CrossMark

ORIGINAL ARTICLE

Characterization of a trichome-specific promoter of the aldehyde dehydrogenase 1 (*ALDH1*) gene in *Artemisia annua*

Meng Liu[1] · Pu Shi[1] · Xueqing Fu[1] · Peter E. Brodelius[2] · Qian Shen[1] ·
Weimin Jiang[1] · Qian He[1] · Kexuan Tang[1]

Abstract Artemisinin is a frequently used anti-malaria drug extracted from glandular trichomes (GSTs) in *Artemisia annua* L. In this study, we report on the characterization of the promoter of aldehyde dehydrogenase 1 (*ALDH1*) involved in the biosynthesis of artemisinin. A 1620-bp promoter fragment was cloned upstream of the *ALDH1* start codon. Putative regulatory *cis*-acting elements are predicted by software, revealing that this gene is affected by complex factors. The activity of the *ALDH1* promoter was analyzed using a reporter gene *GUS*. *GUS* expression showed a spatial difference in leaves at different ages. In young leaves, GUS staining was exclusively discovered in GSTs. In older leaves, both GSTs and T-shaped trichomes (TSTs) showed GUS signals. Only TSTs showed GUS staining in lower leaves. No GUS staining was detected in the bottom leaves. The result demonstrates that the *ALDH1* promoter is trichome-specific. The RT-Q-PCR analysis revealed that both wild-type and recombinant promoters showed similar activity in *A. annua*. After application of exogenous 100 μM methyl jasmonate, 100 μM gibberellin and 100 μM salicylic acid separately, the transcript levels were increased significantly, indicating that *ALDH1* may play an important role in the response to hormones in *A. annua*.

Introduction

Malaria is an infectious disease caused by parasites with high incidence rate and fatality rate. According to World Malaria Report 2014 (WHO 2015), about 3.3 billion people are facing the risk of malaria infection, and 1.2 billion are at high risk. Artemisinin is an endoperoxide sesquiterpene lactone extracted from *Artemisia annua* L. World Health Organization has recommended the artemisinin based combination therapies (ACTs) as the most effective method for the treatment of malaria. The artemisinin content in wild-type *A. annua* is low (0.1–0.8 % dry weight) and cannot meet the large demand around the world (Zhu et al. 2014). Hence, it is of significant importance to find effective strategies to enhance artemisinin content in *A. annua*. Metabolic engineering is a frequently used way to cultivate high artemisinin content plants. According to the recent reports, overexpression of the enzyme genes in artemisinin biosynthesis can enhance artemisinin content remarkably (Waleerat et al. 2010; Lu et al. 2013).

As shown in Fig. 1, all the enzyme genes involved in the biosynthetic pathway of artemisinin in *A. annua* have been discovered and characterized. Firstly, the formation of amorpha-4,11-diene from farnesyl diphosphate (FDP) under the catalysis of amorpha-4,11-diene synthase (ADS) is a critical step of artemisinin biosynthesis (Bouwmeester et al. 1999; Mercke et al. 2000). Next, amorpha-4,11-diene is hydroxylated to artemisinic alcohol with the action of cytochrome P450-dependent hydroxylase (CYP71AV1) and NADPH:cytochrome P450 oxidoreductase (CPR) (Ro

✉ Kexuan Tang
kxtang@sjtu.edu.cn

1 Key Laboratory of Urban Agriculture (South), Ministry of Agriculture, Plant Biotechnology Research Center, School of Agriculture and Biology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

2 Department of Chemistry and Biomedical Sciences, Linnaeus University, 39182 Kalmar, Sweden

Fig. 1 Key enzymes involved in artemisinin biosynthesis pathway in *A. annua*. ADS, amorpha-4,11-diene synthase; CYP71AV1, cytochrome P450 dependent hydroxylase; CPR, cytochrome P450 oxidoreductase; DBR2, double bond reductase 2; ALDH1, aldehyde dehydrogenase
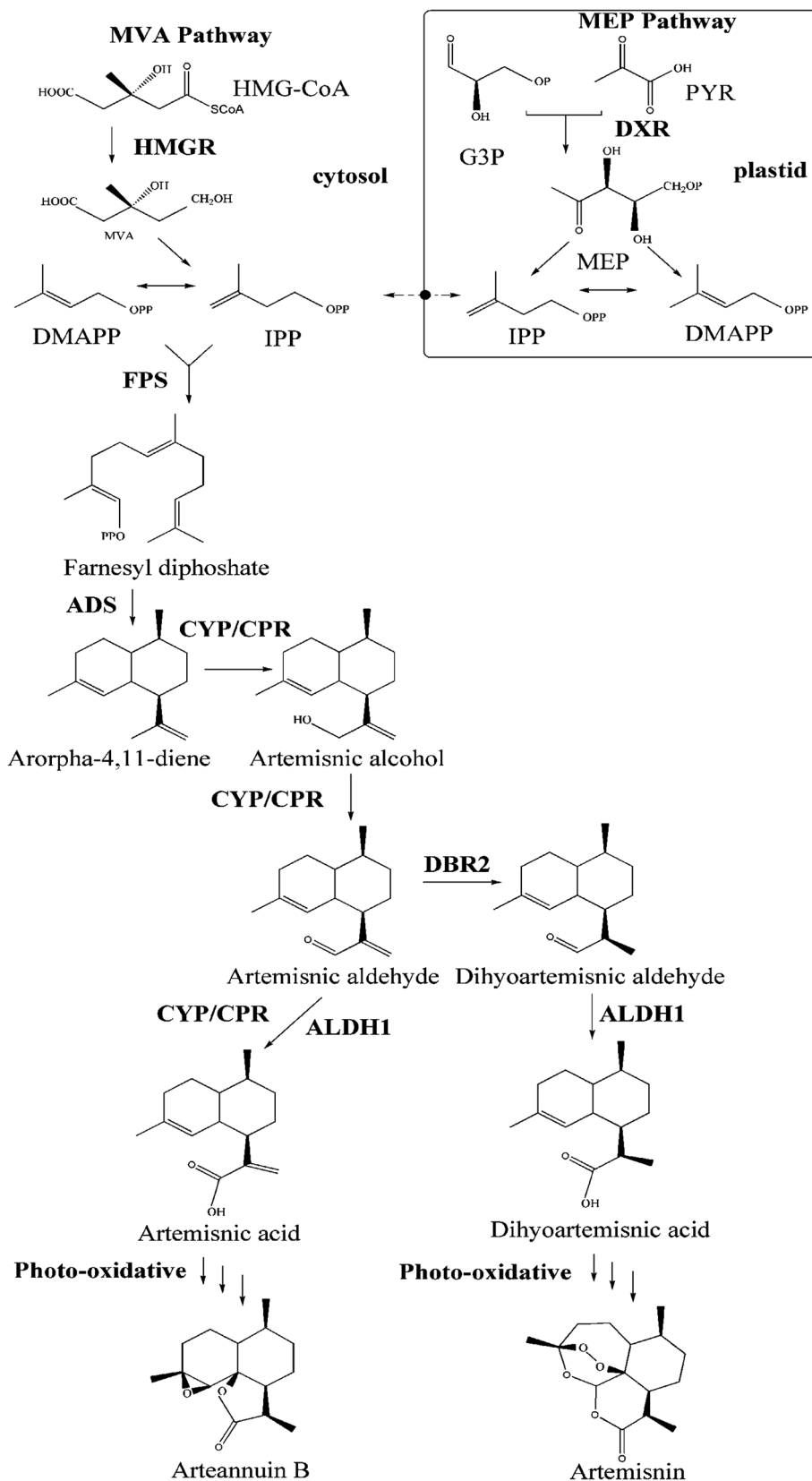
et al. 2006; Teoh et al. 2006). Both CYP71AV1 and CPR participate in the oxidation reaction from artemisinic alcohol to artemisinic aldehyde and further generate artemisinic acid (Lu et al. 2013). Alternatively, artemisinic aldehyde Δ11-(13) reductase (DBR2) can catalyze the reaction from artemisinic aldehyde to dihydroartemisinic aldehyde (Zhang et al. 2008). Aldehyde dehydrogenase 1 (ALDH1) is the last key enzyme which catalyzes the conversion of dihydroartemisinic aldehyde to dihydroartemisinic acid (Teoh et al. 2009). Finally, all the conversions from dihydroartemisinic acid to artemisinin and artemisinic acid to arteannuin B are considered as non-enzymatic reactions (Sy and Brown 2002; Brown and Sy 2004, 2007).

In *A. annua*, there are two types of trichomes, i.e. glandular secretory trichomes (GSTs) and T-shaped trichomes (TSTs). Previous studies have demonstrated that trichomes play an importance role in the formation and storage of artemisinin, especially GSTs (Ferreira and Janick 1996; Kapoor et al. 2007; Arsenault et al. 2010; Nguyen et al. 2011). Researchers have previously revealed that the promoters of enzymatic genes such as *ADS*, *CYP71AV1* and *DBR2* in artemisinin biosynthesis pathway are specifically active in GSTs of young leaves (Wang et al. 2011, 2013; Jiang et al. 2014). The cauliflower mosaic virus 35S promoter (pCaMV-35S) is widely used in metabolic engineering, but it may have adverse effects on plants since foreign proteins are expressed in most tissues of transgenic plants (Bakhsh et al. 2011). Therefore, it is attractive to enhance the artemisinin content by increasing the density of GSTs or to use trichome-specific promoters to control the genes expressing in them (Zhu et al. 2014). However, to our knowledge, there is no report on characterization of the promoter of *ALDH1* in *A. annua*.

To better understand the expression pattern of *ALDH1*, a 1620-bp promoter fragment upstream of the *ALDH1* start codon was cloned from *A. annua*. The promoter was fused with the reporter gene *GUS* and transformed into *A. annua*. Finally, we determined the transcript levels of *ALDH1* in different tissues and after hormonal treatments by RT-Q-PCR.

Materials and methods

Plant material and bacterial strain

Seeds of *A. annua* for this experiment were collected from Chongqing, PR China. They were germinated and grown on Murashige and Skoog (MS) medium with 2 % sucrose and 0.7 % agar and then transplanted into soil after regeneration (Murashige and Skoog 1962). The plants of *A. annua* were grown under the condition of 16 h light/8 h

dark at a constant temperature of 24 °C. *Escherichia coli* (strain DH5α) was used as an intermediate vector for the promoter cloning. *Agrobacterium tumefaciens* (strain EHA105) was used for plant transformation.

Promotor cloning

All the primers used for the cloning of *ALDH1* promoter are listed in Table 1. Genomic DNA was extracted from fresh young leaves of *A. annua* using the cetyltrimethylammonium bromide method (Fütterer et al. 1995). Based on the *ALDH1* gene promoter sequence (GenBank No.: KC118522.1), the promoter fragment of *ALDH1* was amplified by KOD polymerase (Toyobo, Japan), and conducted as follows: 95 °C for 5 min; 35 cycles at 95 °C for 40 s, 55 °C for 40 s, 68 °C for 2 min; 68 °C for 10 min. The PCR products were ligated into the pJET2.1 vector (Thermo Fisher Scientific, Waltham, MA, USA) for sequencing. The nucleotide sequence was used to design primers for vector construction.

Construction of transformation vector

In order to insert the promoter into the plant transformation vector, specific primers were designed to include digestion sites of two restriction enzymes *Hind*III (at the 5′ end) and *Bam*HI (at the 3′ end). We used pCAMBIA1391Z (CambiaLabs) plasmid with the reporter gene *GUS* as a stable transformation vector. The target promoter fragment and the resulting linear vector pCAMBIA 1391Z digested with restriction enzymes *Bam*HI and *Hind*III (Fermentas FastDigest) were ligated at 22 °C using T4 DNA ligase (Thermo Scientific). The reaction system is as follows: 6 μL target DNA fragment, 2 μL linear plasmid, 1 μL T4 DNA ligase, 1 μL 10 × T4 DNA ligase buffer, and conducted for 30 min. The obtained plant transformation vector pCAMBIA1391Z::*pALDH1*::*GUS* as well as an empty pCAMBIA1391Z were introduced into *A. tumefaciens* EHA105 by the freeze–thaw method.

Plant transformation

The two vectors pCAMBIA1391Z::*pALDH1*::*GUS* and pCAMBIA1391Z were introduced into *A. annua* by *A. tumefaciens*-mediated leaf-disc transformation method described in a previous study (Lu et al. 2013). First of all, seeds of *A. annua* were surface sterilized in 70 % ethanol for 1 min, then treated with 20 % (v/v) NaOCl (sodium hypochlorite) for 10 min with a following step of washing 4–5 times using MilliQ water. Subsequently, the seeds were placed on Murashige and Skoog basal medium (Sigma-Aldrich, MO, USA) (Murashige and Skoog 1962) and grown in a growth chamber at 25 °C with a

Table 1 Nucleotide sequence of primes used in this study

No	Name	Application	Sequence (5′–3′)
1	ALDH1-1	Cloning	GCCTTAACGGCCAAATCAATGTCTT
2	ALDH1-2	Cloning	GTTGCTAACACTTCTTCTGTCGCTGGAT
3	ALDH1-HindIII-FP	Recombinant	CCCAAGCTTATGAACCATTAGAAG
4	ALDH1-BamHI-RP	Recombinant	CGCGGATCCCTTTGTTTTTT
5	ALDH1-RT-F	qPCR	GGACTTGCCTCAGGTGTAT
6	ALDH1-RT-R	qPCR	GTGCCTCTAATCCTTGTTC
7	ADS-RT-F	qPCR	AATGGGCAAATGAGGGACAC
8	ADS-RT-R	qPCR	TTTCAAGGCTCGATGAACTATG
9	CYP-RT-F	qPCR	CACCCTCCACTACCCTTG
10	CYP-RT-R	qPCR	GACACATCCTTCTCCCAGC
11	DBR2-RT-F	qPCR	CTTGGGTTACAAGCTGTGGCTCAAG
12	DBR2-RT-R	qPCR	ATATAATCAAAACTAGAGGAGTGACC
13	β-actin-F	qPCR	CCAGGCTGTTCAGTCTCTGTAT
14	β-actin-R	qPCR	CGCTCGGTAAGGATCTTCATCA
15	GUS-RT-F	qPCR	AACCGTTCTACTTTACTGGCTTTGG
16	GUS-RT-R	qPCR	GCATCTCTTCAGCGTAAGGGTAAT

photoperiod of 16 h light/8 h dark for 20–25 days. The *A. annua* plants at a height of approximately 5 cm can be used for transformation. Leaf discs (0.5 cm diameter) were cut and placed on the co-cultivation medium (1/2 MS medium with 100 μmol/L Acetosyringone) containing *A. tumefaciens* at 25 °C for 3 days (McCormick et al. 1986). Then the explants of *A. annua* were transferred to shoot-inducing MS1 medium (MS medium fortified with 5 mg/L 6-benzoyladenine, 0.05 mg/L naphthalene-1-acetic acid, 50 mg/L hygromycin and 500 mg/L carbenicillin) for 2 weeks. After subculturing for 30 days in hygromycin selection medium (change of medium every 10 days), the hygromycin-resistant buds were selected and transferred into root-inducing medium MS2 (1/2 MS + 250 mg/L carbenicillin). After 1–2 months for rooting, the plantlets were transferred into soil in the greenhouse.

GUS assay

Histochemical GUS staining was carried out as previously described (Jefferson et al. 1987). Different tissues from the transgenic and non-transgenic *A. annua* were collected for this experiment. Samples were immersed in GUS reaction buffer (50 mM sodium phosphate buffer, 1 mM $K_3Fe(CN)_6$, 1 mM $K_4Fe(CN)_6$, 0.1 % (v/v) Triton X-100 and 0.1 Mm X-Gluc, pH 7.0) for 12 h at 37 °C in the dark. Ethanol was used to stop the GUS reaction and to eliminate chlorophyll from the tissues at a temperature of 55 °C for 1 h. To obtain good GUS straining effect of *A. annua* tissues, optical microscope (Olympus BX51) was used.

RT-Q-PCR

RNA was extracted from different tissues using plant RNA prep pure Plant kit (Tiangen Biotech, Beijing, China), according to the manufacturer's instruction. The RNA was reversely transcribed by the PrimeScript II 1st Strand cDNA Synthesis Kit (Takara Biotechnology, Dalian, China) to synthesize the first-strand cDNA. Primers used for the RT-Q-PCR are listed in Table 1. The *β-Actin* was used as a standard control gene. RT-Q-PCR was performed with a LightCycler 96 System (Roche, Mannheim, Germany) by the SuperReal PreMix Plus (SYBR Green) kit (Tiangen Biotech, Beijing, PR China) following the manufacturer's instruction. The cycling for RT-Q-PCR was performed at 95 °C (10 min), 40 cycles at 95 °C (20 s), 55 °C (20 s), and 72 °C (20 s). Each cDNA sample was repeated 3 times.

Hormonal treatments

Both wild-type and transgenic plants with pCAM-BIA1391Z::*pALDH1*::*GUS* of *A. annua* were propagated by shoot cutting and grown in soil under the condition of 16 h light/8 h dark at a constant temperature of 24 °C. After 1-month growth, they were subjected to hormonal treatments. The plants were divided into three groups, and each group with 10 individual plants used for biological repeats was sprayed with 100 μM methyljasmonate (MeJA), 100 μM gibberellin (GA) or 100 μM salicylic acid (SA), separately. RT-Q-PCR was used to evaluate the transcript levels at 0, 1, 3, 6, 9 and 12 h after hormonal treatments.

Statistical analysis

Outliers were detected by using Grubbs test (Miller and Miller 2010). Student's t test was used to analyze statistical significance.

Results and discussion

Analysis of promoter nucleotide sequence

A 1620-bp DNA fragment was cloned from genomic DNA of *A. annua*, and the nucleotide sequence is showed in Fig. 2.

We used the TSSP software (http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter) to search more information about this promoter. A putative transcription start site (TSS) of *ALDH1* was found and labeled +1 in Fig. 2, which was located at 49 bp upstream of the *ALDH1* start codon (ATG). The TATA-box (TATA) is a well-conserved core promoter element usually located about 25–32 bp upstream of the TSS in eukaryotes. In this promoter, a putative TATA-box was predicted at the positions of −36 to −28 (ATTATAATA) of the *ALDH1* promoter (Fig. 2). The CAAT-box is another important *cis*-acting element, which is considered to be the binding site for the RNA polymerase. A CAAT-box (CAAT) was also found at −119 to −116 upstream of the putative TSS of *ALDH1* promoter.

To further understand the promoter, we used Plant-CARE software (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict the *cis*-acting elements, and boxes are listed in Table 2. The in silico analysis indicates that the expression of the *ALDH1* gene can be regulated by abiotic stress, biotic stress, multiple hormonal treatments and some transcription factor families.

Plant transcription factor database (http://planttfdb.cbi.pku.edu.cn/index.php?sp=Aan) is a useful software to search the transcription factors, and 28 MYB family transcription factors were found in *A. annua*. We have recently cloned a MYB transcription factor from *A. annua* (unpublished). Overexpression of this MYB gene in GSTs resulted in an increasing transcript levels of *ADS*, *CYP71AV1*, *DBR2* and *ALDH1*, raising biosynthesis of artemisinin in transgenic *A. annua*. One MYB binding site with the sequence CAACTG was discovered at position −1487 to −1482 of *ALDH1* promoter which may be related to drought inducibility.
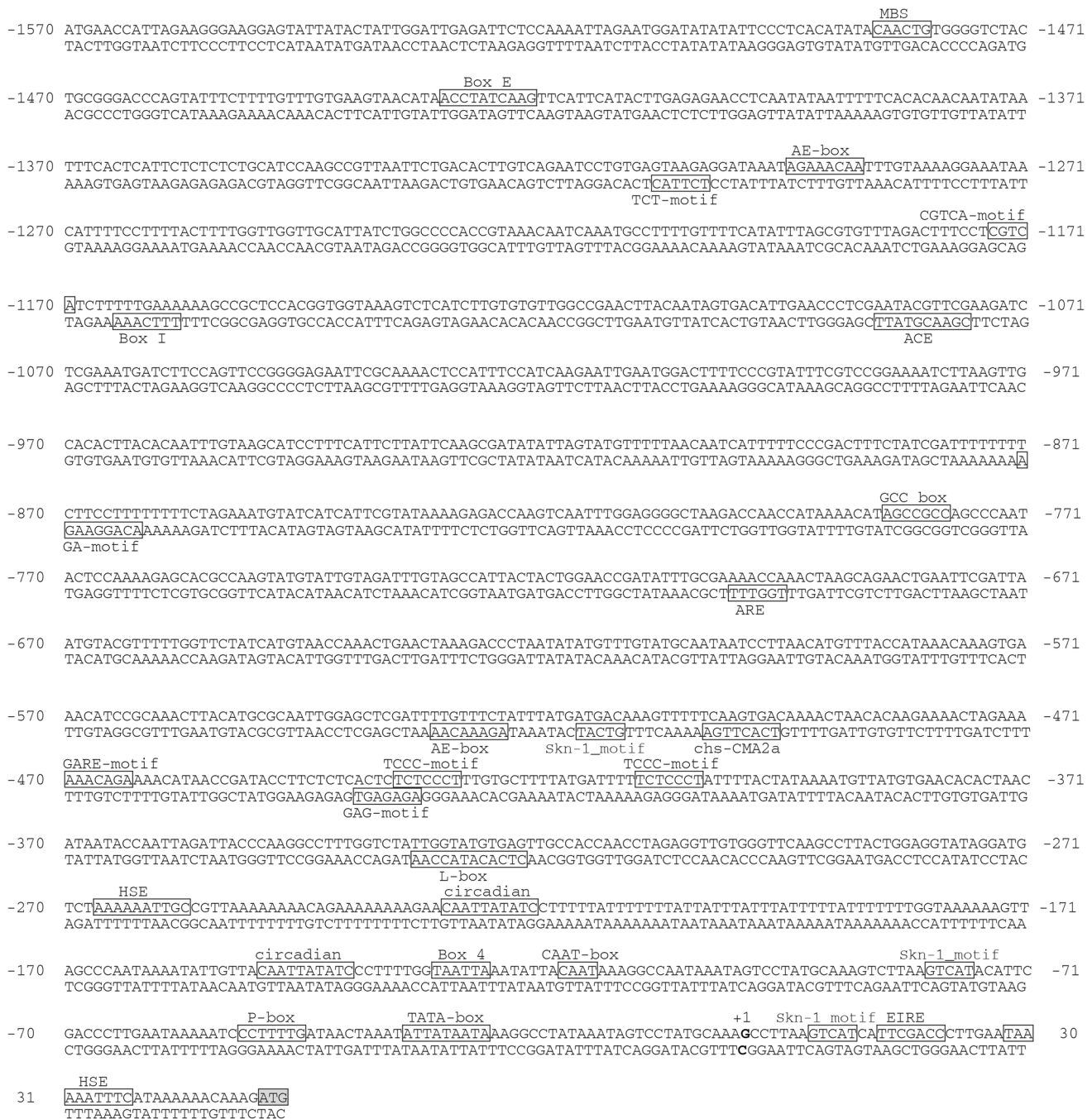
Previous studies showed that treatment of *A. annua* plants with different hormones may increase the production of artemisinin. Treatment with SA (1.0 mM) can up-regulate artemisinin biosynthesis in *A. annua*. The artemisinin content was increased by 54 % compared with the control

after SA treatment (Pu et al. 2009). GA_3 treatment can also improve the production of artemisinin and leaf biomass significantly in *A. annua*, and the expression of *FDS, ADS* and *CYP71AV1* were increased simultaneously (Banyai et al. 2011). Similarly, treatment with MeJA also affects artemisinin production and gene expression in *A. annua* (Caretto et al. 2011). The *ALDH1* promoter carries many *cis*-acting elements putatively regulated by SA, MeJA and GA, which are similar with the promoters of *ADS* and *CYP71AV1* (Wang et al. 2011, 2013). A TCA-element involved in SA responsiveness was found at position −1172 to −1163. A putative MeJA responsive element CGTCA is present in the *ALDH1* promoter at position −1174 to −1170. The GARE-motif (AAACAGA) is involved in the responsiveness to GA. One such *cis*-acting element was found at −470 to −464 upstream of TSS of *ALDH1* promoter. The P-box with the sequence CCTTTTG is also a GA-responsive element, located at −52 to −46 upstream of TSS.

Moreover, the *ALDH1* promoter carries many putative *cis*-elements related to light responsiveness, implying that the expression of *ALDH1* can be regulated by light. These elements include one ACE-box (CTAACGTATT; position 1086 to −1077), one chs-CMA2a (TCACTTGA; position −504 to −497), two TCCC-motifs (TCTCCCT; position −411 to −405 and −436 to −430), one GAG-motif (AGAGAGT; position −434 to −440), one GA-motif (AAAGATGA; position −861 to −870), one L-box (CTCACCTACCAA, position −335 to −324), two AE-boxes (AGAAACAA; position −533 to −526 and −1296 to −1289), one Box 4 (ATTAAT; position −133 to −128), one Box I (TTTCAAA; position −1166 to −1160) and one TCT-motif (TCTTAC; position −1310 to −1305).

In addition, the *ALDH1* promoter contains other putative *cis*-acting regulatory elements. One ARE (TGGTTT) was found at position −702 to −697, which is a putative element essential for the anaerobic induction; two HSE elements involved in heat stress responsiveness were found at the positions of +28 to +37 (AAAAAATTTC) and −267 to −258 (AAAAAATTGC). Moreover, one elicitor-responsive element EIRE (TTCGACC; position +15 to +21) was found in the *ALDH1* promoter, and two circadian elements (CAATTATATC; position −151 to −142 and −232 to −223) involved in circadian control were also discovered. Finally, three Skn-1 motifs (GTCAT; position −517 to −513, −81 to −77, +8 to +12) related to endosperm expression were identified in this promoter, suggesting that *ALDH1* may participate in the seed development. This support the research that artemisinin precursors can be detected in seeds (Brown et al. 2003).

Based on a large number of putative *cis*-acting elements present in the *ALDH1* promoter, we may conclude that the expression of *ALDH1* is under complex control.

```
                                                                                                         MBS
-1570 ATGAACCATTAGAAGGGAAGGAGTATTATACTATTGGATTGAGATTCTCCAAAATTAGAATGGATATATATTCCCTCACATATACAACTGTGGGGTCTAC -1471
      TACTTGGTAATCTTCCCTTCCTCATAATATGATAACCTAACTCTAAGAGGTTTTAATCTTACCTATATATAAGGGAGTGTATATGTTGACACCCCAGATG

                                          Box E
-1470 TGCGGGACCCAGTATTTCTTTTGTTTGTGAAGTAACATAACCTATCAAGTTCATTCATACTTGAGAGAACCTCAATATAATTTTTCACACAACAATATAA -1371
      ACGCCCTGGGTCATAAAGAAAACAAACACTTCATTGTATTGGATAGTTCAAGTAAGTATGAACTCTCTTGGAGTTATATTAAAAAGTGTGTTGTTATATT

                                                                                        AE-box
-1370 TTTCACTCATTCTCTCTCTGCATCCAAGCCGTTAATTCTGACACTTGTCAGAATCCTGTGAGTAAGAGGATAAATAGAAACAATTTGTAAAAGGAAATAA -1271
      AAAGTGAGTAAGAGAGAGACGTAGGTTCGGCAATTAAGACTGTGAACAGTCTTAGGACACTCATTCTCCTATTTATCTTTGTTAAACATTTTCCTTTATT
                                                            TCT-motif
                                                                                          CGTCA-motif
-1270 CATTTTCCTTTTACTTTTGGTTGGTTGCATTATCTGGCCCCACCGTAAACAATCAAATGCCTTTTGTTTTCATATTTAGCGTGTTTAGACTTTCCTCGTC -1171
      GTAAAAGGAAAATGAAAACCAACCAACGTAATAGACCGGGGTGGCATTTGTTAGTTTACGGAAAACAAAAGTATAAATCGCACAAATCTGAAAGGAGCAG


-1170 ATCTTTTTGAAAAAAGCCGCTCCACGGTGGTAAAGTCTCATCTTGTGTGTTGGCCGAACTTACAATAGTGACATTGAACCCTCGAATACGTTCGAAGATC -1071
      TAGAAAAACTTTTTTCGGCGAGGTGCCACCATTTCAGAGTAGAACACACAACCGGCTTGAATGTTATCACTGTAACTTGGGAGCTTATGCAAGCTTCTAG
          Box I                                                                          ACE

-1070 TCGAAATGATCTTCCAGTTCCGGGGAGAATTCGCAAAACTCCATTTCCATCAAGAATTGAATGGACTTTTCCCGTATTTCGTCCGGAAAATCTTAAGTTG -971
      AGCTTTACTAGAAGGTCAAGGCCCCTCTTAAGCGTTTTGAGGTAAAGGTAGTTCTTAACTTACCTGAAAAGGGCATAAAGCAGGCCTTTTAGAATTCAAC


-970 CACACTTACACAATTTGTAAGCATCCTTTCATTCTTATTCAAGCGATATATTAGTATGTTTTTAACAATCATTTTTCCCGACTTTCTATCGATTTTTTTT -871
     GTGTGAATGTGTTAAACATTCGTAGGAAAGTAAGAATAAGTTCGCTATATAATCATACAAAAATTGTTAGTAAAAGGGCTGAAAGATAGCTAAAAAAAA

                                                                                       GCC_box
-870 CTTCCTTTTTTTTTCTAGAAATGTATCATCATTCGTATAAAAGAGACCAAGTCAATTTGGAGGGGCTAAGACCAACCATAAAACATAGCCGCCAGCCCAAT -771
     GAAGGACAAAAAGATCTTTACATAGTAGTAAGCATATTTTCTCTGGTTCAGTTAAACCTCCCCGATTCTGGTTGGTATTTTGTATCGGCGGTCGGGTTA
     GA-motif

-770 ACTCCAAAAGAGCACGCCAAGTATGTATTGTAGATTTGTAGCCATTACTACTGGAACCGATATTTGCGAAAACCAAACTAAGCAGAACTGAATTCGATTA -671
     TGAGGTTTTCTCGTGCGGTTCATACATAACATCTAAACATCGGTAATGATGACCTTGGCTATAAACGCTTTTGGTTTGATTCGTCTTGACTTAAGCTAAT
                                                                            ARE

-670 ATGTACGTTTTTGGTTCTATCATGTAACCAAACTGAACTAAAGACCCTAATATATGTTTGTATGCAATAATCCTTAACATGTTTACCATAAACAAAGTGA -571
     TACATGCAAAAACCAAGATAGTACATTGGTTTGACTTGATTTCTGGGATTATATACAAACATACGTTATTAGGAATTGTACAAATGGTATTTGTTTCACT


-570 AACATCCGCAAACTTACATGCGCAATTGGAGCTCGATTTTGTTTCTATTTATGATGACAAAGTTTTTCAAGTGACAAAACTAACACAAGAAAACTAGAAA -471
     TTGTAGGCGTTTGAATGTACGCGTTAACCTCGAGCTAAAACAAAGATAAATACTACTGTTTCAAAAGTTCACTGTTTTGATTGTGTTCTTTTGATCTTT
     GARE-motif                                     AE-box     Skn-1_motif    chs-CMA2a
                                                                                   TCCC-motif
-470 AAACAGAAAACATAACCGATACCTTCTCTCACTCTCTCCCTTTGTGCTTTTATGATTTTTCTCCCTATTTTACTATAAAATGTTATGTGAACACACTAAC -371
     TTTGTCTTTTGTATTGGCTATGGAAGAGAGTGAGAGAGGGAAACACGAAAATACTAAAAAGAGGGATAAAATGATATTTTACAATACACTTGTGTGATTG
     TCCC-motif                             GAG-motif

-370 ATAATACCAATTAGATTACCCAAGGCCTTTGGTCTATTGGTATGTGAGTTGCCACCAACCTAGAGGTTGTGGGTTCAAGCCTTACTGGAGGTATAGGATG -271
     TATTATGGTTAATCTAATGGGTTCCGGAAACCAGATAACCATACACTCAACGGTGGTTGGATCTCCAACACCCAAGTTCGGAATGACCTCCATATCCTAC
                                                                 L-box
                HSE                                             circadian
-270 TCTAAAAAATTGCCGTTAAAAAAAACAGAAAAAAAAGAACAATTATATCCTTTTTATTTTTTTATTATTTATTTATTTTTATTTTTTTGGTAAAAAAGTT -171
     AGATTTTTTAACGGCAATTTTTTTGTCTTTTTTTTCTTGTTAATATAGGAAAATAAAAAAATAATAAATAAATAAAAATAAAAAAACCATTTTTTCAA

                     circadian                   Box 4         CAAT-box                       Skn-1_motif
-170 AGCCCAATAAAATATTGTTACAATTATATCCCTTTTGGTAATTAAATATTACAATAAAGGCCAATAAATAGTCCTATGCAAAGTCTTAAGTCATACATTC -71
     TCGGGTTATTTTATAACAATGTTAATAAATAGGGAAAACCATTAATTTATAATGTTATTTCCGGTTATTTATCAGGATACGTTTCAGAATTCAGTATGTAAG

                   P-box            TATA-box                               +1  Skn-1 motif   EIRE
-70 GACCCTTGAATAAAAATCCCTTTTGATAACTAAATATTATAATAAAGGCCTATAAATAGTCCTATGCAAAGCCTTAAGTCATCATTCGACCCTTGAATAA 30
    CTGGGAACTTATTTTTAGGGAAAACTATTGATTTATAATATTATTTCCGGATATTTATCAGGATACGTTTCGGAATTCAGTAGTAAGCTGGGAACTTATT

         HSE
31 AAATTTCATAAAAAACAAAGATG
   TTTAAAGTATTTTTTGTTTCTAC
```

Fig. 2 Nucleotide sequence and putative *cis*-acting elements in the 1620-bp *ALDH1* promoter region. The putative transcript start site is in *bold*. The *cis*-acting elements and ATG start codon are in *box*

Study on the activity of *ALDH1* promoter

In our study, we analyzed the *ALDH1* promoter using a reporter gene *GUS*-fusion, and the recombinant plasmids were transformed into *A. annua*. Positive T_0 transgenic lines confirmed by PCR showed a similar GUS staining pattern while the control plants did not. This indicates that the *ALDH1* promoter drives the downstream *GUS* gene expression in *A. annua*.

According to Fig. 3, we found that *GUS* expression showed a spatial difference in leaves at different ages. In young leaves, GUS staining is only observed in GSTs with strong activity (Fig. 3a–c). As the plant grows older, GSTs still showed *GUS* expression, but near the leaf veins, GUS

Table 2 Putative *cis*-acting elements involved in the promoter of *ALDH1*

Motif	Nucleotide sequence	Description
ACE	CTAACGTATT	Cis-acting element involved in light responsiveness
AE-box	AGAAACAA	Part of a module for light response
ARE	TGGTTT	Cis-acting regulatory element essential for the anaerobic induction
Box I	TTTCAAA	Light responsive element
CGTCA-motif	CGTCA	Cis-acting regulatory element involved in the MeJA-responsiveness
GA-motif	AAAGATGA	Part of a light responsive element
GAG-motif	AGAGAGT	Part of a light responsive element
GARE-motif	AAACAGA	Gibberellin-responsive element
HSE	AAAAAATTTC	Cis-acting element involved in heat stress responsiveness
L-box	CTCACCTACCAA	Part of a light responsive element
MBS	CAACTG	MYB binding site involved in drought inducibility
P-box	CCTTTTG	Gibberellin-responsive element
Skn-1 motif	GTCAT	Cis-acting regulatory element required for endosperm expression
TCA-element	CCATCTTTTT	Cis-acting element involved in salicylic acid responsiveness
TCCC-motif	TCTCCCT	Part of a light responsive element
TCT-motif	TCTTAC	Part of a light responsive element
TGACG-motif	TGACG	Cis-acting regulatory element involved in the MeJA-responsiveness
Chs-CMA2a	TCACTTGA	Part of a light responsive element
Circadian	CAANNNNATC	Cis-acting regulatory element involved in circadian control

staining was also observed in TSTs (Fig. 3d–f). In leaves at nodes 7–8, changes have occurred and only TSTs exhibited GUS staining (Fig. 3g–i). In the bottom leaves, no GUS staining was observed (Fig. 3l). *GUS* expression pattern shows similarities with results from studies on the *CYP71AV1* promoter (Wang et al. 2013). Transgenic plants with the empty pCAMBIA1391Z vector did not show any GUS staining (Fig. 3m). Furthermore, we did not detect any promoter activity in the TSTs of stems (Fig. 4a–c). No GUS signals in roots were found yet (Fig. 4d–f).

Histochemical GUS staining revealed strong blue signal in GSTs of young leaves. This indicated that the cloned promoter contained *cis*-acting regulatory elements related to specific expression in GSTs. The finding raises the possibility that the fastest period of artemisinin precursors' production (i.e. dihydroartemisinic acid and other upstream precursors) is in young leaf cycle (Lommen et al. 2006). During leaf development, the *GUS* expression showed a shift from GSTs to TSTs. This phenomenon is in accordance with a previous study that *ALDH1* was expressed at low levels in T-shaped trichomes (Soetaert et al. 2013). Furthermore, some linear short aldehydes can also be oxidized by ALDH1 (Teoh et al. 2009), indicating that other reactions oxidized by ALDH1 may take place in the TSTs of old leaves. The collapse of trichomes can explain why no GUS staining was found in old leaves, and contribute to the decline of artemisinin during leaf development (Lommen et al. 2006). As shown above and described in the literature, we can conclude that the *ALDH1* is specifically expressed in different trichomes at different time.

Expression pattern of *pALDH1-ALDH1* and *pALDH1-GUS* in *A. annua*

The relative activities of wild-type (*pALDH1-ALDH1*) and recombinant *ALDH1* (*pALDH1-GUS*) promoters were analyzed at different leaf positions by RT-Q-PCR. In order to make a comprehensive investigation of artemisinin biosynthetic pathway, we examined other genes involved (*ADS*, *CYP71AV1*, *DBR2*) in wild-type plants. Data of transcript levels were estimated by $2^{-\triangle\triangle Ct}$ method (Livak and Schmittgen 2001). The relative expression in leaf 1 was set to 1.0 for each transcript. According to the results (Fig. 5), the highest relative activity of wild-type promoter was in young leaves (leaf 1), and it dropped gradually during leaf development (leaf 4 and leaf 9). In the old leaves (leaf 16) no expression could be detected. The observed expression of *pALDH1-GUS* was similar with *pALDH1-ALDH1*. This is in line with the expression pattern of *ADS* and *CYP71AV1* and *DBR2*.

As shown in Fig. 6, the expression patterns of *pALDH1-ALDH1* and *pALDH1-GUS* in different tissues were also quantified by RT-Q-PCR. The relative expression in young leaf
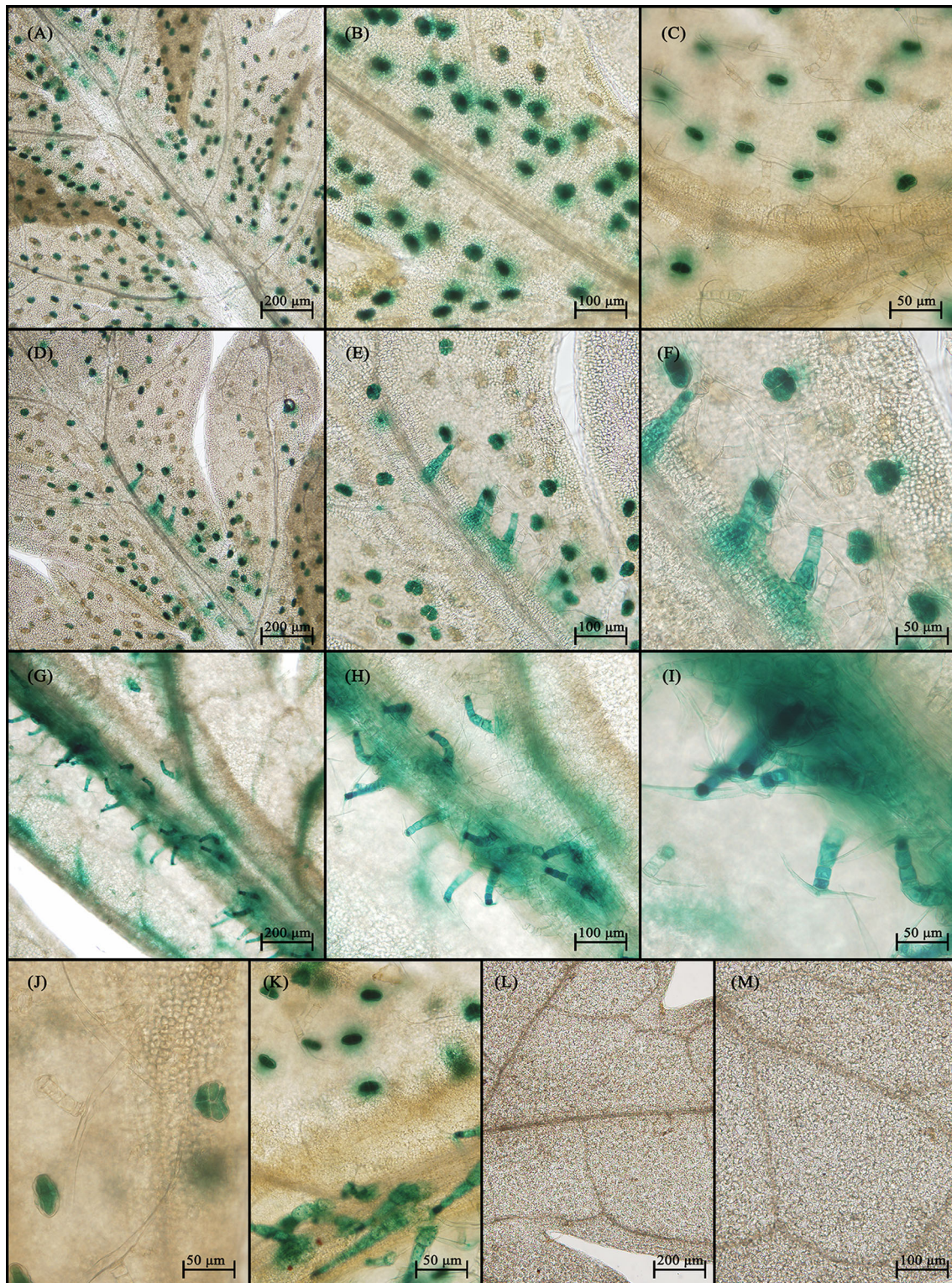
Fig. 3 GUS staining in transgenic *A. annua*: **a** leaf primordia of transgenic *A. annua* with a GUS-fusion, **b** magnification (2×) of leaf primordia, **c** magnification (4×) of leaf primordia, **d** leaf 3 of transgenic *A. annua* with a GUS-fusion; **e** magnification (2×) of leaf 3, **f** magnification (4×) of leaf 3, **g** leaf 8 of transgenic *A. annua* with a GUS-fusion, **h** magnification (2×) of leaf 8, **i** magnification (4×) of leaf 8, **j** leaf 1 (4×) of transgenic *A. annua* with a GUS-fusion, **k** magnification (4×) of leaf 4, **l** magnification (4×) of leaf 16 and **m** control with an empty pCAMBIA1391Z
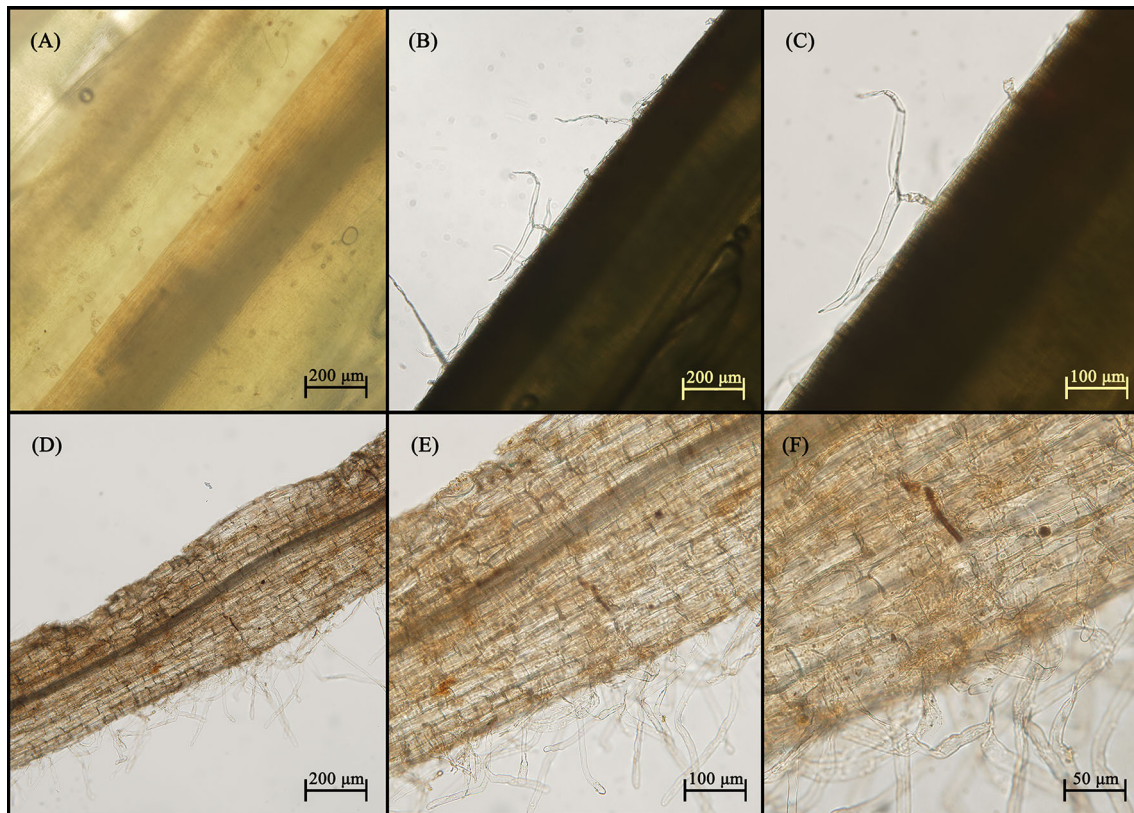
Fig. 4 GUS staining in other tissues of transgenic *A. annua*: **a** stems of transgenic *A. annua*, **b** TSTs in stems of transgenic *A. annua*, **c** magnification (2×) of TSTs in stems, **d** roots of transgenic *A. annua*, **e** magnification (2×) of roots in transgenic *A. annua*, **f** magnification (4×) of roots in transgenic *A. annua*

was set to 1.0. The data showed that both the expressions of *pALDH1-GUS* and *pALDH1-ALDH1* exhibited a spatial difference. The highest relative expression was seen in leaf primordia and young leaves, and poorly showed in old leaves, roots and stems. The other three genes, *ADS*, *CYP71AV1* and *DBR2*, showed similar expression patterns. This finding may be related to the distribution of trichomes in different tissues of *A. annua*.

The results of RT-Q-PCR analysis indicate that the cloned promoter is effective for controlling the gene expression in *A. annua*. The consistence of GUS assay and expression analysis also reflected that the *ALDH1* gene specifically expressed in GSTs of young leaves and that the artemisinin precursors may be synthesized in relatively young leaves within GSTs in early developmental stages (Teoh et al. 2006; Olsson et al. 2009).

Response to hormones

Previous studies showed that spaying exogenous hormones on *A. annua* may up-regulate the artemisinin biosynthesis (Pu et al. 2009; Waleerat et al. 2010; Guo et al. 2010). In order to determine the expression pattern of the *pALDH1-*

ALDH1 and *pALDH1-GUS* after hormonal treatments, we sprayed *A. annua* with MeJA, GA and SA respectively, and performed the analysis by RT-Q-PCR. The relative expression at 0 h was set to 1.0. According to Fig. 7, the transcript levels of *pALDH1-GUS* at 3 h were about two-folds than that of the control after MeJA treatment, and then declined gradually. The recombinant promoter carries many motifs that are involved in GA-induced expression, and the highest expression of *pALDH1-GUS* was at 9 h after GA treatment. A similar expression pattern was also found after SA treatment. Further, transcripts of *pALDH1-ALDH1* were also induced by these hormones (Fig. 7). These results are in accordance with the in silico analysis that the *ALDH1* promoter carries many hormonal responsive *cis*-acting elements and the results suggest that up-regulation of *ALDH1* by hormonal treatments may be a key factor to enhance artemisinin content in *A. annua*.
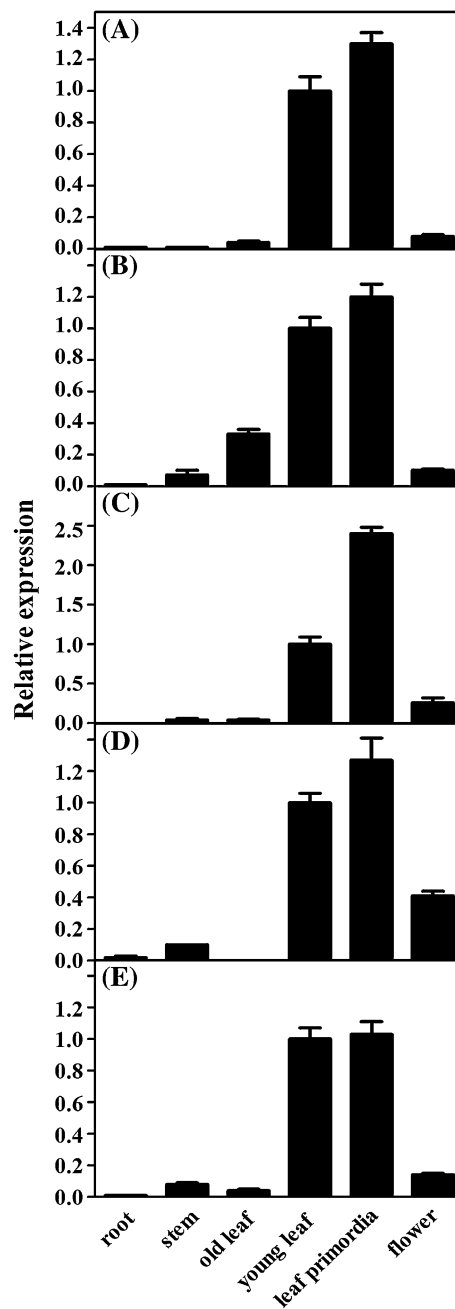
Conclusion

A 1620-bp *ALDH1* promoter was cloned, putative *cis*-acting elements related to biotic and abiotic factors were predicted using software. The *GUS* expression analysis
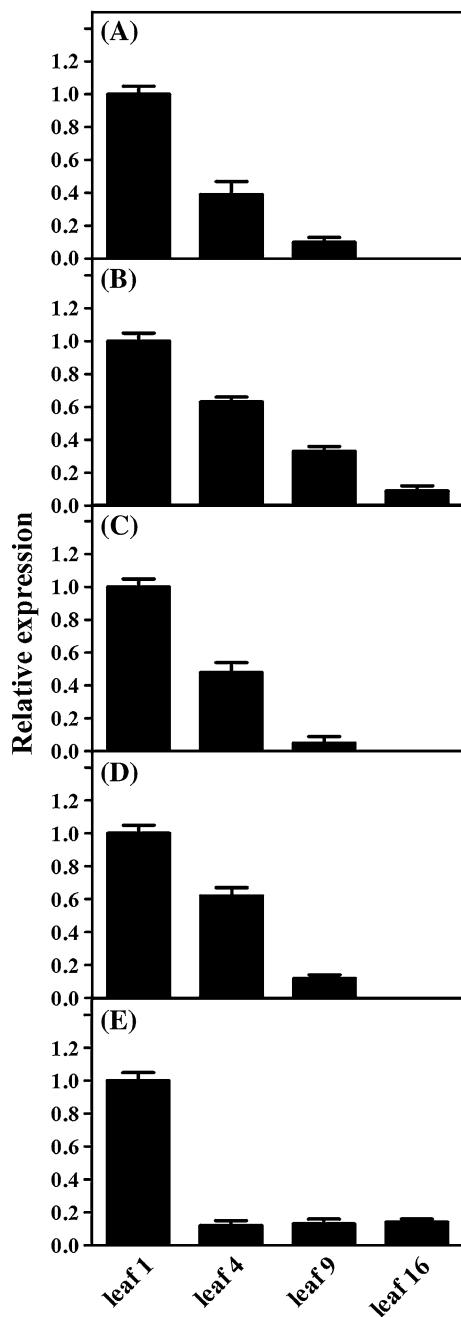
Fig. 5 Expression patterns in different nodes (leaf 1, leaf 4, leaf 9, leaf 16) of leaves: **a** relative expression of *pALDH1-ALDH1*, **b** relative expression of *pALDH1-GUS*, **c** relative expression of *ADS*, **d** relative expression of *CYP71AV1*, **e** relative expression of *DBR2*

Fig. 6 Expression patterns in different tissues (root, stem, old leaf, young leaf, leaf primordia and flower): **a** relative expression of *pALDH1-ALDH1*, **b** relative expression of *pALDH1-GUS*, **c** relative expression of *ADS*, **d** relative expression of *CYP71AV1*, **e** relative expression of *DBR2*

shows that the recombinant promoter is GST-specific in young leaves, and it changes to TST expression in older leaves. No GUS staining was observed in stems, roots or very old leaves. It is most likely that the *ALDH1* promoter is involved in some biosynthetic processes in TSTs in addition to being involved in artemisinin biosynthesis in GSTs. Further studies on the relationship between GSTs and TSTs should be conducted. Studies on the binding of

different transcription factors to *ALDH1* promoter will reveal the regulatory mechanism of trichomes. Expression analysis of *ALDH1* in different tissues indicated that *ALDH1* showed high transcript levels in young leaves, and other enzymatic genes involved in the biosynthesis of artemisinin (*ADS*, *CYP71AV1* and *DBR2*) have similar characteristics. Taken together, we can conclude that the
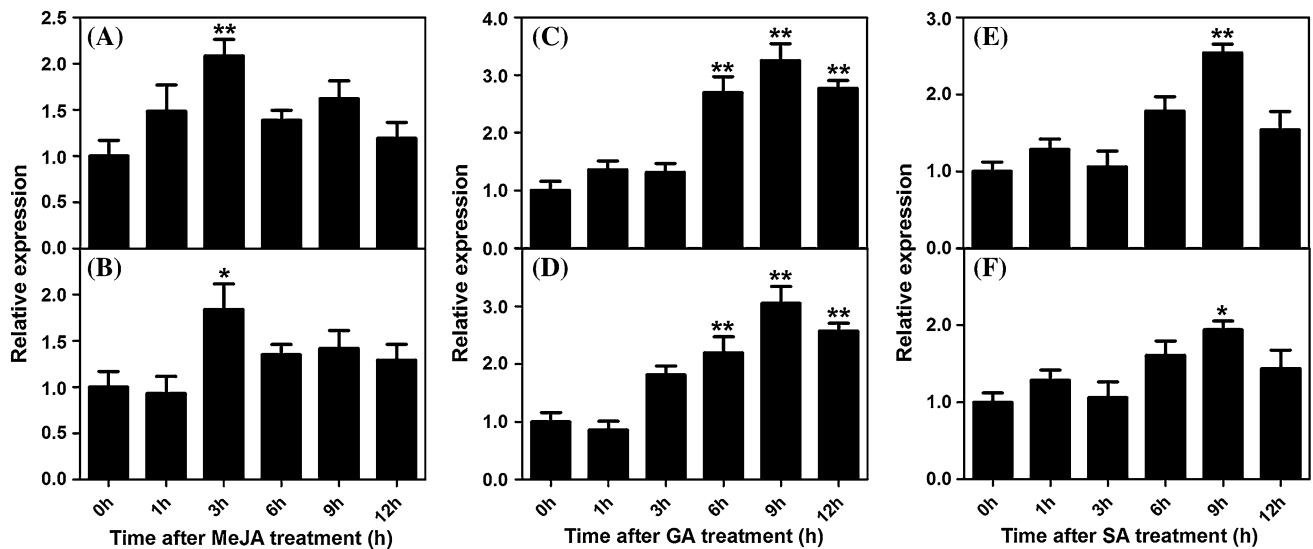
Fig. 7 Expression patterns of *ALDH1* in response to hormones: relative expression of *pALDH1-ALDH1* (**a**) and *pALDH1-GUS* (**b**) after MeJA; relative expression of *pALDH1-ALDH1* (**c**) and *pALDH1-GUS* (**d**) after GA, **c** relative expression of *pALDH1-ALDH1* (**e**) and *pALDH1-GUS* (**f**) after SA

artemisinin precursors may be synthesized in the GSTs of young leaves. Furthermore, the expression of *ALDH1* can be up-regulated by hormones which make a great contribution to the enhancement of artemisinin content. Using the *ALDH1* promoter to guide the expression of a given gene in trichomes provides a new strategy for metabolic engineering of plants.

References

Arsenault PR, Vail D, Wobbe KK, Erickson K, Weathers PJ (2010) Reproductive development modulates gene expression and metabolite levels with possible feedback inhibition of artemisinin in *Artemisia annua*. Plant Physiol 154:958–968

Bakhsh A, Rao QA, Shamim Z, Husnain T (2011) Rubisco small subunit as a strong, green tissue-specific promoter. Arch Biol Sci 63:299–307

Banyai W, Mii M, Supaibulwatana K (2011) Enhancement of artemisinin content and biomass in *Artemisia annua* by exogenous GA3 treatment. Plant Growth Regul 63:45–54

Bouwmeester HJ, Wallaart TE, Janssen MH et al (1999) Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. Phytochem 52:843–854

Brown GD, Sy LK (2004) In vivo transformations of dihydroartemisinic acid in *Artemisia annua* plants. Tetrahedron 60:1139–1159

Brown GD, Sy LK (2007) In vivo transformations of artemisinic acid in *Artemisia annua* plants. Tetrahedron 63:9548–9566

Brown GD, Liang GY, Sy LK (2003) Terpenoids from the seeds of *Artemisia annua*. Phytochem 64:303–323

Caretto S, Quarta A, Durante M, Nisi R, De Paolis A, Blando F, Mita G (2011) Methyl jasmonate and miconazole differently affect arteminisin production and gene expression in *Artemisia annua* suspension cultures. Plant Biol 13:51–58

Ferreira JFS, Janick J (1996) Distribution of artemisinin in *Artemisia annua*. In: Janick J (ed) Progress in new crops. ASHS, Arlington, pp 579–584

Fütterer J, Gisel A, Iglesias V, Klöti A, Kost B, Mittelsten Scheid O, Neuhaus G, Neuhaus-Url G, Schrott M, Shillito R, Spangenberg G, Wang ZY (1995) Standard molecular techniques for the analysis of transgenic plants. In: Potrykus I, Spangenberg G (eds) Gene transfer to plants. Springer, Berlin, pp 215–263

Guo XX, Yang XQ, Yang RY, Zeng QP (2010) Salicylic acid and methyl jasmonate but not Rose Bengal enhance artemisinin production through invoking burst of endogenous singlet oxygen. Plant Sci 178:390–397

Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusion: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plant. EMBO J 6:3901–3907

Jiang WM, Lu X, Qiu B, Zhang FY, Shen QY, Lv ZY, Fu XQ, Yan TX, Gao ED, Zhu MM, Chen LX, Zhang L, Wang GF, Sun XF, Tang KX (2014) Molecular cloning and characterization of a trichome-specific promoter of artemisinic aldehyde △11(13) reductase (DBR2) in *Artemisia annua*. Plant Mol Biol Rep 32:82–91

Kapoor R, Chaudhary V, Bhatnagar AK (2007) Effects of arbuscular mycorrhiza and phosphorus application on artemisinin concentration in *Artemisia annua* L. Mycorrhiza 17:581–587

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCt method. Methods 25:402–408

Lommen WJM, Schenk E, Bouwmeester HJ, Verstappen FWA (2006) Trichome dynamics and artemisinin accumulation during development and senescence of *Artemisia annua* leaves. Planta Med 72:336–345

Lu X, Shen Q, Zhang L, Zhang FY, Jiang WM, Lv ZY, Yan TX, Fu XQ, Wang GF, Tang KX (2013) Promotion of artemisinin biosynthesis in transgenic *Artemisia annua* by overexpressing ADS, CYP71AV1 and CPR genes. Ind Crop Prod 49:380–385

McCormick S, Niedermeyer J, Fry J, Barnason A, Horsch R, Fraley R (1986) Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. Plant Cell Rep 5(2):81–84

Mercke P, Bengtsson M, Bouwmeester HJ, Posthumus MA, Brodelius PE (2000) Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. Arch Biochem Biophys 381:173–180

Miller JN, Miller JC (2010) Statistics and chemometrics for analytical chemistry, 6th edn. Prentice Hall, England

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Nguyen KT, Arsenault PR, Weathers PJ (2011) Trichomes plus roots plus ROS=artemisinin: regulating artemisinin biosynthesis in *Artemisia annua* L. In Vitro Cell Dev Biol-Plant 47:329–338

Olsson ME, Olofsson LM, Lindahl A-L, Lundgren A, Brodelius M, Brodelius PE (2009) Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of *Artemisia annua* L. Phytochemistry 70:1123–1128

Pu GB, Ma DM, Chen JL, Ma LQ, Wang H, Li GF, Ye HC, Liu BY (2009) Salicylic acid activates artemisinin biosynthesis in *Artemisia annua* L. Plant Cell Rep 28(7):1127–1135

Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MC, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–943

Soetaert SS, Van Neste CM, Vandewoestyne ML, Head SR, Goossens A, Van Nieuwerburgh FC, Deforce DL (2013) Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. BMC Plant Biol 13(1):220

Sy LK, Brown GD (2002) The mechanism of the spontaneous autoxidation of dihydroartemisinic acid. Tetrahedron 58:897–908

Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS (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

Teoh KH, Polichuk DR, Reed DW, Covello PS (2009) Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis in *Artemisia annua*. Botany 87:635–642

Waleerat B, Chalermpol K, Masahiro M, Kanyaratt S (2010) Overexpress of farnesyl pyrophosphate synthase (FPS) gene affected artemisinin content and growth of *Artemisia annua* L. Plant Cell Tiss Org Cult 103:255–265

Wang H, Olofsson F, Lundgren A, Brodelius PE (2011) Trichome-specific expression of amorpha-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L., as reported by a promoter-GUS fusion. Am J Plant Sci 2:619–628

Wang H, Han J, Kanagarajan S, Lundgren A, Brodelius PE (2013) Trichome-specific expression of the amorpha-4,11-diene 12-hydroxylase (cyp71av1) gene, encoding a key enzyme of artemisinin biosynthesis in *Artemisia annua*, as reported by a promoter-GUS fusion. Plant Mol Biol 81(1–2):119–138

World Health Organization (2015) World malaria report 2014. WHO, Geneva

Zhang YS, Teoh KH, Reed DW, Maes L, Goossens A, Olson DJ, Ross AR, Covello PS (2008) The molecular cloning of artemisinic aldehyde 11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. J Biol Chem 283:21501–21508

Zhu M, Zhang F, Lv Z, Shen Q, Zhang L, Lu X, Jiang WM, Fu XQ, Yang TX, Chen LX, Wang GF, Tang KX (2014) Characterization of the promoter of *Artemisia annua* amorpha-4,11-diene synthase (ADS) gene using homologous and heterologous expression as well as deletion analysis. Plant Mol Biol Rep 32:406–418