequences after assembled into transcripts by computational methods (Table S1). Three monoterpene synthases (AaMTS1, AaMTS2 and AaMTS3) and two sesquiterpene synthases (AaSTS1 and AaSTS2) genes were confirmed by PCR amplification with gene-specific primers based on EST sequence information. The cDNA used as template was synthesized from flower bud. PCR amplification was performed with the DNA polymerase (Takara) and following PCR conditions: 2 min at 94 °C and then 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, followed by 72 °C for 6 min. The PCR products were separated on a 2 % agarose ethidium bromide gel, purified from the gel (Omega) and ligated into T-vector (Takara).

Detection of AaTPSs in tissues by RT-PCR

Alignment of A. annua TPS cDNA sequences allowed the design of gene-specific primers for their PCR amplification, while β-actin was used as an internal control (Table S3). The cDNA templates of root, stem, young and old leaves and flower bud were used to detect the expression of AaTPSs through semi-quantitative RT-PCR. The cycling protocol was used as follows: 94 °C for 2 min, 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, followed by 72 °C for 6 min. For all experiments of semi-quantitative RT-PCR products were separated on 2 % agarose ethidium bromide gel and exposed to UV light for documentation.

Comparative transcript profiling of AaTPSs in leaves and trichomes by qRT-PCR

All primers were confirmed firstly by RT-PCR to check if the primers can successful amplify PCR products in samples of young and old leaves (Table S3). Then primers succeeding in amplifying bands were tested for real-time quantitative RT-PCR (qRT-PCR). The expression profiling of 4 monoterpene synthase (AaMTS), 7 sesquiterpene synthase (AaSTS) and AaSQS genes were detected in young and old leaves and their trichomes by qRT-PCR. The qRT-PCR was performed using diluted cDNA, 200 nM of gene-specific primers, and SYBR Green PCR mix on 96-well optical PCR plates using the 7500 Sequence Detection System (Applied Biosystems, USA). The qRT-PCR cycling was performed as follows: 50 °C for 2 min, 95 °C for 10 min, 40 cycles at 95 °C for 15 s, 60 °C for 1 min and finally a dissociation stage at 95 °C for 15 s, 60 °C for 1 min, 95 °C for 15 s. Triplicates of all samples were run, and a negative control of the Master Mix in addition of primers was performed in all qRT-PCR runs.

Expression analysis of artemisinin biosynthetic genes by RT-PCR and qRT-PCR

The cDNA templates of root, stem, young and old leaves and flower bud were used to detect the expression of artemisinin biosynthetic genes (AaADS, AaCYP71AV1, AaDBR2, AaCPR and AaALDH1) through RT-PCR and qRT-PCR (Table 3). The cycling protocol of RT-PCR was used as follows: 94 °C for 2 min, 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, followed by 72 °C for 6 min. Then primers succeeding in amplifying bands were tested for qRT-PCR. The cycling protocol of qRT-PCR was performed as follows: 50 °C for 2 min, 95 °C for 10 min, 40 cycles at 95 °C for 15 s, 60 °C for 1 min and finally a dissociation stage at 95 °C for 15 s, 60 °C for 1 min, 95 °C for 15 s. Triplicates of all samples were run, and a negative control of the Master Mix in addition of primers was performed in all qRT-PCR runs.

Data analysis

Multiple sequence alignments were generated with the ClustalW program. Unrooted phylogenetic tree and molecular evolutionary analyses were conducted using MEGA version 5.05 (Tamura et al., 2011). The gene-specific primers for RT-PCR and qRT-PCR analysis were designed using Primer 3 software (http://bioinfo.ut.ee/primer3), and the specificity of primer pairs was confirmed by BLASTn with the nucleotide sequences of the non-redundant set of A. annua transcripts generated. The primer sequences of all the genes used in this study are listed in Table S3. The specificity of qRT-PCR reaction was verified by dissociation curve analysis. Two independent biological replicates for each sample and three technical replicates of each biological replicate were analyzed for qRT-PCR analysis. The transcript levels of each gene in different tissue samples were normalized with the transcript level of the most suitable internal control gene, β-actin. Relative expression levels were calculated using the REST 2009 software V. 2.0.13 (Qiagen, Hilden, Germany) (Pfaffl et al., 2002).
Conclusion

The present studies were to increase our understanding of terpene metabolism in the plant A. annua, and to evaluate the expression patterns for the pathogenesis to improve the biosynthesis of artemisinin. The expression of majority AaTPS genes were much lower than AaADS in young leaves and flower bud, and their influences on artemisinin yield appear to be relatively limited. However, a similar tissue-specific expression pattern of the six genes, AaADS, AaGaS, AaST1, AaBPS, AaMTS1 and AaMTS2, respectively, was observed between young and old leaves trichomes producing artemisinin and its precursors, down-regulation of these AaTPSs except AaADS will most likely increase artemisinin production in A. annua. On the other hand, five AaTPS (AaCP5, AaECS, AaBFS, AaST1 and AaLS) and AaQS genes have completely different tissue-specific expression patterns, these genes should have no obvious effect on the yield of artemisinin. Characteristics of sequence motifs, phylogenetic relationships and transcript levels of AaTPSs provide a foundation for future functional analysis. The aim of our studies contributes more data on the tissue-specific expression of AaTPSs in A. annua, and provides clues to dissect the concerted regulation for AaTPS expression and the possible competitor for farnesyl dipiphosphate in trichomes.

Acknowledgements

This research was supported by grants from the Scientific Research Startup of Guangdong Pharmaceutical University (no. 43555022), China Postdoctoral Science Foundation (no. 20090460780) and Administration of Traditional Chinese Medicine of Guangdong Province, China (NO 20122107).

References


Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucl Acids Res. 30:e36.


