Article type : Original Article

Overexpression of a type I isopentenyl pyrophosphate isomerase of *Artemisia annua* in the cytosol leads to high arteannuin B production and artemisinin increase

Dongming Ma[1,2], Gui Li[1], Fatima Alejos-Gonzalez[1], Yue Zhu[1], Zhen Xue[3], Aimin Wang[3], Hui Zhang[3], Xing Li[4], Hechun Ye[3], Hong Wang[5], Benye Liu[6], De-Yu Xie[1]*

[1] Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA

[2] Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine, Guangzhou, China

[3] Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China

[3] Institute of Botany, Chinese Academy of Sciences, Beijing, China

[5] University of Chinese Academy of Sciences, Beijing, China

[6] Institute of Pharmaceutical Biology, Technical University Braunschweig, Braunschweig, Germany

*Corresponding author: dxie@ncsu.edu

Running title: Metabolic engineering of arteannuin B and artemisinin

Key words: *Artemisia annua*; artemisinin; arteannuin B; artemisinic acid; Isopentenyl pyrophosphate isomerase; isoprene

Summary

We recently characterized a gene-terpene network that is associated with artemisinin biosynthesis in self-pollinated (SP) *Artemisia annua*, an effective antimalarial plant. We hypothesize that an alteration of gene expression in the network may improve the production of artemisinin and its precursors. In this study, we cloned an isopentenyl pyrophosphate isomerase (*IPPI*) cDNA, *AaIPPI1*, from *Artemisia annua* (Aa). The full-length cDNA encodes a type I IPPI containing a plastid transit peptide (PTP) at its amino terminus. After removal of the PTP, the recombinant truncated AaIPPI1 isomerized isopentenyl pyrophosphate (IPP) to dimethyl allyl pyrophosphate (DMAPP) and *vice versa*. The steady-state equilibrium ratio of IPP/DMAPP in the enzymatic reactions was approximately 1:7. The truncated AaIPPI1 was overexpressed in the cytosol of the SP *A. annua* variety. The leaves of transgenic plants produced approximately 4% arteannuin B (g/g, dry weight, dw) and 0.17-0.25% artemisinin (g/g, dw), the levels of which were significantly higher than those in the leaves of wild-type plants. In addition, transgenic plants showed an increase in artemisinic acid production of more than 1% (g/g, dw). In contrast, isoprene formation was significantly reduced in transgenic plants. These results provide evidence that the overexpression of this *AaIPPI1* in the cytosol can lead to metabolic alterations of terpenoid biosynthesis and show that its transgenic plants have the potential to yield high production levels of arteannuin B as a new precursor source for artemisinin.

Introduction

Isopentenyl pyrophosphate isomerase (IPPI, also called isopentenyl diphosphate isomerase, IDPI) catalyzes the interconversion of isopentenyl pyrophosphate (IPP, or isopentenyl diphosphate, IDP) and dimethylallyl pyrophosphate (DMAPP, or dimethylallyl diphosphate, DMADP) (Ramos-Valdivia *et al.*, 1997, Oh *et al.*, 2000, Berthelot *et al.*, 2012, Zhou *et al.*, 2013). IPPI is classified into types I and II according to the cofactors required for its catalytic activity. Catalysis by type I IPPI, which is found in animals, fungi, some bacteria, and plants, requires Mg^{2+} and zinc cofactors (Carrigan and Poulter, 2003, Lee and Poulter, 2006). Catalysis by type II IPPI, a flavoprotein found in archaea and some bacteria, requires the reduced form of flavin mononucleotide and Mg^{2+} (Kaneda *et al.*, 2001, Sharma *et al.*, 2010).

In all plants investigated to date, type I has been found to be the form of the enzyme that maintains the steady-state levels of IPP and DMAPP (Fig. 1a) (Albrecht and Sandmann, 1994, Kajiwara *et al.*, 1997, Gallagher *et al.*, 2003, Phillips *et al.*, 2008). Over the past 10 years, fundamental progress has been achieved in localizing IPPI in the cytosol, plastids, mitochondria, and peroxisomes (Okada *et al.*, 2008, Phillips *et al.*, 2008, Sapir-Mir *et al.*, 2008, Vranova *et al.*, 2013). These studies have shown that plant cells employ a complicated metabolic compartmentalization of isoprenoids. Numerous studies have also shown that the formation of DMAPP in the cytosol and in plastids occurs through different enzymatic mechanisms. The MVA pathway has been demonstrated to produce IPP only (Fig. 1a) (Hoeffler *et al.*, 2002); thus, DMAPP in the cytosol of plants results from the isomerization of IPP catalyzed by a type I IPPI. In plastids, in addition to the main production of DMAPP via IPP isomerization, recent reports have shown that HDR (Fig. 1a) in the MEP pathway converts 4-hydroxy-3-methyl-but-2-enyl pyrophosphate (HMBPP) to both IPP and DMAPP at a steady-state ratio of approximately 6:1 (Rohdich *et al.*, 2003, Brandt *et al.*, 2004, Eisenreich *et al.*, 2004). Given that DMAPP is the primary starter C5 molecule (Fig. 1a), its pool size affects the biosynthesis of all downstream isoprenoids in plants, including the isoprene, monoterpenes (C10), and diterpenes (C20) synthesized in plastids and the sesquiterpenes and triterpenes synthesized in the cytosol (Croteau *et al.*, 2000, Okada *et al.*, 2008, Phillips *et al.*, 2008). Therefore, IPPI activity and the IPP/DMAPP ratio play essential roles in the regulation of plant terpenoid metabolism.

Artemisia annua produces artemisinin, an effective antimalarial endoperoxide sesquiterpene lactone (Liu *et al.*, 1979, Klayman, 1985, Miller and Su, 2011). Artemisinin-based combination therapy (ACT) is the first-line treatment for malaria and saves the lives of millions of patients yearly (WHO, 2006, WHO, 2010, Ansari *et al.*, 2013, Taylor, 2013). *A. annua* is a diploid cross-hybridized species (Delabays *et al.*, 1993, Xie *et al.*, 1995, Graham *et al.*, 2010, Alejos-Gonzalez *et al.*, 2011) that contains nine pairs of chromosomes. Over the past three decades, numerous breeding studies have been undertaken to create multiple hybrid cultivars of this species (Delabays *et al.*, 1993, Delabays *et al.*, 2001, Graham *et al.*, 2010, Paul *et al.*, 2014), many of which have reported an increase in the production of artemisinin.

However, the yield of artemisinin from hybrid cultivars does not yet meet the demand for ACT (Paddon *et al.*, 2013). In addition, the price of artemisinin has been unstable, and the cost of therapy is high for patients from countries in which malaria is endemic (White, 2008, Shretta and Yadav, 2012, Weathers and Towler, 2014).

The use of artemisinin precursors from *A. annua* for *in vitro* semi-synthesis has recently emerged as an efficient alternative approach to the production of artemisinin (Paddon *et al.*, 2013). Over the past 10 years, artemisinic acid (Fig. 1a) has been the main synthetic precursor used in the metabolic engineering of artemisinin. The biosynthetic pathway from armorpha-4,11-diene to artemisinic acid and dihydroartemisinic acid has been biochemically characterized by research efforts in enzyme analysis, gene cloning, and synthetic approaches (Bouwmeester *et al.*, 1999, Chang *et al.*, 2000, Ro *et al.*, 2006, Teoh *et al.*, 2006). As a significant result, the recent success of the synthetic biology of artemisinic acid coupled with the semi-synthesis of artemisinin has led to the industrial-scale production of these compounds to supplement ACT (Ro *et al.*, 2006, Teoh *et al.*, 2006, Paddon *et al.*, 2013, Turconi *et al.*, 2014), although the cost of this synthetic technology remains to be improved (Peplow, 2016). As is well understood, *A. annua* also produces arteannuin B and a wide array of diverse terpenoids (Fig. 1a) (Brown, 2010, Ma *et al.*, 2015). Chemical synthesis has demonstrated that arteannuin B is another appropriate precursor molecule for the semi-synthesis of artemisinin (Nowak and Lansbury, 1998) and for biotransformation to artemisinin (Akhila *et al.*, 1987, Nair and Basile, 1993, Dhingra and Lakshmi Narasu, 2001). These studies have demonstrated that the availability of a rich arteannuin B source may offer a new route to the production of sufficient artemisinin for the treatment of malaria.

To innovate new approaches to increase artemisinin production, we recently developed a self-pollinated (SP) variety of *A. annua* that produces artemisinin, artemisinic acid, and arteannuin B (Alejos-Gonzalez *et al.*, 2011). In addition, the SP plants are capable of regeneration via both organogenesis and somatic embryogenesis from plant tissue culture (Alejos-Gonzalez *et al.*, 2013, Ma *et al.*, 2015) and are highly transformable for genetic modification. Furthermore, an integrative metabolomics and transcriptomic study of *A. annua*,

which characterized its gene-to-terpene landscape, showed that the biosynthesis of artemisinin and artemisinic acid is closely associated with the plant's development from the juvenile to the flowering phase. Unlike the content of artemisinin and artemisinic acid, which is highly dependent upon the spatial position and development of the plant, the content of arteannuin B is approximately 0.7-1.3% g/g dry weight (dw) in all analyzed tissues, including the leaves of young seedlings as well as later flowers (Ma *et al.*, 2015). These superior properties allow the development of a new platform of SP plants to improve the production of artemisinin and its precursors. In addition, Weather *et al.* recently demonstrated that dried leaf products (such as leaf tablets) are an effective and inexpensive therapy for the treatment of malaria (Weathers *et al.*, 2014a, Weathers *et al.*, 2014b, Weathers and Towler, 2014). Based on these discoveries, our SP plants may provide a new resource for the development of leaf products for the treatment of malaria.

Recently, we characterized a gene-terpene network that controls artemisinin biosynthesis in SP *A. annua* (Ma *et al.*, 2015). An uncharacterized *IPPI* homolog (namely, *AaIPPI1*) together with *ADS*, *CYP71AV1*, *DBR2*, and *CPR1* is located in the network (Fig. 1b). We hypothesize that an overexpression of this *AaIPPI1* may increase artemisinin biosynthesis in plants. In this study, an *AaIPPI1* cDNA was cloned from *A. annua*. Sequence analysis of *AaIPPI1* showed that it encodes a type I IPPI. A recombinant truncated AaIPPI1, (-tp)AaIPPI1, was purified for enzymatic analysis. The (-tp)AaIPPI1 catalyzed the reversible isomerization of IPP and DMAPP at a final steady-state equilibrium ratio of approximately 1:7. GFP fusion and confocal microscopic analysis further demonstrated that the truncated *(-tp)AaIPPI1* cDNA fragment encodes an enzyme that is localized in the cytosol. The *(-tp)AaIPPI1* cDNA fragment was overexpressed to increase the level of the enzyme in transgenic self-pollinated *A. annua* plants. Metabolic analysis showed that transgenic plants yielded approximately 4% arteannuin B (g/g, dw) and that they contained significantly increased contents of artemisinin. In addition, isoprene formation was significantly reduced in the transgenic plants. This study provides evidence showing that overexpression of genes in the network can fundamentally increase the production of artemisinin and its precursors. In addition, *AaIPPI1* transgenic plants could supply a new resource to produce arteannuin B for artemisinin synthesis.

Results

Isolation of *AaIPPI1* cDNA, sequence characterization, and enzymatic activity

An expressed sequence tag (EST) fragment was obtained from a glandular trichome cDNA library (Ma *et al.*, 2009). Sequence alignment of this fragment revealed its high identity to *IPPI* cDNA homologs in the gene bank curated by NCBI. The full-length sequence (1148 bp) contains an 852-bp-long open reading frame (ORF) that encodes 283 amino acids (Fig. S1). We named this cDNA *AaIPPI1* (Fig. 1b). An alignment with the amino acid sequences of seven IPPI homologs revealed that the first 59 N-terminal amino acids of AaIPPI1 did not align with four other reported homologs (Fig. S1). Further analysis using the TargetP program (Emanuelsson *et al.*, 2007) indicated that these 59 amino acids formed a plastid transit peptide (PTP) encoded by the 5' end of the ORF. Moreover, the alignment showed that the amino acid sequence of AaIPPI1 contained a conserved cysteine residue in a TNTCCSHPL motif and a conserved glutamate residue in a WGEHELDY motif, which are typical properties of the type I IPPI family (Hahn *et al.*, 1999). Thus, AaIPPI1 is a member of the type I IPPI family. Gene expression profiling showed that *AaIPPI1* was expressed in five tissues, with relatively high expression levels in trichomes, leaves, and flowers (Fig. 2a-b).

The full-length ORF of the *AaIPPI1* cDNA, *(tp)AaIPPI1*, was cloned into a pET30a (+) vector (Fig. S2a) and used to express a recombinant protein in *E. coli*. However, after lysis of *E. coli*, the resulting recombinant protein, (tp)AaIPPI1, was found only in inclusion bodies, but it was not detected in the extraction buffer (Fig. S2b). To obtain soluble protein, the first 177 nucleotides, which encode the first 59 amino acids forming the PTP, were removed from the 5' end of the ORF cDNA to produce a truncated cDNA fragment, *(-tp)AaIPPI1*. When the *(-tp)AaIPPI1* cDNA fragment was introduced into the pET30a (+) vector (Fig. S2a), expression of the construct in *E. coli* yielded a soluble protein in the extraction buffer. This protein was further purified (Fig. S2c) for the enzyme assays.

An enzyme assay of the recombinant (-tp)AaIPPI1 was performed using IPP and DMAPP, two allylic prenyl diphosphate compounds, as substrates. Under alkaline phosphatase

conditions, IPP and DMAPP are readily converted into 3-methyl-3-buten-1-ol (isoprenol) and 3-methyl-2-buten-1-ol, respectively (Ramos-Valdivia *et al.*, 1997, Hoshino *et al.*, 2006). These two alcohol compounds can be easily measured by GC and GC-MS analysis alone; therefore, the activity of (-tp)AaIPPI1 was measured using these two alcohol compounds. After incubation of IPP with recombinant (-tp)AaIPP1, the steady-state equilibrium products of isomerization included IPP and DMAPP, which were converted to their corresponding isoprenol and 3-methyl-2-buten-1-ol, respectively (Fig. 3a and b). Measurement of the levels of isoprenol and 3-methyl-2-buten-1-ol showed that their final steady-state ratio was approximately 1:7 (Fig. 3a), indicating that approximately 87.5% of the IPP was isomerized to DMAPP. After incubation of DMAPP with recombinant (-tp)AaIPP1, the final steady-state equilibrium products of the isomerization reactions also included IPP and DMAPP, which were converted to their corresponding isoprenol and 3-methyl-2-buten-1-ol, respectively (Fig. 3b). The ratio of isoprenol to 3-methyl-2-buten-1-ol was also approximately 1:7 (Fig. 3b), indicating that 12.5% of the DMAPP substrate was isomerized to IPP. Similar results were obtained after multiple repetitions of the enzymatic assay using various reaction times (e.g., 1 and 2 h) and substrate concentrations (Fig. S3). These results demonstrated that the recombinant (-tp)AaIPPI1 catalyzed the interconversion of IPP and DMAPP, reaching a final steady-state equilibrium ratio of approximately 1:7. In addition to phosphatase, hydrochloric acid (1 N) was used to dephosphorylate IPP and DMAPP isomerized by recombinant (-tp)AaIPPI1. The resulting data supported the ratio of approximate 1:7 of IPP to DMAPP (Fig. S3, a-f). Experiments using two concentrations of IPP or DMAPP (20 and 40 µg/0.1 mL) also showed a ratio of approximate 1:7 (Fig. S3, a-f). Experiments using 10 µg/0.1 mL of IPP or DMAPP and dephosphorylation with alkaline phosphatase also resulted in a steady-state equilibrium ratio of IPP/DMAPP of approximately 1:7 (Fig. S3, g-h).

Cytosolic localization of (-tp)AaIPPI1

The cytosolic localization of (-tp)AaIPPI1 was analyzed by confocal microscopy. After removal of its stop codon (TAA), the resulting *(-tp)AaIPPI1* was adjacently fused to the ATG of an *EGFP* in the pSITEII-EGFP-N1 binary vector to obtain a new binary plasmid, pSITEII-(-tp)AaIPPI1-EGFP (Fig. 4a). This vector was introduced into *Nicotiana*

benthamiana for transient expression. Observation of infected plants by confocal microscopy showed that the fused (-tp)IPPI1-EGFP protein was localized to the cytosol (Fig. 4a).

Overexpression of *(-tp)AaIPPI1* in self-pollinated *A. annua*

The *(-tp)AaIPPI1* cDNA with its stop codon (TAA) was introduced into another binary vector, PMDC84. In the resulting recombinant binary plasmid, PMDC84-(-tp)AaIPPI1 (Fig. 4b), *(-tp)AaIPPI1* was driven by a 2 x 35S promoter but was not fused to GFP protein (Fig. 4b). The binary vector PMDC84-(-tp)AaIPPI1 was introduced into self-pollinated *A. annua*. Multiple transgenic plants were obtained from genetic transformation. Among these, three lines, namely, OE-T1, OE-T15, and OE-T17 (Fig. 4b), were selected for further transgene expression analysis. Both semi-quantitative RT-PCR and qRT-PCR analyses using gene-specific primers (Table S1) revealed the expression of *AaIPPI1* in both wild-type and transgenic plants (Fig. 4c). The expression level of *AaIPPI1* was much higher in transgenic plants than in wild-type plants (Fig. 4c), which indicated that the overexpression of the *(-tp)AaIPPI1* transgene increased the level of *AaIPPI1* messenger RNA. Western blot analysis using an AaIPPI1 antibody further showed that the level of AaIPPI1 was much higher in transgenic than in wild-type plants (Fig. S2d). This result demonstrated that overexpression of the *(-tp)AaIPPI1* transgene significantly increased the total content of AaIPPI1 protein in transgenic plants (Fig. S2d). As shown in Fig. 4b, the growth phenotypes of transgenic and wild-type plants in the phytotron were similar. In addition to the three aforementioned transgenic lines, parallel transgenic lines (e.g., OE-T2 and OE-T13), which were demonstrated by semi-quantitative RT-PCR, qRT-PCR, and western blot analysis (Fig. S4b-d), also showed a similar growth phenotype to WT plants. When *(-tp)AaIPPI1* transgenic plants were transferred to the greenhouse for the production of seeds, their growth phenotypes were similar to those of wild-type (Fig. S4a). In our investigation, empty vector and other artemisinin pathway gene (such as *ADS and HDR*) transgenic plants (Ma *et al.*, 2015, Ma *et al.*, 2017) were routinely generated as a control to assess potential effects of different transgenes on plant growth and other features. *A. annua HDR* (*AaHDR1*) transgenic plants (Ma *et al.*, 2017) were used as control to evaluate effects of binary construct on *AaIPPI1* expression and plant growth. As observed for OE-T15 and OE-T17 (Fig. 4c), the

growth of *AaHDR1* transgenic plants was similar to wild-type plants (Fig. S5a). Quantitative RT-PCR analysis showed that the expression levels of *AaIPPI1* in wild-type and *AaHDR1* transgenic were similar, but significantly lower than in *(-tp)AaIPPI1* transgenic (OE-T15, 17, and 3) plants (Fig. S5b). These data demonstrated that the high *AaIPPI1* expression level in OE-T1, 2, 13, 15, 17, and 23 (Fig. 4c, Fig. S4b and c, and Fig. S5b) resulted from the *(-tp)AaIPPI1* overexpression.

GC-MS and LC-MS analyses of artemisinic acid, arteannuin B, and artemisinin

GC-MS analysis was performed to identify and measure the levels of artemisinic acid and arteannuin B in transgenic and wild-type plants. Metabolites were ionized using electron ionization. Total ion chromatography showed that the peak size of artemisinic acid was greater in two transgenic lines, OE-T15 and OE-T17, than in wild-type plants (Fig. 5a). Based on a standard curve (Fig. S6), quantification showed that the content of artemisinic acid in the OE-T15 and OE-T17 lines was approximately 1.3% (g/g, dry weight), which was slightly higher than that in wild-type plants (Fig. 6a). Total ion chromatography also showed that the peak size of arteannuin B was much greater in transgenic than in wild-type plants (Fig. 5b). Based on a standard curve (Fig. S6), quantification showed that the content of arteannuin B was 2.6-4% (g/g, dry weight) in three transgenic lines (OE-T1, OE-T15, and OE-T17), which was significantly higher than that in wild-type plants (Fig. 6b). The arteannuin B content in other transgenic plant lines, such as OE-T2 and OE-T13, was also increased (Fig. S4e), with values within this range.

LC-MS analysis was conducted to identify and measure artemisinin. The selected ion chromatography [M+1] recorded at $[283]^+$ showed that the peak size of artemisinin was greater in transgenic than in wild-type plants (Fig. S7). Tandem MS analysis also showed that the abundance of the MS/MS fragment was greater in transgenic than in wild-type plants (Fig. S8). Based on an artemisinin standard curve (Fig. S6), the calculation showed that the content of artemisinin was 0.17-0.26% (g/g, dry weight) in the three transgenic lines, OE-T1, OE-T15, and OE-T17, which was 2- to 3-fold higher than that in the wild-type plants (Fig. 6b). Quantification of artemisinin also demonstrated the increase in artemisinin contents (Fig. S4d)

in the two other transgenic lines (OE-T2 and OE-T13) that were grown in a greenhouse (Fig. S4a).

Decrease in isoprene formation in transgenic plants

Isoprene is directly biosynthesized from DMAPP (Fig. 1a). We used a fast isoprene sensor (FIS) to analyze isoprene emission from *A. annua* as described previously (Xi *et al.*, 2016). In FIS, isoprene is converted to photons, which are recorded by a computer. Multiple live recordings showed that not only the entire seedling but also the excised leaves emitted isoprene (Fig. 6c). At the beginning of recording (defined as 0 sec), when the samples were placed in a sealed chamber, no isoprene was detected. After a few or 10 sec, FIS began to detect isoprene. When leaves of transgenic plants were placed in the chamber, the emission trend in the beginning was similar to that of wild-type samples (Fig. 6c). In comparison to the wild-type control samples, isoprene emission from the leaves and seedlings of transgenic lines, e.g., OE-T15 and OE-T17, was reduced (Fig. 6c). The average levels of isoprene emission per second over a 40-min period were calculated. The resulting data showed that the emission of isoprene from leaves of OE-T15 and OE-T17 plants was significantly reduced (Fig. 6d).

Discussion

Increasing artemisinin production is one of global efforts in the fight against malaria. The transgenic plants developed in this study have the potential to provide arteannuin B as an alternative precursor for the semi-synthesis of artemisinin in the future. Both chemical synthesis and bioconversion have been used to convert arteannuin B to artemisinin (Nair and Basile, 1993, Nowak and Lansbury, 1998). Nair and Basile (1993) observed that under optimized buffer conditions, a crude enzyme extract could convert 80% of arteannuin B to artemisinin. This conversion rate is much higher than 40-45%, the rate obtained by the pharmaceutical industry to convert artemisinic acid to artemisinin (Paddon *et al.*, 2013). However, to date, arteannuin B has not been used as a precursor for the semi-synthesis of artemisinin, which can be explained by two scenarios inferred from previous studies. One explanation is likely associated with economic infeasibility, although the total synthesis of

arteannuin B in the laboratory has been reported (Lansbury and Mojica, 1986); the other is the low production of arteannuin B in plants. Nair and Basile (1993) were able to extract only 23 mg arteannuin B from 1 kg dry leaves, representing approximately 0.0023% (g/g, dry weight). Woerdenbag *et al.* reported the presence of approximately 0.1% (g/g, dry weight) arteannuin B in leaves, flowers, and young stems (Woerdenbag *et al.*, 1991) and of 0.08% arteannuin B in an artemisinin high-production variety (Woerdenbag *et al.*, 1994). These studies show that it is not economically feasible to isolate arteannuin B for industrial-scale artemisinin semi-synthesis. Therefore, to use arteannuin B as an alternative precursor source, increasing its production is necessary. We recently observed that the arteannuin B content in the above-ground tissues of our self-pollinated *A. annua* plants ranges from 0.7-1.3% (g/g, dw) during different developmental stages. In addition, the content of arteannuin B is not obviously affected by plant development from the early young seedling stages to the late flowering stage (Ma *et al.*, 2015). Furthermore, this property is also consistent in the progeny. Therefore, self-pollinated plants provide an appropriate platform for metabolic engineering to produce a high yield of arteannuin B. In the present study, the leaves of 7-week-old seedlings of transgenic plants (Fig. 4b) showed a significant increase in arteannuin B content compared with wild-type plants, and two lines, OE-T15 and OE-T17, produced approximately 4% (g/g, dw) arteannuin B (Fig. 6a). This high production suggests that transgenic plants may provide a viable resource for the extraction of arteannuin B in the future. In addition, the contents of artemisinic acid were increased in these two lines (Fig. 5a and Fig. 6a). Certainly, different transgenic effects on artemisinic acid contents were observed in multiple transgenic lines. For example, unlike in OE-T15 and OE-T17 lines, the content of artemisinic acid was not significantly increased in line OE-T1 (Fig. 5a and Fig. 6a). This result was likely associated with the lower protein level in OE-T1 compared with OE-T15 and OE-T17 (Fig. S2d).

The high production of arteannuin B and the increased levels of artemisinin that resulted from the increase in AaIPPI1 levels in the cytosol of transgenic plants described herein are related to the reduction of isoprene levels. IPPI catalyzes the interconversion of IPP and DMAPP, the C_5 building blocks of all short- and long-chain terpenoids, including monoterpenoids, sesquiterpenoids, diterpenoids, steroids, and cholesterols (Fig. 1a) (Muehlbacher and Poulter,

1988, Anderson *et al.*, 1989, RamosValdivia *et al.*, 1997, Croteau *et al.*, 2000, Berthelot *et al.*, 2012, Heuston *et al.*, 2012, Vranova *et al.*, 2013, Zhou *et al.*, 2013, Rodriguez-Concepcion and Boronat, 2015). DMAPP is the primary starter molecule to which IPP is linked by prenyltransferases to form short- and long-chain isoprenoids, e.g., sesquiterpenoids resulting from the condensation of two IPP and one DMAPP. Our *in vitro* experiments showed that (-tp)AaIPPI1 isomerized IPP and DMAPP at a final ratio of approximately 1:7 at steady-state equilibrium. Based on this property of the enzymatic reaction, it can be hypothesized that the overexpression of (-tp)AaIPPI1 increases the content of DMAPP available for artemisinin biosynthesis in the cytosol. This hypothesis is supported by the observation that the transgenic plants displayed a highly increased production of arteannuin B and enhanced artemisinin content (Fig. 6a and b). This result actually raises an important question: if more IPP molecules are isomerized to DMAPP, leading to increased production of arteannuin B and artemisinin in the cytosol, what is the source of the additional IPP? Based on the literature reported for multiple plant species, IPP can be transported from plastids to the cytosol (Fig. 1a) (Vranova *et al.*, 2013). Accordingly, it can be hypothesized that an increase in the transport of IPP from plastids to the cytosol may lead to a reduction of DMAPP levels in plastids. To investigate this possibility, we performed an FIS analysis. The resulting FIS data showed that the emission of isoprene by transgenic plants was reduced (Fig. 6c and d). Given that isoprene is the direct product of DMAPP in plastids (Fig. 1a), the results of the FIS analysis indicated a reduction in the levels of this substrate. This result is supported by our recent study of a synthetic plant-insect geranyl pyrophosphate synthase (GPPS) that condenses IPP and DMAPP to downstream terpenoid molecules (Fig. 1a). When the synthetic GPPS was overexpressed in the cytosol, the transgenic *Camelina* plants showed a significant decrease in isoprene emission and an increase in triterpenes, and enhanced plant growth (Xi *et al.*, 2016), suggesting that the transport of IPP from the plastids to the cytosol was increased. Our results also suggest that overexpression of AaIPPI1 in the cytosol may be useful for improving the production of artemisinin in other *A. annua* varieties. Genetic breeding studies have shown that there are two chemotypes of *A. annua*, those with high-artemisinin production (HAP) and those with low-artemisinin production (LAP) (Ting *et al.*, 2013). Our self-pollinated plants belong to the LAP chemotype. Based on the data

reported herein, it is likely that the overexpression of *(-tp)AaIPPI1* in HAP plants would greatly increase artemisinin levels.

Numerous studies have demonstrated a significant role of glandular trichomes in the biosynthesis of artemisinin (Duke and Paul, 1993, Duke *et al.*, 1994, Tellez *et al.*, 1999, Lommen *et al.*, 2005, Teoh *et al.*, 2006, Covello *et al.*, 2007, Xiao *et al.*, 2016). A recent study reported that an increase in the number of trichomes significantly enhanced artemisinin production (Singh *et al.*, 2016). ADS, CYP71AV1, CPR1, DBR2, CYB5, ADH1, and ALDH1 (Fig. 1a) were originally identified in glandular trichomes (Bouwmeester *et al.*, 1999, Ro *et al.*, 2006, Weathers *et al.*, 2006, Zhang *et al.*, 2008, Teoh *et al.*, 2009, Zhang *et al.*, 2009, Paddon *et al.*, 2013). In our study, we also isolated *AaIPPI1* from glandular trichomes. RT-PCR analysis showed that glandular trichomes had the highest levels of *AaIPPI1* transcription, although it was also highly expressed in other tissues (Fig. 2a-b). These data suggest that AaIPPI1 activity may be preferentially associated with the biosynthesis of artemisinin, arteannuin B, and artemisinic acid in glandular trichomes. To date, numerous studies including ours have indicated that the engineering of trichomes may offer a promising approach to the improvement of the artemisinin yield for fighting against malaria.

Experimental Procedures

Plant materials and growth conditions

Two different ecotypes of *A. annua* were used in this study. One, collected from Sichuan Province, China, was named strain #001. This strain is a heterozygous variety that has been grown in greenhouses for leaf materials (Chen *et al.*, 2000, Han *et al.*, 2006). It was used to clone the *AaIPPI1* gene. The other strain is a novel self-pollinated *A. annua* variety (Alejos-Gonzalez *et al.*, 2011, Ma *et al.*, 2015). The progenies of the self-pollinated plants were grown in a phytotron for seeds as reported previously (Alejos-Gonzalez *et al.*, 2011). F3 progeny seedlings grown in the phytotron were used for genetic transformation and functional analysis of *AaIPPI1*. The photoperiod and temperature used in the phytotron were 16/8 (h) and 25°C (Supplementary Methods). Leaves from plant nodes #7-12 (Fig. 4b) were harvested and pooled for gene expression and metabolite analyses (Table S2).

Isolation of *AaIPPI1* cDNA

An EST cDNA library was constructed from glandular secretory trichomes of strain #001 (supplementary methods) (Ma *et al.*, 2009). A cDNA sequence that was highly homologous to that of the isopentenyl diphosphate isomerase gene was identified from the library. Its full-length cDNA was obtained by RACE RT-PCR as reported previously (Ma *et al.*, 2009); it is denoted herein as *AaIPPI1*. Sequence alignment was completed using ClustalW at EMBnet (http://www.ch.embnet.org/software/ClustalW.html).

Heterogeneous expression of *AaIPPI1* and purification of recombinant protein

The ORF of *AaIPPI1* was amplified for heterogeneous expression using a pair of gene-specific primers, 5'-GATGGATCCATGGCAGAAACGTTGGTCTCA-3' (forward primer, BamHI restriction site underlined) and 5'-GTACTCGAGTTACAAGTTATGGATTGTCTT -3' (reverse primer, XhoI restriction site underlined). The amplified cDNA fragment was digested with *Bam*HI and *Xho*I, ligated to a pET-30a (+) vector that had also been digested with *Bam*HI and *Xho*I (Novagen, Billerica, USA), and then purified. The resulting plasmid, pET-*(tp)AaIPPI1* (Fig. S2a), could express a His-tag recombinant protein. In addition, an N-terminal truncated fragment of the *AaIPPI1* cDNA was amplified using another primer pair: forward 5'-GATGGATCCATGACGACAATCTTGACTGAT -3' and reverse 5'-GTACTCGAGTTACAAGTTATGGATTGTCTT-3', which contained *BamH*I and *Xho*I restriction sites (underlined), respectively. The resulting cDNA lacked the first 177 nucleotides of the ORF. This cDNA was ligated to pET30a(+) to generate pET-*(-tp)AaIPPI1* (Fig. S2a). The pET-*(tp)AaIPPI1* and pET*(-tp)AaIPPI1* plasmids were separately transformed into competent cells of *Escherichia coli* strain BL21 (DE3). Two BL21 (DE3) colonies, one containing the pET-*(tp)AaIPPI1* plasmid and the other containing the pET-*(-tp)AaIPPI1* plasmid, were selected for protein expression. *E. coli* cells containing pET-*(tp)AaIPPI1* or pET-*(-tp)AaIPPI1* were cultured in liquid LB medium supplied with 50 mg L^{-1} ampicillin at 37°C. When the OD value of the suspension culture was approximately 0.6 at 600 nm, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture to a final concentration

of 0.2 mM. The suspension culture was then incubated for an additional 6 h at 28°C. The cells were harvested by centrifugation and re-suspended in 10 mL of 0.05 M sodium phosphate buffer (pH 7.5) containing 20 mM imidazole; the suspension was sonicated on ice for 5 min at 50% pulses using an Ultrasonic Crasher. The homogenate was centrifuged for 10 min at 10,000 × g at 4°C. The soluble crude protein was loaded onto a Ni-NTA His-Bind™ Resin column (Novagen, Billerica, US) to purify the recombinant (-tp)AaIPPI1 according to the manufacturer's protocol. The column was washed with 0.05 M sodium phosphate buffer (pH 7.5) containing 20 mM imidazole to remove untargeted proteins. Then, the washed column was eluted with 2.5ml 0.05 M sodium phosphate buffer (pH 7.5) containing 300 mM imidazole to collect recombinant protein. The resulting recombinant (-tp)AaIPPI1 was loaded onto PD-10 columns (Amersham Pharmacia Biotech, Uppsala, Sweden) to further remove salt. The recombinant (-tp)AaIPPI1 was eluted from the PD-10 column in 0.1 M Tris–HCl (pH 7.5) containing 10% (v/v) glycerol and stored at -80°C. SDS-PAGE (10%) was used to assess the quality of purification. The protein concentration was determined by the Bradford method based on a standard curve of bovine serum albumin.

Enzyme assays

Enzyme assays were carried out in a 100-μL volume consisting of 10 μg recombinant (-tp)AaIPPI1, 100 mM Tris-HCl buffer (pH 7.5), 5 mM $MgCl_2$, 10 mM DTT, and 20 μg IPP or DMAPP in a 1.5-mL Eppendorf tube. In addition, 10 μg and 40 μg IPP or DMAPP was used to characterize the isomerization activity of recombinant (-tp)AaIPPI1. Different reaction times from 0.5-3 h were tested to optimize the reaction assays to characterize the isomerization properties. A duration of 3 h was found to be appropriate for the enzyme assay to characterize the steady-state ratio of IPP and DMAPP. After 3 h of incubation at 37°C, alkaline phosphatase (1 U) was added, and the incubation mixture was gently vortexed. The mixtures were incubated for an additional 3 h at 37°C to allow the enzyme to remove diphosphate from IPP and DMAPP to form their corresponding isoprenol and 3-methyl-2-buten-1-ol. After dephosphorylation, 200 μl diethyl ether was quickly added to each tube, followed by six times of rapid and thorough shaking to extract metabolites. The tubes were centrifuged for 5 min at 10000 × g, and the supernatants were pipetted into new

tubes for analysis of metabolites using GC-MS as described below and for GC analyses on an HP-GCD instrument (Agilent). The metabolites were separated using a DB-5 silicon capillary column (30 m × 0.25 mm, Agilent). Helium was used as the carrier gas at a flow rate of 1 mL/min. A splitless mode was used in the inlet. The injection temperature was set to 250°C. The initial oven temperature was set to 40°C for 3 min; the temperature was then ramped to 250°C at a rate of 5°C /min. A positive electron ionization mode was used to characterize the enzymatic products. The mass detector was set to scan the mass range from 40 to 240 m/z.

In addition, dephosphorylation using hydrochloric acid was performed to characterize the isomerization catalyzed by AaIPPI1. For this assay, 20 and 40 µg substrates were used in the 2-h reaction incubation. All other reaction conditions were the same as those described above. After the reaction, the enzymatic products were dephosphorylated for 30 min with alkaline phosphatase (1 U) and then for 40 min with 1 N hydrochloric acid to analyze the ratios of isoprenol to 3-methyl-2-buten-1-ol. After dephosphorylation, 200 µl diethyl ether was quickly added to each tube, which was then was rapidly and thoroughly shaken for six times to extract metabolites. To separate the metabolites, the GC-MS instrument and oven temperature programs were different from those described above. GC-MS analysis was performed using a GC-QqQ MS (7890a-5975b, Agilent, USA) (using an HP-5ms column, 30 m x 0.25 mm x 0.10 µm, Agilent, USA). The initial oven temperature was set to 35°C for 2 min and then ramped to 40°C for 5 min, followed by ramping to 325°C at a rate of 70°C /min and maintenance for 3 min. All other instrumental conditions were the same as those described above.

Transient expression of (-tp)AaIPPI1 in *Nicotiana benthamiana*

The stop codon TAA was removed from the truncated *(-tp)IPPI* cDNA fragment. The resulting cDNA fragment was fused to the upstream of and adjacent to the starter codon ATG of *GFP* in the binary vector pSITEII-EGFP-N1 using the gateway technique described below. The resulting new binary vector, pSITEII-(-tp)AaIPPI1–EGFP (Fig. 4a), was obtained for the protein localization assays. The vector was introduced into *Agrobacterium tumefaciens* strain LBA4404. A positive colony was used to infect *N. benthamiana* leaves according to an

agroinfiltration protocol (Martin *et al.*, 2009). After 48 h, the infected leaves were examined for the green fluorescence of GFP using an LSM 5 PASCAL (Carl Zeiss) confocal laser scanning microscope.

Development of plant expression vectors and transformation of *Agrobacterium tumefaciens*

A gateway technique was used to construct binary vectors for gene expression in plants. The truncated cDNA *(-tp)IPPI* fragment containing the stop codon was cloned into an entry vector, pENTR/D-TOPO (Invitrogen, CA), to obtain a pENTR-(-tp)AaIPPI1 plasmid. The destination vector used herein was PMDC84, which contains a 2 × 35S promoter and a GFP reporter gene (Curtis and Grossniklaus, 2003). The pENTR-(-tp)AaIPPI1 and PMDC-84 plasmids were digested using LR Clonase II enzyme mix (Invitrogen, Carlsbad, USA) and ligated according to the manufacturer's protocol. After the ligation reaction, a new binary vector, PMDC84-(-tp)AaIPPI1 (Fig. 4b), was obtained, in which the *(-tp)AaIPPI1* cDNA fragment was inserted at the immediate downstream of the 2 × 35S promoter but was not fused to *GFP*. The resulting PMDC84-(-tp)AaIPPI1 was used to express transgenic protein in the cytosol of plant cells. By following the protocol described previously (Xie *et al.*, 2004), this binary vector was introduced into *Agrobacterium tumefaciens* strain LBA4404 for genetic transformation of *A. annua* as described below.

Seeds of the F3 progeny of self-pollinating plants were sterilized and then germinated on agar-solidified MS basal medium (Murashige and Skoog, 1962). The preparation and activation of LBA4404 harboring the PMDC84-(-tp)AaIPPI1 vector was performed as described previously (Xie *et al.*, 2004). In brief, leaves at the 5, 6, and 7 nodes of one-month-old seedlings were used as explants for infection by activated LBA4404. The infected leaf discs were inoculated on an agar-solidified selection medium, which was composed of MS basal medium supplemented with 1.0 mg/L N6-benzyladenine (BAP), 0.05 mg/L naphthalene-1-acetic acid (NAA), 20 mg/L hygromycin, and 200 mg/L timentin. Selected leaf discs were transferred to newly prepared selection medium every 10 days. After approximately 45 days of selection, multiple hygromycin-resistant adventitious shoots were

formed from calli induced from those infected leaf discs. Resistant shoots with 3-4 leaves were excised from the calli and inoculated on an optimized rooting medium (Alejos-Gonzalez *et al.*, 2013) consisting of half-strength MS basal medium supplemented with 0.5 mg/L NAA, 20 mg/L hygromycin, and 200 mg/L timentin. After 1-2 weeks of root induction, hygromycin-resistant plantlets were obtained and then planted in pot soil as described previously (Alejos-Gonzalez *et al.*, 2013). In addition, the same binary vector has been used to overexpress other artemisinin pathway genes (such as *AaHDR1*) in *A. annua* (Ma *et al.*, 2017). *AaHDR1* transgenic plants (Fig. S5a) were used as control to assess whether or not the used binary vector itself could affect the *AaIPPI1* expression.

Semi-quantitative RT–PCR, real time quantitative RT-PCR, and western blot analysis

Total RNA was isolated from young leaves of transgenic candidates and wild-type plants using an RNeasy Plant Mini Kit (Qiagen, CA, USA). All total RNA samples were treated with DNase with on-column DNase treatment (Qiagen, CA, USA) to remove genomic DNA as reported previously (Shi and Xie, 2010). The first-strand cDNA was synthesized using 1.0 μg of DNA-free RNA using PowerScript reverse transcriptase (RT, Clontech, USA). One microliter of cDNA was used as a template for PCR with Taq polymerase (Promega, Madison, USA). These three types of experiments were performed according to the protocols of the three manufacturers. RT-PCR was performed for the *(-tp)AaIPPI1* transgene and two positive controls, an HMG-CoA reductase (HMGR) gene and a farnesyl pyrophosphate synthase (FPS) gene. In addition, the housekeeping gene *β*-actin was used as a reference control. Gene-specific primers and PCR thermal cycles used for these genes are described in Table S1.

For the quantitative analysis, real-time qPCR was performed with Power SYBR Green PCR Master Mix (Applied Biosystems) according to the manufacturer's guidelines. The thermal cycles were composed of 50°C for 2 min and 95°C for 10 min, followed by 45 cycles of 95°C for 15 sec and 60°C for 1 min, and then 40°C. Five replicates and three replicates were evaluated for *AaIPPI11* and *ACTIN*, respectively. *ACTIN* was used as an internal standard for normalization. The relative expression of *AaIPPI1* was calculated using the ddCt algorithm.

Based on the AaIPPI1 amino acid sequence, an oligomeric peptide consisting of CKAPSDGKWGEHELD was synthesized by GenScript (Piscataway, NJ, USA). This oligomeric peptide, which was optimized to exclude the conserved domains of this enzyme family using the Antigen Design Tool, was used for the production of polyclonal antibodies. The resulting antibodies were used to specifically analyze the expression of the (-tp)AaIPPI1 protein in plants. Total crude protein extraction and immunoblot analyses were performed according to a previously reported method for deoxy-D-xylulose-5-phosphate synthase (Munoz-Bertomeu *et al.*, 2006). Briefly, total protein was extracted from the leaves of transgenic *A. annua* and wild-type plants. Crude protein samples were separated by SDS-PAGE, transferred to nitrocellulose membranes, and probed with rabbit anti-AaIPPI1 antibodies. An anti-rabbit IgG HPR conjugate was used as a secondary antibody (Promega, Madison, USA). A SuperSignal West Pico Trial Kit (Thermo Scientific, Rockford, IL, USA) was used to detect the immunoblotting signal on X-ray film according to the manufacturer's protocol.

Extraction of metabolites and GC-MS analysis

Based on our previous approaches to the metabolomic characterization of the landscape of isoprenoids in *A. annua* (Ma *et al.*, 2015), hexane was used to extract non-polar metabolites, and GC-MS analysis was performed using a gas chromatograph 6890 coupled with 5975C MSD (Agilent Technologies, USA) (Ma *et al.*, 2015). An RTX-5 capillary column (30 m × 0.25 mm × 0.25 μm) was used to separate metabolites. A splitless mode was used in the inlet. The injection temperature was set to 250°C. The oven temperature was initially set at 60°C and then ramped to 260°C at a constant rate of 10°C/min and held at 260°C for 25 min. Pure helium was used as the carrier gas at a flow rate of 1 mL/min. A positive electron impact ion source (70 EV) was used to ionize compounds, and mass fragments were scanned in the range from 40-800 (m/z) with a 4-min solvent delay. Ion chromatographic peaks detected by GC-MS were deconvoluted using Agilent MassHunter Mass Profiler (MHMP) and Mass Profiler Professional (MPP) software, as described previously (Ma *et al.*, 2015).

An authentic standard of arteannuin B (purchased from ALB Technology Limited, Hong Kong, category #: ALB-RS-0427) was used to establish a standard curve (Fig. S6) to quantify this metabolite in plants. In addition, authentic standards of artemisinic acid and artemisinin (purchased from Sigma, St. Louis, MO, USA) were used to establish standard curves (Fig. S6).

HPLC-MS and LC-Q-TOF-MS/MS analysis of artemisinin

Artemisinin extraction and quantification using HPLC-MS analysis on a 2010 eV LC/UV/ESI/MS instrument (Shimadzu) were performed as described previously (Alejos-Gonzalez *et al.*, 2011, Alejos-Gonzalez *et al.*, 2013). Artemisinin was further analyzed using LC-Q-TOF-MS/MS on an Agilent 6520 Q Time-of-Flight Tandem Mass Spectrometer coupled with an HPLC 1200 instrument (Agilent, CA, USA). The method used in this study has been reported recently (He *et al.*, 2016).

Analysis of isoprene using the fast isoprene sensor

The fast isoprene sensor (FIS) (Guenther and Hills, 1998, Exton *et al.*, 2010) has been developed to specifically measure isoprene emission in our laboratory (Xi *et al.*, 2016). In this study, wild-type, OE-T15, and OE-T17 plants (Fig. 4b) were used to measure isoprene. The measurements were conducted using the leaves of 7-week-old plants (Fig. 4b) according to the protocol developed by our group (Xi *et al*. 2016). Briefly, three leaves were excised from plants as a biological sample and placed in a sealed chamber connected to the FIS (Hills Scientific, Inc., Boulder, CO, USA). After the FIS was calibrated using an isoprene standard, a continuous recording was performed for 40 min to characterize the isoprene emissions. All measurements were performed between 9:00 AM and 3:00 PM at 25°C under a 16-h photoperiod supplemented with 50 µmol m^{-2} s^{-1} light intensity. To reduce the potential effects of measurement time on isoprene emission, the following steps were developed: the first measurement was performed using biological sample 1 in the order OE-T17, wild-type, and OE-T15; the second measurement was performed using biological sample 2 in the order OE-T15, OE-T17, and wild-type; the third measurement was performed using biological sample 3 in the order wild-type, OE-T15, and OE-T17. In FIS, isoprene is converted to

photons in the presence of an ozone flow. The emitted photons were recorded every 5 sec by a computer using FIS software. At the end of the experiment, the biological samples were carefully weighed to determine their fresh weights. The number of photons recorded over a 40-min period was averaged to estimate the mean value of isoprene emitted per gram of fresh tissue per second.

Statistical analysis

Student's t-test was used to evaluate the statistical significance of differences in plant growth and in the levels of various metabolites. P-values less than 0.05 were considered statistically significant.

Acknowledgements

This research was primarily supported by the North Carolina Biotechnology Center (grant #550031 and reference #2009-MRG-1117) and the CALS Dean's Enrichment Grants Program, College of Agriculture and Life Science, North Carolina State University. In addition, experiments for gene cloning and protein expression were financially supported by the Knowledge Innovation Program of the Chinese Academy of Sciences (grant #KSCX2-SW-329). The authors declare no conflicts of interest.

Short Supporting Information Legends

Fig. S1 Alignment of the deduced amino acids of *Artemisia annua* isopentenyl diphosphate isomerase 1 and six homologs using ClustalW (Thompson *et al*. 1994).

Fig. S2. Recombinant protein expression in *E. coli* for enzymatic assay and western blotting analysis of transgenic protein expression in plants.

Fig. S3. Ion chromatograph showing isomerization of IPP and DMAPP catalyzed by recombinant (-tp)AaIPPI1 and dephosphorylated by alkaline phosphatase or alkaline phosphatase and1 N hydrochloric acid (HCl).

Fig. S4. Overexpression of (-tp)AaIPPI1 targeted to the cytosol increases artemisinin content.

Fig. S5. Expression comparison of *AaIPPI1* in wild-type, *AaIPPPI1* overexpression, and control (*AaHDR1*) transgenic plants.

Fig. S6. Standard curves established using arteannuin B, artemisinic acid, and artemisinin for quantification of the three compounds from plants.

Fig. S7. Selected ion chromatograms of $[283]^+$ from the HPLC-MS analysis showing higher levels of artemisinin in the leaves of OE-T17 plants in comparison to wild-type plants.

Fig. S8. Tandem MS profiles of artemisinin. Artemisinin from transgenic plants, wild-type plants, and authentic standard samples were analyzed using tandem MS on a LC-Q-TOF-MS/MS instrument.

Table S1. Primer pairs designed for RT-PCR and thermal cycle programs for cDNA amplification.

Table S2. Sampling for analysis of arteannuin B, artemisinin, and artemisinic acid, RNA isolation, and protein extraction.

Supplementary Methods

Plant Growth in the phytotron and sampling.

Isolation of glandular trichomes.

References

Akhila, A., Thakur, R.S. and Popli, S.P. (1987) Biosynthesis of artemisinin in *Artemisia annua*. *Phytochemistry*, **26**, 1927-1930.

Albrecht, M. and Sandmann, G. (1994) Light-Stimulated Carotenoid Biosynthesis during Transformation of Maize Etioplasts Is Regulated by Increased Activity of Isopentenyl Pyrophosphate Isomerase. *Plant Physiol*, **105**, 529-534.

Alejos-Gonzalez, F., Perkins, K., Winston, M.I. and Xie, D.-Y. (2013) Efficient somatic embryogenesis and organogenesis of self-pollination *Artemisia annua* progeny and artemisinin formation in regenerated plants. *Am J Plant Sci*, **4**, 2206-2217

Alejos-Gonzalez, F., Qu, G.S., Zhou, L.L., Saravitz, C.H., Shurtleff, J.L. and Xie, D.Y. (2011) Characterization of development and artemisinin biosynthesis in self-pollinated *Artemisia annua* plants. *Planta*, **234**, 685-697.

Anderson, M.S., Muehlbacher, M., Street, I.P., Proffitt, J. and Poulter, C.D. (1989) Isopentenyl diphosphate: dimethylallyl diphosphate isomerase. *J Biol Chem*, **264**, 19169-19175.

Ansari, M.T., Saify, Z.S., Sultana, N., Ahmad, I., Saeed-Ul-Hassan, S., Tariq, I. and Khanum, M. (2013) Malaria and artemisinin derivatives: An updated review. *Mini-Rev Med Chem*, **13**, 1879-1902.

Berthelot, K., Estevez, Y., Deffieux, A. and Peruch, F. (2012) Isopentenyl diphosphate

isomerase: A checkpoint to isoprenoid biosynthesis. *Biochimie*, **94**, 1621-1634.

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Posthumus, M.A., Schmidt, C.O., De Kraker, J.-W., Konig, W.A. and Franssen, M.C.R. (1999) Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry*, **52**, 843-854.

Brandt, W., Dessoy, M.A., Fulhorst, M., Gao, W., Zenk, M.H. and Wessjohann, L.A. (2004) A proposed mechanism for the reductive ring opening of the cyclodiphosphate MEcPP, a crucial transformation in the new DXP/MEP pathway to isoprenoids based on modeling studies and feeding experiments. *Chembiochem*, **5**, 311-323.

Brown, G.D. (2010) The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of *Artemisia annua* L. (Qinghao). *Molecules*, **15**, 7603-7698.

Carrigan, C.N. and Poulter, C.D. (2003) Zinc is an essential cofactor for type I isopentenyl diphosphate:dimethylallyl diphosphate isomerase. *J Am Chem Soc*, **125**, 9008-9009.

Chang, Y.-J., Song, S.-H., Park, S.-H. and Kim, S.-U. (2000) Amorpha-4,11-diene synthase of *Artemisia annua*: cDNA isolation and bacterial expression of a terpene synthase Involved in artemisinin biosynthesis. *Arch Biochem Biophys*, **383**, 178-184.

Chen, D.-H., Ye, H.-C. and Li, G.-F. (2000) Expression of a chimeric farnesyl diphosphate synthase gene in *Artemisia annua L*. transgenic plants via *Agrobacterium tumefaciens*-mediated transformation. *Plant Sci*, **155**, 179-185.

Covello, P.S., Teoh, K.H., Polichuk, D.R., Reed, D.W. and Nowak, G. (2007) Functional genomics and the biosynthesis of artemisinin. *Phytochemistry*, **68**, 1864-1971.

Croteau, R., Kutchan, T. and Lewis, N. (2000) Natural Products (Secondary Metabolites). In *Biochemistry & Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L. eds), pp. 1250-1342.

Curtis, M.D. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol*, **133**, 462-469.

Delabays, N., Benakis, A. and Collet, G. (1993) Selection and breeding for high artemisinin (qinghaosu) yielding strains of *Artemisia annua*. *Acta Horticult*, **330**, 203-207.

Delabays, N., Simonnet, X. and Gaudin, M. (2001) The genetics of artemisinin content in *Artemisia annua* L. and the breeding of high yielding cultivars. *Curr Med Chem*, **8**, 1795-1801.

Dhingra, V. and Lakshmi Narasu, M. (2001) Purification and characterization of an enzyme Involved in biochemical transformation of arteannuin B to artemisinin from *Artemisia annua*. *Biochem Biophy Res Commun*, **281**, 558-561.

Duke, M.V., Paul, R.N., Elsohly, H.N., Sturtz, G. and Duke, S.O. (1994) Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua L*. *Int J Plant Sci*, **155**, 365-372.

Duke, S.O. and Paul, R.N. (1993) Development and fine structure of the glandular trichomes of *Artemisia annua* L. *Int J Plant Sci*, **154**, 107-118.

Eisenreich, W., Bacher, A., Arigoni, D. and Rohdich, F. (2004) Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell Mol Life Sci*, **61**, 1401-1426.

Emanuelsson, O., Brunak, S., von Heijne, G. and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*, **2**, 953-971.

Exton, D.A., Smith, D.J., McGenity, T.J., Steinke, M., Hills, A.J. and Suggett, D.J. (2010)

Application of a fast Isoprene sensor (FIS) for measuring isoprene production from marine samples. *Limnol Oceanogr Meth*, **8**, 185-195.

Gallagher, C.E., Cervantes-Cervantes, M. and Wurtzel, E.T. (2003) Surrogate biochemistry: use of Escherichia coli to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation. *Appl Microbiol Biotechnol*, **60**, 713-719.

Graham, I.A., Besser, K., Blumer, S., Branigan, C.A., Czechowski, T., Elias, L., Guterman, I., Harvey, D., Isaac, P.G., Khan, A.M., Larson, T.R., Li, Y., Pawson, T., Penfield, T., Rae, A.M., Rathbone, D.A., Reid, S., Ross, J., Smallwood, M.F., Segura, V., Townsend, T., Vyas, D., Winzer, T. and Bowles, D. (2010) The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science*, **327**, 328-331.

Guenther, A.B. and Hills, A.J. (1998) Eddy covariance measurement of isoprene fluxes. *J Geophys Res-Atmos*, **103**, 13145-13152.

Hahn, F.M., Hurlburt, A.P. and Poulter, C.D. (1999) *Escherichia coli* open reading frame 696 Is idi, a nonessential gene encoding isopentenyl diphosphate isomerase. *J Bacteriol*, **181**, 4499-4504.

Han, J.L., Liu, B.Y., Ye, H.C., Wang, H., Li, Z.Q. and Li, G.F. (2006) Effects of overexpression of the endogenous farnesyl diphosphate synthase on the artemisinin content in *Artemisia annua* L. *J Integr Plant Biol*, **48**, 482-487.

He, X., Li, Y., Lawson, D. and Xie, D.-Y. (2016) Metabolic engineering of anthocyanins in dark tobacco varieties *PHYSIOLOGIA PLANTARUM*, **doi:10.1111/ppl.12475**.

Heuston, S., Begley, M., Gahan, C.G.M. and Hill, C. (2012) Isoprenoid biosynthesis in bacterial pathogens. *Microbiology-(UK)*, **158**, 1389-1401.

Hoeffler, J.F., Hemmerlin, A., Grosdemange-Billiard, C., Bach, T.J. and Rohmer, M. (2002) Isoprenoid biosynthesis in higher plants and in Escherichia coli: on the branching in the methylerythritol phosphate pathway and the independent biosynthesis of isopentenyl diphosphate and dimethylallyl diphosphate. *Biochem J*, **366**, 573-583.

Hoshino, T., Tamegai, H., Kakinuma, K. and Eguchi, T. (2006) Inhibition of type 2 isopentenyl diphosphate isomerase from Methanocaldococcus jannaschii by a mechanism-based inhibitor of type 1 isopentenyl diphosphate isomerase. *Bioorg Med Chem*, **14**, 6555-6559.

Kajiwara, S., Fraser, P.D., Kondo, K. and Misawa, N. (1997) Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in *Escherichia coli*. *Biochem J*, **324**, 421-426.

Kaneda, K., Kuzuyama, T., Takagi, M., Hayakawa, Y. and Seto, H. (2001) An unusual isopentenyl diphosphate isomerase found in the mevalonate pathway gene cluster from Streptomyces sp. strain CL190. *Proc Natl Acad Sci U S A*, **98**, 932-937.

Klayman, D.L. (1985) Qinghaosu (Artemisinin) - An antimalarial drug from China. *Science*, **228**, 1049-1055.

Lansbury, P.T. and Mojica, C.A. (1986) Total synthesis of (+/-)-arteannuin B. *Tetrahedron Lett*, **27**, 3967-3970.

Lee, S. and Poulter, C.D. (2006) Escherichia coli type I isopentenyl diphosphate isomerase: structural and catalytic roles for divalent metals. *J Am Chem Soc*, **128**, 11545-11550.

Liu, J.-M., Ni, M.-Y., Fan, J.-F., Tu, Y.-Y., Wu, Z.-H., Wu, Y.-L. and Chou, W.-S. (1979) Structure and reaction of arteannuin. *Acta Chim Sin*, **37**, 129-143.

Lommen, W.J., Schenk, E., Bouwmeester, H.J. and Verstappen, F.W. (2005) Trichome dynamics and artemisinin accumulation during development and senescence of *Artemisia annua* leaves. *Planta Med*, **71**, 336.

Ma, D.-M., Wang, Z., Wang, L., Alejos-Gonzales, F., Sun, M.-A. and Xie, D.-Y. (2015) A genome-wide scenario of terpene pathways in self-pollinated *Artemisia annua*. *Mol. Plant.*, **8**, 1580-1598.

Ma, D.M., Li, G., Zhu, Y. and Xie, D.Y. (2017) Overexpression and suppression of *Artemisia annua* 4-hydroxy-3-methylbut-2-enyl diphosphate reductase 1 gene (AaHDR1) differentially regulate artemisinin and terpenoid biosynthesis. *Front Plant Sci*, **8**.

Ma, D.M., Pu, G.B., Lei, C.Y., Ma, L.Q., Wang, H.H., Guo, Y.W., Chen, J.L., Du, Z.G., Wang, H., Li, G.F., Ye, H.C. and Liu, B.Y. (2009) 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*, **50**, 2146-2161.

Martin, K., Kopperud, K., Chakrabarty, R., Banerjee, R., Brooks, R. and Goodin, M.M. (2009) Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. *Plant J*, **59**, 150-162.

Miller, Louis H. and Su, X. (2011) Artemisinin: Discovery from the Chinese Herbal Garden. *Cell*, **146**, 855-858.

Muehlbacher, M. and Poulter, C.D. (1988) Isopentenyl diphosphate isomerase - inactivation of the enzyme with active-site-directed irresversible inhibitors and transition-state analogs. *Biochemistry*, **27**, 7315-7328.

Munoz-Bertomeu, J., Arrillaga, I., Ros, R. and Segura, J. (2006) Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant Physiol*, **142**, 890-900.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. *PHYSIOLOGIA PLANTARUM*, **15**, 473-497.

Nair, M.S.R. and Basile, D.V. (1993) Bioconversion of arteannuin B to artemisinin. *Journal of Natural Products*, **56**, 1559 - 1566.

Nowak, D.M. and Lansbury, P.T. (1998) Synthesis of (+)-artemisinin and (+)-deoxoartemisinin from arteannuin B and arteannuic acid. *Tetrahedron*, **54**, 319-336.

Oh, S.K., Kang, H. and Shin, D.H. (2000) Molecular cloning and characterization of a functional cDNA clone encoding isopentenyl diphosphate isomerase from Hevea brasiliensi. *J. Plant Physiol*, **157**.

Okada, K., Kasahara, H., Yamaguchi, S., Kawaide, H., Kamiya, Y., Nojiri, H. and Yamane, H. (2008) Genetic evidence for the role of isopentenyl diphosphate isomerases in the mevalonate pathway and plant development in *arabidopsis*. *Plant Cell Physiol*, **49**, 604-616.

Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell,

M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S. and Newman, J.D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, **496**, 528-+.

Paul, S., Khanuja, S.P.S. and Gupta, M.M. (2014) Breeding strategy for genetic improvement up to four generations in relation to artemisinin with canopy and other secondary metabolites in *Artemisia annua* L. *Ind Crop Prod*, **56**, 67-73.

Peplow, M. (2016) Synthetic malaria drug meets market resistance. *Nature*, **530**, 389-390.

Phillips, M.A., D'Auria, J.C., Gershenzon, J. and Pichersky, E. (2008) The *Arabidopsis thaliana* type I isopentenyl diphosphate isomerases are targeted to multiple subcellular compartments and have overlapping functions in isoprenoid biosynthesis. *Plant Cell*, **20**, 677-696.

Ramos-Valdivia, A.C., van der Heijden, R. and Verpoorte, R. (1997) Isopentenyl diphosphate isomerase: a core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. *Nat Prod Rep*, **14**, 591-603.

RamosValdivia, A.C., vanderHeijden, R. and Verpoorte, R. (1997) Isopentenyl diphosphate isomerase: a core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. *Nat Prod Rep*, **14**, 591-603.

Ro, D.-K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R. and Keasling, J.D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature*, **440**, 940-943.

Rodriguez-Concepcion, M. and Boronat, A. (2015) Breaking new ground in the regulation of the early steps of plant isoprenoid biosynthesis. *Curr Opin Plant Biol*, **25**, 17-22.

Rohdich, F., Zepeck, F., Adam, P., Hecht, S., Kaiser, J., Laupitz, R., Grawert, T., Amslinger, S., Eisenreich, W., Bacher, A. and Arigoni, D. (2003) The deoxyxylulose phosphate pathway of isoprenoid biosynthesis: studies on the mechanisms of the reactions catalyzed by IspG and IspH protein. *Proc Natl Acad Sci U S A*, **100**, 1586-1591.

Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D. and Eyal, Y. (2008) Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant Physiol*, **148**, 1219-1228.

Sharma, N.K., Pan, J.J. and Poulter, C.D. (2010) Type II isopentenyl diphosphate isomerase: probing the mechanism with alkyne/allene diphosphate substrate analogues. *Biochemistry*, **49**, 6228-6233.

Shi, M.-Z. and Xie, D.-Y. (2010) Features of anthocyanin biosynthesis in *pap1-D* and wild-type *Arabidopsis* thaliana plants grown in different light intensity and culture media conditions. *Planta*, **231**, 1385-1400.

Shretta, R. and Yadav, P. (2012) Stabilizing supply of artemisinin and artemisinin-based

combination therapy in an era of wide-spread scale-up. *Malaria J*, **11**, 399.

Singh, N.D., Kumar, S. and Daniell, H. (2016) Expression of beta-glucosidase increases trichome density and artemisinin content in transgenic *Artemisia annua* plants. *Plant Biotechnol J*, **14**, 1034-1045.

Taylor, S. (2013) Artemisia annua, Artemisinin, ACTs and Malaria Control in Africa: Tradition, Science and Public Policy. *Soc Hist Med*, **26**, 799-801.

Tellez, M.R., Canel, C., Rimando, A.M. and Duke, S.O. (1999) Differential accumulation of isoprenoids in glanded and glandless *Artemisia annua L*. *Phytochemistry*, **52**, 1035-1040.

Teoh, K.H., Polichuk, D.R., Reed, D.W. and Covello, P.S. (2009) Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. *Botany*, **87**, 635-642.

Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G. and Covello, P.S. (2006) *Artemisia annua L*. (*Asteraceae*) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett*, **580**, 1411-1416.

Ting, H.-M., Wang, B., Ryden, A.-M., Woittiez, L., van Herpen, T., Verstappen, F.W.A., Ruyter-Spira, C., Beekwilder, J., Bouwmeester, H.J. and van der Krol, A. (2013) The metabolite chemotype of Nicotiana benthamiana transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytol.*, **199**, 352-366.

Turconi, J., Griolet, F., Guevel, R., Oddon, G., Villa, R., Geatti, A., Hvala, M., Rossen, K., Goller, R. and Burgard, A. (2014) Semisynthetic artemisinin, the chemical path to industrial production. *Org Process Res Dev*, **18**, 417-422.

Vranova, E., Coman, D. and Gruissem, W. (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu Rev Plant Biol*, **64**, 665-700.

Weathers, P., Elkholy, S. and Wobbe, K. (2006) Artemisinin: The biosynthetic pathway and its regulation in *Artemisia annua*, a terpenoid-rich species. *In Vitro Cell Dev Biol - Plant*, **42**, 309.

Weathers, P.J., Elfawal, M.A., Towler, M.J., Acquaah-Mensah, G.K. and Rich, S.M. (2014a) Pharmacokinetics of artemisinin delivered by oral consumption of *Artemisia annua* dried leaves in healthy vs. *Plasmodium chabaudi*-infected mice. *J Ethnopharmacol*, **153**, 732-736.

Weathers, P.J., Jordan, N.J., Lasin, P. and Towler, M.J. (2014b) Simulated digestion of dried leaves of *Artemisia annua* consumed as a treatment (pACT) for malaria. *J Ethnopharmacol*, **151**, 858-863.

Weathers, P.J. and Towler, M.J. (2014) Changes in key constituents of clonally propagated *Artemisia annua* L. during preparation of compressed leaf tablets for possible therapeutic use. *Ind Crop Prod*, **62**, 173-178.

White, N.J. (2008) Qinghaosu (Artemisinin): The price of success. *Science*, **320**, 330-334.

WHO (2006) Meeting on the production of artemisinin and artemisinin-based combination therapies 6-7 June 2005 Arusha, United Republic Tanzania

WHO (2010) Good procurement practices for artemisinin-based antimalarial medicines. In *WHO Global Malaria Programme*.

Woerdenbag, H., Pras, N., Chan, N.G., Bang, B.T., Bos, R., van Uden, W., Van Y, P., Van Boi, N., Batterman, S. and Lugt, C. (1994) Artemisinin, related sesquiterpenes, and essential oil in *Artemisia annua* during a vegetation period in Vietnam. *Planta Med*, **60**, 272-274.

Woerdenbag, H.J., Bos, N.P.R., Hendriks, J.F.V.H. and Malingré, T.M. (1991) Analysis of artemisinin and related sesquiterpenoids from *Artemisia annua* L. by combined gas chromatography/mass spectrometry. *Phytochem Analysis*, **2**, 215-219.

Xi, J., Rossi, L., Lin, X. and Xie, D.-Y. (2016) Overexpression of a synthetic insect–plant geranyl pyrophosphate synthase gene in Camelina sativa alters plant growth and terpene biosynthesis. *Planta*, **in press**, 1-16.

Xiao, L., Tan, H.X. and Zhang, L. (2016) Artemisia annua glandular secretory trichomes: the biofactory of antimalarial agent artemisinin. *Sci bull*, **61**, 26-36.

Xie, D.-Y., Jackson, L.A., Cooper, J.D., Ferreira, D. and Paiva, N.L. (2004) Molecular and biochemical analysis of two cDNA clones encoding dihydroflavonol-4-reductase from *Medicago truncatula*. *Plant Physiol.*, **134**, 979-994.

Xie, D.-Y., Kang, N.-l. and Li, G.-z. (1995) Studies on the karyotype of *Artemisia annua*. *Chinese Bulletin of Botany*, **12 (Supplement)**, 71-72.

Zhang, Y., Teoh, K.H., Reed, D.W. and Covello, P.S. (2009) Molecular cloning and characterization of Dbr1, a 2-alkenal reductase from *Artemisia annua*. *Botany*, **87**, 643-649.

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J.H., Ross, A.R.S. and Covello, P.S. (2008) The molecular cloning of artemisinic aldehyde Δ-11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J Biol Chem*, **283**, 21501-21508.

Zhou, C.F., Li, Z.R., Wiberley-Bradford, A.E., Weise, S.E. and Sharkey, T.D. (2013) Isopentenyl diphosphate and dimethylallyl diphosphate/isopentenyl diphosphate ratio measured with recombinant isopentenyl diphosphate isomerase and isoprene synthase. *Anal Biochem*, **440**, 130-136.

Figure Legends

Fig. 1. Biosynthetic pathway of artemisinin and the metabolic network characterized in *Artemisia annua*. (**a**) Schematic showing the pathways to artemisinin and other compounds starting from the mevalonate and 2-C-methyl-D-erythritol 4-phosphate (MEP) pathways in *A. annua*. Metabolite abbreviations: G3P, glycerol-3-phosphate; HMBPP, 4-hydroxy-3-methyl-but-2-enyl pyrophosphate; MVAPP: mevalonate pyrophosphate; IPP: isopentyl pyrophosphate; DMAPP: dimethylallyl pyrophosphate; GPP: geranyl pyrophosphate; FPP: farnesyl pyrophosphate; and GGPP: geranylgeranyl pyrophosphate. Enzyme abbreviations: HDR, 4-hydroxy-3-methyl-but-2-enyl pyrophosphate reductase (the wide arrow extending from HMBPP to IPP indicates HDR activity that is preferential for IPP); MPDC, mevalonate 5-pyrophosphate decarboxylase; IPPI: isopentenyl diphosphate isomerase; ADS, armorpha-4,11-diene synthase, ADH1: alcohol dehydrogenase 1, ALDH1: aldehyde dehydrogenase, CPR1: cytochrome P450 reductase 1, CYB5: cytochrome b5 monooxygenase, CYP71AV1: cytochrome P450 monooxygenase; DBR2: artemisinic aldehyde delta-11(13)-double bond reductase. Question marks (?) indicate unknown. (**b**) A gene-metabolite network established from six tissues showing a metabolic association between *AaIPPI1* and four artemisinin pathway genes (*ADS*, *CYP71AV1*, *CPR1*, and *DBR2*) and artemisinic acid (aa) and arteannuin B (a-b) (Ma *et al.* 2015).

Fig. 2. Expression profiles of *AaIPPI1*in different tissues. Semi-quantitative RT-PCR (**a**) and qRT-PCR analyses (**b**) showing *AaIPPI1* expression levels in various organs of *A. annua*. Bars labeled with different numbers of "*" means significant difference from the bar of roots (Student's *t*-test, "*": P-value less than 0.05, "**": P-value less than 0.01).

Fig. 3. Characterization of products from the enzymatic reaction of recombinant (-tp)AaIPPI1 with IPP and DMAPP for 3 h. Products were analyzed after hydrolysis using alkaline phosphatase. (**a)** Gas chromatogram showing two reaction products obtained from the incubation of recombinant (-tp)AaIPPI1 and IPP substrate. (**b)** Gas chromatogram showing two reaction products obtained from the incubation of recombinant (-tp)AaIPPI1 and the

DMAPP substrate. (**c)** Gas chromatogram showing 3-methyl-3-buten-1-ol (isoprenol), the only product from the incubation of denatured recombinant (-tp)AaIPPI1 and the IPP substrate as a control. (**d)** Gas chromatogram showing 3-methyl-2-buten-1-ol, the only product from the incubation of denatured recombinant (-tp)AaIPPI1 and the DMAPP substrate as a control. (**e)** Mass spectrum for peak #1 in (**a**), (**b**), and (**c**). (**f**) Mass spectrum for peak 2 labeled in (**a)**, (**b)**, and (**d)**. (**g**) Mass spectrum for the 3-methyl-3-buten-1-ol standard in the NIST library. (**h)** Mass spectrum for the 3-methyl-2-buten-1-ol standard in the NIST library.

Fig. 4. Overexpression of (-tp)AaIPPI1 is targeted to the cytosol. (**a)** The cassette of the *(-tp)AaIPPI1*/*EGFP* fusion construct was used to examine the protein localization, and confocal images showed the subcellular localization of (-tp)AaIPPI1-EGFP in the cytosol of the leaves of *N. benthamiana*. (**b)** T-DNA cassette of a binary construct in which a 2 x 35S promoter was used to overexpress *(-tp)AaIPPI1* in *A. annua*; seven-week-old plants of three transgenic lines (after transfer from selection medium to soil) and the wild-type control were photographed to show their similar growth phenotype. (**c)** Semi-quantitative RT-PCR images (left) and qRT-PCR plot (right) showing overexpression of the *(-tp)AaIPPI1* transgene in the transgenic plants. *HMGR* and *FPS*, two other genes in the pathway used as a reference control in the semi-quantitative RT-PCR, showed equivalent expression in transgenic and wild-type control plants. *ACTIN* was used as a reference gene to assess the RT-PCR efficiency in the transgenic and wild-type plants. Bars labeled with "*" indicate a significant difference (Student's t-test, P-value less than 0.01) between transgenic and wild-type plants.

Fig. 5. Ion chromatographs and mass spectra of artemisinic acid and arteannuin B in transgenic and wild-type plants. (**a-b**) Ion chromatographs of GC-MS comparing the peak size of artemisinic acid from transgenic and wild-type plants (**a**) and MS profiles of artemisinic acid from transgenic plants and an authentic standard; (**c-d**) ion chromatographs of GC-MS comparing the peak sizes of arteannuin B from transgenic and wild-type plants (**c**) and MS profiles of arteannuin B from transgenic plants and an authentic standard (**d**). OE-T1, OE-T15, and OE-T17 transgenic lines; WT: wild-type.

Fig. 6. Contents of artemisinin, artemisinic acid, and arteannuin B and isoprene emission. (**a**)-(**b**) Contents of artemisinic acid and arteannuin B **(a)** and artemisinin (**b**) in the leaves of OE-T1, OE-T15, OE-T17, and wild-type control seedlings. (**c**)-(**d**) Isoprene emission dynamics recorded over a 40-min period (**c**) and average levels per second (**d**) indicated by photon accounts generated by a fast isoprene sensor. OE-T1, T15, and T17: three transgenic lines overexpressing truncated *(-tp)AaIPPI1* (lacking the plastid transit peptide); WT: wild-type. Student's *t*-test was performed to calculate P-values less than 0.05 or 0.01. In "**a**", bars labeled with "*" and "**" mean P-values less than 0.05 and 0.01, respectively; bars labelled with "+ and #" mean that the values of artemisinic acid were compared among transgenic and WT plants, but were not compared with those of arteannuin B in plants. Bars with "+" means no significant difference between transgenic and WT plants, while bar with "#" means significant difference between transgenic and WT plants. In "**b**", bars labelled with "*" mean P-values less than 0.01. In "**d**", bars labelled with "*" and "**" mean P-values less than 0.05 and 0.01, respectively, indicating significant difference between transgenic and WT plants.