© Indian Academy of Sciences

RESEARCH ARTICLE

Comparative analysis of *ADS* gene promoter in seven *Artemisia* species

MOJTABA RANJBAR[1,2], MOHAMMAD REZA NAGHAVI[1]* and HOSHANG ALIZADEH[1]

[1]*Agronomy and Plant Breeding Department, College of Agriculture, University of Tehran, Karaj, 31587-11167, Iran*
[2]*Faculty of Biotechnology, Amol University of Special Modern Technologies, Amol, 46168-49767, Iran*

Abstract

Artemisinin is the most effective antimalarial drug that is derived from *Artemisia annua*. Amorpha-4,11-diene synthase (ADS) controls the first committed step in artemisinin biosynthesis. The *ADS* gene expression is regulated by transcription factors which bind to the *cis*-acting elements on the ADS promoter and are probably responsible for the *ADS* gene expression difference in the *Artemisia* species. To identify the elements that are significantly involved in *ADS* gene expression, the *ADS* gene promoter of the seven *Artemisia* species was isolated and comparative analysis was performed on the ADS promoter sequences of these species. Results revealed that some of the *cis*-elements were unique or in terms of number were more in the high artemisinin producer species, *A. annua*, than the other species. We have reported that the light-responsive elements, W-box, CAAT-box, 5′-UTR py-rich stretch, TATA-box sequence and tandem repeat sequences have been identified as important factors in the increased expression of *ADS* gene.

[Ranjbar M., Naghavi M. R. and Alizadeh H. 2014 Comparative analysis of *ADS* gene promoter in seven *Artemisia* species. *J. Genet*. **93**, 767–774]

Introduction

Based on the documented reports, World Health Organization has estimated that there were 216 million cases of malaria in 2010 resulting in 655,000 deaths, equivalent to roughly 2000 deaths everyday (Graham *et al*. 2010; WHO 2010). Strains of *Plasmodium falciparum* displayed high levels of resistance to almost all commonly used antimalarial drugs, especially chloroquine, quinine, amodiaquine and sulphadoxine-pyrimethamine, therefore this threat has presented a major barrier to successful disease management in malaria endemic areas (Kokwaro 2009).

In 1967, Chinese researchers had began to identify new antimalarial drugs from plants that were used in traditional Chinese medicine and succeeded in isolating artemisinin from *A. annua* that was able to kill *Plasmodium berghei* in mice (Klayman 1985; Hien and White 1993). Numerous studies have confirmed that artemisinin and its derivatives show a remarkable activity against the previously used antimalarial drugs resistant and sensitive strains of *P. falciparum*. WHO recommends artemisinin-based combination therapies (ACTs) for uncomplicated *falciparum* malaria (Nosten *et al.* 2007). Artemisinin is also used as a potential therapeutic agent against other parasitic, viral and cancer diseases (Xiao 2005; Lai and Singh 2006; Romero *et al.* 2006). However, the amount of artemisinin in *A. annua* is very low and this limits its massive commercialization (Jing *et al.* 2009). Therefore, researchers have adopted different methods to increase artemisinin production, such as treatment with hormones and elicitors, genetic engineering and identifying new sources of artemisinin (Chen *et al.* 2000; Jing *et al.* 2009; Mannan *et al.* 2010).

Previous results have shown that two genes in the artemisinin pathway, *ADS* and *DBR2*, have a significant positive correlation with the amount of artemisinin in *Artemisia* species (Pu *et al.* 2009; Lei *et al.* 2011). Promoter is the region of a genomic sequence, which is responsible for the generation of RNA (Pedersen *et al.* 1999). The key *cis*-acting regulatory region of the promoter controls the transcription of adjacent coding regions into mRNA and they are recognized by transcription factors that are key regulators of gene expression (Liang *et al.* 2002; Smale and Kadonaga 2003; Hahn 2004). Features of the promoter can be detected by the intrinsic relations between promoters and their corresponding genes, and this is useful for understanding gene transcription (Xie *et al.* 2006). Therefore, promoter identification and isolation is a fundamental and an important step in gene annotation. In this study, we have isolated, cloned and sequenced the ADS promoter in seven *Artemisia* species and then these sequences were compared on the basis of

*For correspondence. E-mail: mnaghavi@ut.ac.ir.

Keywords. *Artemisia*; ADS promoter; *cis*-elements; artemisinin.

presence, number and location of regulatory elements. Based on a comparative analysis and previously available information, we have identified a number of regulatory elements that might have a key role in enhancing *ADS* gene expression.

Materials and methods

Seeds of seven *Artemisia* species (*A. annua, A. spicigera, A. absinthium, A. sieberi, A. diffusa, A. scoparia* and *A. vulgaris*) were obtained from the *Artemisia* collection of the Iranian Biological Resource Center (IBRC, Tehran, Iran). Plants were grown in a growth chamber at 23°C under a 16 h photoperiod and light intensity of 5000 lux. Genomic DNA was extracted from fresh young leaves of seven *Artemisia* species using the CTAB method (Pirttila *et al.* 2001). DNA quality and quantity were determined by 1% agarose gel electrophoresis and its concentration was adjusted to 70 ng/μL.

Sequences of *ADS* gene promoter from *A. annua* (GenBank accession no. Dy448297, Dy448296, Ay528931 and Fj613423) were retrieved from GenBank and aligned with Clustal W for the identification of conserved regions. Based on the conserved regions, two pairs of primers (ADS1 and ADS2) were designed for the isolation of ADS promoter from seven *Artemisia* species (table 1) and the primers validity were confirmed by Oligo Analyzer (Treemu Kuulasmaa, Finland).

Amplified PCR reactions were performed in a final volume of 20 μL containing 140 ng/μL of the genomic DNA, 10 pmol of forward and reverse primers (ADS1 and ADS2), 1 U of Ex *Taq* DNA polymerase (Takara, Shiga, Japan, TPs (2.5 mM each dNTP) and 2 μL of 10× PCR buffer. The temperature was programmed according to a touchdown protocol, beginning with 4 min at 94°C, followed by 9 cycles (94°C for 1 min, 60°C for 40 s (reduced by 1°C for each cycle), 72°C for 90 s); 24 standard cycles (94°C for 1 min, 52°C for 40 s, 72°C for 90 s) and a final elongation step (72°C for 10 min). The amplified products were separated by 1.2% agarose gel electrophoresis and purified from the gel by Glass milk method (Geneclean Spink kit, Bio 101, Vista). The recovered product was cloned into the pBluescript vector for sequencing. The plasmid DNA was isolated and the cloned fragments in pBluescript vector were confirmed by reamplification and restriction enzymes of *Hin*dIII and *Eco*RI. DNA sequences were obtained from the automated fluorescent sequencing system (performed by Seqlab Biotech, Germany. http://www.seqlab.de/). The accuracy of the obtained

Table 1. Nucleotide sequences of the primers in this study.

Name	Primer sequence
ADS1-F	5′-CTCGTTTTAGTCGGCTTTCC-3′
ADS1-R	5′-TGGGCCTTTCTAAACACTCC-3′
ADS2-F	5′-GGAGTGTTTAGAAAGGCCCA-3′
ADS2-R	5′-CAATGGGGCGAATAGGTT-3′

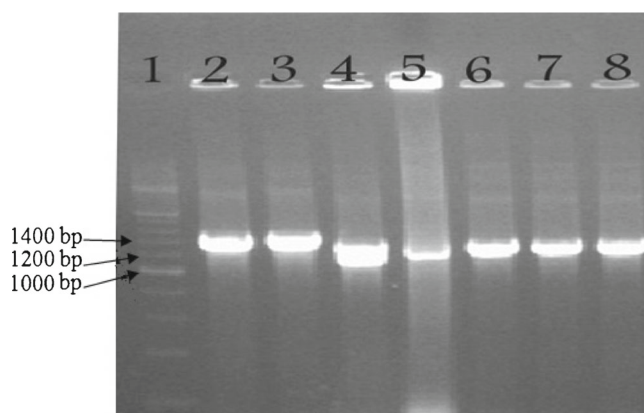

Figure 1. DNA products isolated by ADS1 primer set (PCR products: 2, *A. annua*; 3, *A. scoparia*; 4, *A.absinthium*; 5, *A. diffusa*; 6, *A. sieberi*; 7, *A. spicigera*; 8, *A. vulgaris*; 1, DNA marker).

promoter sequences was confirmed with GenBank *Artemisia* database using the BLASTn sequence alignment program (http://blast.nlm.nih.gov/Blast.cgi).

To identify the regulatory regions important for ADS promoter activity, putative *cis*-elements of the promoter were predicted by using the softwares PlantCARE, PlantPAN and Softberry (Lescot *et al.* 2002; Chang *et al.* 2008; Wang *et al.* 2011). In addition, cluster analysis of *ADS* gene promoter sequences from seven *Artemisia* species were constructed using Phylogeny. fr software (http://www.phylogeny.fr/) (Dereeper *et al.* 2008).

Results

Fragments of 1410–1529 bp were isolated by PCR using ADS1 primer set in seven *Artemisia* species (figure 1). Similarly, ADS2 set amplified fragments of 950–1050 bp were also isolated in the studied species (figure 2). Merging the sequences resulted from above PCRs let to the following sequences: *A. annua* 2522 bp (GenBank accession no. JX409873), *A. absinthium* 2512 bp (GenBank accession no. KC315910), *A. diffusa* 2409 bp (GenBank accession no.

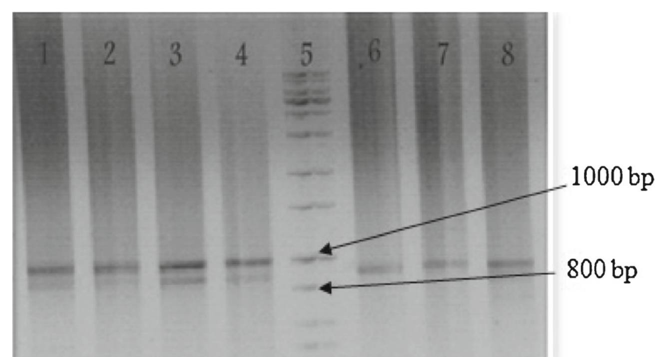

Figure 2. DNA products isolated by ADS2 primer set (PCR products: 1, *A. annua*; 2, *A. scoparia*; 3, *A.absinthium*; 4, *A. diffusa*; 6, *A. sieberi*; 7, *A. spicigera*; 8, *A. vulgaris*; 5, DNA marker).

Table 2. Types and number of regulatory elements of ADS promoter in seven *Artemisia* species.

Function	Site	Number of regulatory element						
		A. annua	*A. absinthium*	*A. diffusa*	*A. scoparia*	*A. sieberi*	*A. spicigera*	*A. vulgaris*
Light responsive element	AE-box, Box4, BoxI, G-box,GAG-motif, MNF1, GT1-motif, chs-CAM1a, MRE, ACE, ATCT-motif, SPI, TCT-motif, ATCC-motif, GATA-motif, CATT-motif, I-box, as-2-box, Gap-box, LAMP-element, CG-motif	27	25	15	21	20	15	13
Cis-acting element conferring high transcription levels	5-UTR py-rich stretch	2	1	0	1	0	0	0
Elicitation wounding and pathogen responsiveness binds WRKY type transcription factor	GCC-motif, W-Box	2	2	0	2	1	1	0
MyB-binding site involved in drought-inducibility	MBS	1	1	1	2	2	3	2
cis-acting element involved in heat stress responsiveness	HSE	1	0	0	2	2	3	3
Cis-acting element involved in low-temperature responsiveness	LTR	0	0	2	0	1	1	2
Cis-acting element involved in defense and stress responsiveness	TC-rich repeat	1	0	0	2	0	1	0
Regulatory element for endosperm expression	Skn-1-motif, GCN4-motif	15	11	5	7	5	6	7
Gibberellin-responsive element	GARE-motif, P-box, TATC-box	1	2	2	0	0	2	2
Auxin-responsive element	TGA-element, AUXRR-core	1	0	3	2	2	1	6
Ethylene-responsive element	ERE	1	0	0	0	0	1	0
Cis-acting element involved in salisylic acid responsiveness	TCA-element, SARE	1	1	1	3	1	2	1
Cis-acting regulatory CGTCA-motif MeJA-responsiveness	CGTCA-motif	0	1	3	2	1	1	1
Cis-acting element involved in the abscisic acid responsiveness	ABRE	0	0	0	0	0	1	1
Cis-element for induction upon fungal elicitation	BoxE	1	0	0	0	0	0	2

KC315911), *A. scoparia* 2516 bp (GenBank accession no. KC315912), *A. sieberi* 2393 bp (GenBank accession no. KC315913), *A. spicigera* 2423 bp (GenBank accession no. JX867774) and *A. vulgaris* 2435 bp (GenBank accession no. KC315914).

For the identification of putative transcription start site (TSS) of the ADS promoter in seven *Artemisia* species two methods were used. In the first method, the TSS of ADS promoter of seven *Artemisia* species were predicted using the TSSP software (http://linux1.softberry.com/berry.phtml), and in the second method promoter sequences were compared with the ADS promoter that was previously reported by Wang *et al.* (2011).

TSS of ADS promoter of *A. annua, A. spicigera, A. absinthium, A. sieberi, A. diffusa, A. scoparia* and *A. vulgaris* were predicted in 50, 51, 52, 46, 52, 50 and 64 bp upstream of the translation initiation ATG-codon, respectively. A putative TATA-box of ADS promoter of *A. annua* was detected at positions −29 to −24 (TATAAA) upstream of the TSS. Moreover, putative TATA-box location and sequences of other species were as follows: *A. vulgaris* at positions −29 to −24 (TATACA), *A. scoparia* at positions −30 to −25 (TATAAC), *A. absinthium* at positions −29 to −24 (TATAGA), *A. diffusa* at positions −29 to −24 (TATAAC), *A. sieberi* at positions −25 to −20 (TATACT) and *A. spicigera* at positions −30 to −25 (TATACA) upstream of the TSS.

Two putative CAAT boxes were identified inside 150 bp upstream of the TATA-box of the ADS promoter in *A. annua* (AATCAATTG or AATTCAATTGT at positions −77 to −69 and −145 to −137 upstream of the TSS), *A. absinthium* (TTCCAATCG or TTTCAATTG at positions −140 to −132 and −149 to −141 upstream of the TSS) and *A. scoparia* (ACCCAATTC or AATCAATTG at positions −39 to −30 and −146 to −137 upstream of the TSS). Further, putative CAAT-box was identified in *A. sieberi* (ATTCAATCT at positions −177 to −169 upstream of the TSS). However, we could not find any CAAT-box inside 150 bp upstream of the TATA-box of ADS promoter in *A. diffusa*, *A. vulgaris* and *A. spicigera*. In addition to the TATA-box and CAT-boxes, several other regulatory elements such as light responsive elements, gibberellin responsive elements, auxin responsive elements were found. Besides, some *cis*-acting elements involved in high transcription levels, ethylene-responsive, heat stress responsiveness, defense and stress responsiveness, salicylic acid and abscisic acid responsiveness, induction upon fungal elicitation, endosperm expression, MeJA-responsiveness, low-temperature responsiveness, elicitation, wounding and pathogen responsiveness were also found in ADS promoters of all or some of the seven *Artemisia* species (table 2).

A tandem repeat sequence with copy number of two and period size of the repeat of 68 at positions −168 to −37 upstream of the TSS was identified by PlantPAN in the ADS promoter of *A. annua*. In addition, a tandem repeat sequence with copy number of 5.6 and period size of the repeat of 17 at positions −2408 to −2313 upstream of the TSS was identified in the ADS promoter of *A. absinthium*.

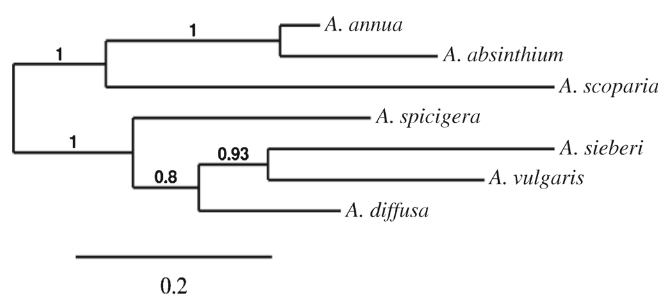

Figure 3. Cluster analysis reveals the relationship between the *ADS* gene promoter sequences in the seven species of *Artemisia.*

The cluster analysis for ADS promoter of seven *Artemisia* species categorized the species into two clusters (figure 3). In the first cluster, the ADS promoter of *A. annua*, *A. absinthium* and *A. scoparia* were included. While, the second cluster composed of *A. spicigera*, *A. sieberi*, *A. vulgaris* and *A. diffusa*.

Discussion

ADS gene is one of the most important genes for artemisinin production. Different studies have suggested that the biosynthesis of artemisinin is regulated by the transcription factors, which bind to the *cis*-acting elements on their promoters (Ptashne and Gann 1997; Hong *et al.* 2009). Therefore, we hypothesized that the promoter of ADS was responsible for the *ADS* gene expression discrepancy in the studied species. To screen for elements that are significantly involved in the higher *ADS* gene expression, comparative analysis was performed on the ADS promoter sequences of seven *Artemisia* species. The comparative analysis results support the concept that some of the *cis*-elements identified in this study were unique to the high artemisinin producer species (*A. annua*) or were different in terms of the number from the other species (table 2). These elements include W-box, CAAT-box, 5′-UTR py-rich stretch, light responsive elements and TATA-box sequences (figure 4).

In addition to the elements mentioned in this study, tandem repeat sequences have also been identified as an important factor to increase the expression of *ADS* gene (figure 4). TATA-box is a *cis*-regulatory element with the conventional consensus sequence 5′-TATAAA-3′ that is found in the core promoter region of most genes transcribed by the eukaryotic RNA polymerase II. This *cis*-regulatory element seems to be important to accurately determine the position at which the transcription is initiated. Our results showed that the conventional TATA-box has been identified only in *A. annua*. Several studies have investigated the effect of mutations in the conventional TATA-box on gene expression and promoter disorder (Dia *et al.* 1996; Fang *et al.* 1989; Stewart and Stargell 2001). To determine the functional

Figure 4. (*continues*)

architecture of the TATA-box sequence of a rice phenylalanine ammonia-lyase (PAL) promoter, Zhu *et al.* (1995) have analysed the transcription of PAL promoter-GUS templates by whole-cell extracts of rice cell suspension cultures. Substitution of the conventional TATA-box sequence with GCGGGTT, TCGTTAA and TATGGAA altered the structure of the minor groove to prevent binding of the *trans* factor and inactivated the minimal promoter. The mechanistic differences of the two sequences: TATAAA (conventional) and CATAAA (nonconventional) have been investigated by Stewart and Stargell (2001) and the functional activity of two sequences indicated that the stability of the TFIIA-TBP complex was strongly dependent on the sequence of the core promoter element. The single mutation in the prototype 13 bp

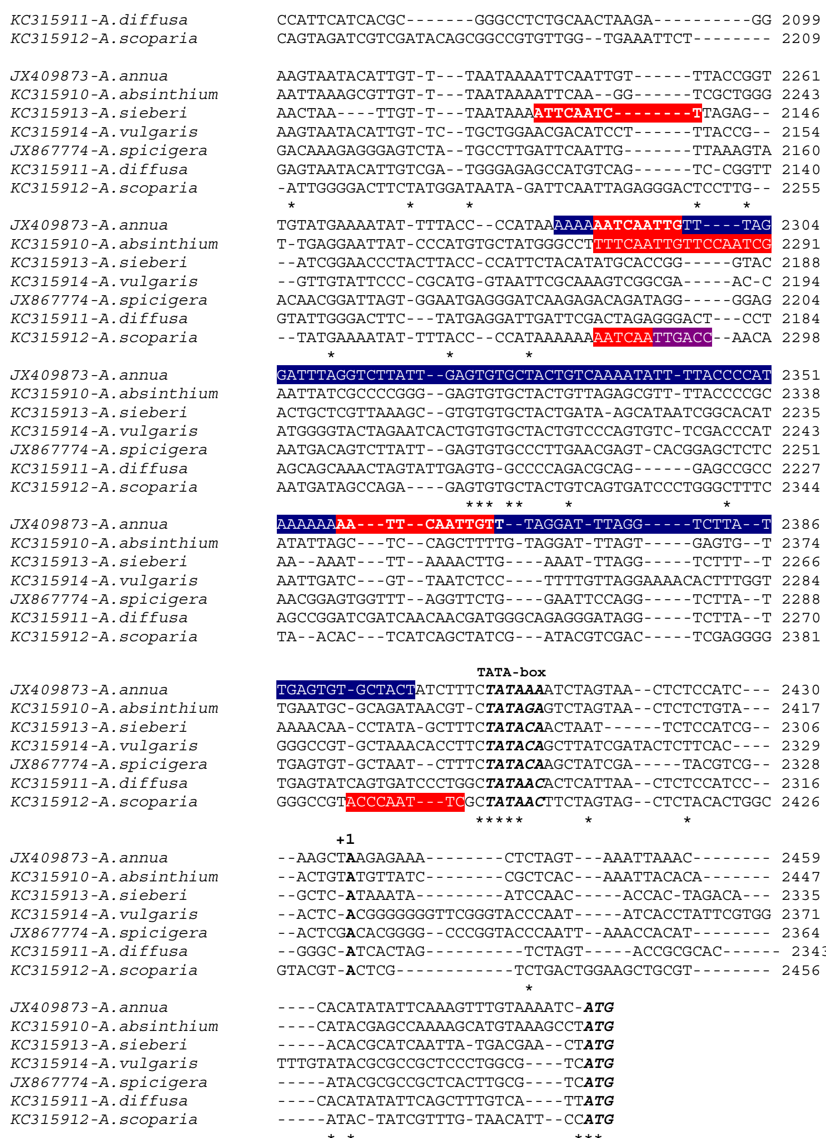

Figure 4. Sequence analysis of ADS promoter in seven *Artemisia* species. The TATA-box and TSS are shown in bold italics. The TSS is marked as +1 and CAAT-box, 5′UTR py-rich stretch, W-box and tandem repeat sequence are shown in red, dark yellow, violet and dark blue, respectively. The v that are conserved in all seven species are indicated by asterisks (*) and dots represents similar sequences in *Artemisia* species between different distances.

TATA-box sequence, TCACTATATATAG, demonstrated that the TATA-box sequence in the basal promoter contributed to determine the light dependent gene expression in plants (Kiran *et al.* 2006). Based on the previously mentioned issues, nonconventional TATA-box sequence has been introduced as one of the possible reasons for low *ADS* gene expression in the studied species by the structural change in minor groove or weak stability of the TFIIA–TBP–DNA complex.

The CAAT-box is found in the promoter region of many eukaryotic genes. Specifically, the CAAT-box is located about 150 bases upstream of the TATA-box. It binds the transcription factors and thereby stabilizes the adjacent preinitiation complex for easier binding of RNA polymerase. Therefore, it is a key element in regulating the frequency of transcription of the gene. Our results showed that the CAAT-box was detected only in *A. annua* and *A. absinthium*. Further, the importance and role of CAAT-box is understood by investigating the previous works (Fang *et al.* 1989; Dia *et al.* 1996; Dai *et al.* 1999). The deletion analysis of the CaMV 35s promoter showed that the 90 to −46 region (three CAAT-box) had little transcriptional activity when fused upstream of the 35s TATA-box promoter but it played an accessory role by increasing the transcriptional activity of its upstream regulatory sequences (Fang *et al.* 1989). A series of 5′ deletion analyses of the rpL34 promoter of tobacco indicated

that deletion of the promoter up to −104 (CAAT-box region) abolished the promoter activity (Dia *et al.* 1996). In addition, deletion analysis and point mutation in the CAAT-box of the nose promoter significantly reduced the promoter response in transgenic tobacco plants and calli to wound, H_2O_2, methyl jasmonate and 2,4-D 9 (Dai *et al.* 1999).

Table 2 shows a large number of regulatory elements involved in response to light, and also the highest number of light responsive elements were observed in the ADS promoter region of *A. annua*. The previous studies have suggested that light was one of the most important environmental factors affecting plant growth and formation of plant products including both primary and secondary metabolites development (Dewdney *et al.* 1993; Korth *et al.* 2000; Lopez-Ochoa *et al.* 2007). Dewdney *et al.* (1993) have examined the effects of different light spectra on the expression of GapA and GapB in dark-adaptation of *Arabidopsis thaliana* and their results showed that the mRNA level of *GapA* and *GapB* genes were significantly increased by blue and white light than red light. Light-responsive *cis*-elements (LREs) are required for light-controlled transcriptional activity. Further, the spacing and combination of LRE had a significant effect on the gene expression (Gilmartin and Chua 1990; Korth *et al.* 2000; Lopez-Ochoa *et al.* 2007). To identify the potential *cis*-acting sequence elements involved in light and organ-specific regulation of the *GapA* gene in *A. thaliana*, deletion analysis was performed by Conley *et al.* (1994) and the obtained results indicated that the −359 to −110 bp region of the *GapA* gene was necessary for light responsiveness. Liu *et al.* (2002) studied the effects of light irradiation on growth and production of artermisinin in hairy root cultures of *A. annua* and their results demonstrated that the artemisinin content of hairy roots increased with the increase in light intensity up to 3000 lux of light irradiation using an ordinary fluorescent lamp. Cryptochrome1 (CRY1) was a blue light receptor of *A. thaliana* and its overexpression increased the accumulation of secondary metabolites. Subsequent studies demonstrated that the overexpression of cryptochrome1 gene (*CRY1*) derived from *A. thliana* in *A. annua* led to increased accumulation of both artemisinin and anthocyanin. Further, artemisinin accumulation was associated with the increased expression of *FDS*, *CYP 71AV1* and *ADS* genes (Hong *et al.* 2009).

In addition, the 5-pyrimidine-rich stretch is a *cis*-acting element conferring high transcription levels. This element was predicted in *A. annua* (at positions −59 to −68 and −1207 to −1216), *A. absinthium* (at positions −1194 to −1202) and *A. scoparia* (at positions −1208 to −1217). This element was also found at the regulatory region of *ACT2* gene in *Arabidopsis* and also *HMGR* gene in tomato. Mutagenesis and deletion analysis confirmed that this element played a regulatory role in *Arabidopsis* and tomato (Daraselia *et al.* 1996; An and Meagher 2010). We propose that 5-pyrimidine-rich stretch plays an important role in regulating the expression of *ADS* gene.

In this study, two W-box elements were found in *A. annua*, *A. absinthium* and *A. scoparia*. WRKY transcription factors were identified as transcriptional regulatory factors which bind to the W-box element and play an important role in response to many kinds of biotic and abiotic stresses. Previous studies have demonstrated that the expression levels of most of the genes involved in the artemisinin biosynthesis pathway (HMGR, ADS, CYP71AV1 and DBR2) were significantly increased with the transient expression of AaWRKY1 cDNA in leaves of *A. annua* (Ma *et al.* 2009).

Cluster analysis revealed a good homology in *ADS* gene promoter sequences of *A. annua* and *A. absinthium* (figure 3). In addition, these two species also showed many common regulatory elements (table 2). The higher amount of artemisinin production may be due to the higher activity of *ADS* gene (Pu *et al.* 2009; Lei *et al.* 2011), therefore, more studies need to be carried out to determine the activity of *ADS* gene of *A. absinthium* in artemisinin biosynthesis.

Acknowledgements

Authors would like to acknowledge the Iran National Foundation (INSF) for the financial support of this work, through a grant no. 90002271. We also thank the Iranian Biological Resource Center (IBRC) for providing seeds.

References

An Y. Q. C. and Meagher R. B. 2010 Strong expression and conserved regulation of ACT2 in *Arabidopsis thaliana* and *Physcomitrella patens*. *Plant Mol. Biol. Rep.* **28**, 481–490.

Chang W. C., Lee T. Y., Huang H. D., Huang H. Y. and Pan R. L. 2008 PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics* **9**, 561.

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

Conley T. R., Park S. C., Kwon H. B., Peng H. P. and Shih M. C. 1994 Characterization of cis-acting elements in light regulation of the nuclear gene encoding the A subunit of chloroplast isozymes of glyceraldehyde-3-phosphate dehydrogenase from *Arabidopsis thaliana*. *Mol. Cell. Biol.* **14**, 2525–2533.

Dai Z., An K., Edward G. E. and An G. 1999 Functional role of CAAT Box element of the nopaline synthase (nos) promoter. *J. Plant Biol.* **42**, 181–185.

Daraselia N. D., Tarchevskaya S. and Narita J. O. 1996 Thekpromoter for tomato 3-hydroxy-3-methylglutaryl coenzyme a reductase gene 2 has unusual regulatory elements that direct high-leve1 expression. *Plant Physiol.* **11**, 727–733.

Dereeper A., Guignon V., Blanc G., Audic S., Chevenet F., Dufayard J. F. *et al.* 2008 *Phylogeny.fr*: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **1**, 36.

Dewdney J., Conley T. R., Shih M. C. and Goodman H. M. 1993 Effects of blue and red light on expression of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase of Arabidopsis thaliana. *Plant Physiol.* **103**, 1115–1121.

Dia Z., Gao J., An K., Lee J. M., Edwards G. E. and An G. 1996 Promoter elements controlling developmental and environmental

regulation of a tobacco ribosomal protein gene L34. *Plant Mol. Biol.* **32**, 1055–1065.

Fang R. X., Nagy F., Jivasubramanian A. M. and Chua N. H. 1989 Multiple cis regulatory elements for maximal expression of the cauliflower mosaic virus 35S Promoter in transgenic plants. *Plant Cell* **1**, 141–150.

Gilmartin P. M. and Chua N. H. 1990 Spacing between GT-1 binding sites within a light-responsive Element 1s critical for transcriptional activity. *Plant Cell* **2**, 447–455.

Graham I. A., Besser K., Blumer S., Branigan C. A., Czechowski T., Elias L. *et al.* 2010 The genetic map of Artemisia annua L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science* **327**, 328–331.

Hahn S. 2004 Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* **11**, 394–403.

Hien T. T. and White N. J. 1993 Qinghaosu. *Lancet* **341**, 603–608.

Hong G. J., Hu W. L., Li J. X., Chen X. Y. and Wang L. J. 2009 Increased accumulation of artemisinin and anthocyanins in *Artemisia annua* expressing the Arabidopsis blue light receptor CRY1. *Plant Mol. Biol. Rep.* **27**, 334–341.

Jing F., Zhang L., Li M., Tang Y., Wang Y., Wang Y. *et al.* 2009 Abscisic acid (ABA) treatment increases artemisinin content in *Artemisia annua* by enhancing the expression of genes in artemisinin biosynthetic pathway. *Biologia* **64**, 319–323.

Kiran K., Ansari S. A., Srivastava R., Lodhi N., Chaturvedi C. P., Sawant S. V. *et al.* 2006 The TATA-Box sequence in the basal promoter contributes to determining light-dependent gene expression in plants. *Plant Physiol.* **142**, 364–376.

Klayman D. L. 1985 Qinghaosu (artemisinin): an antimalarial drug from China. *Science* **228**, 1049–1055.

Kokwaro G. 2009 Ongoing challenges in the management of malaria. *Malaria J.* **8**, S2.

Korth K. L., Jaggard D. A. and Dixon R. A. 2000 Developmental and light regulated post translational control of 3-hydroxy-3-methylglutaryl-CoA reductase levels in potato. *Plant J.* **23**, 507–516.

Lai H. and Singh N. P. 2006 Oral artemisinin prevents and delays the development of 7,12-dimethylbenz[a]anthracene (DMBA)-induced breast cancer in the rat. *Cancer Lett.* **231**, 43–48.

Lei C., Ma D., Pu G., Qiu X., Du Z., Wang H. *et al.* 2011 Foliar application of chitosan activates artemisinin biosynthesis in Artemisia annua L. *Ind. Crop Prod.* **33**, 176–182.

Lescot M., Déhais P., Thijs G., Marchal K., Moreau Y., Van de Peer Y. *et al.* 2002 Plant CARE: A database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**, 325–327.

Liang P. H., Ko T. P. and Wang A. H. 2002 Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* **269**, 3339–3354.

Liu C., Guo C., Wang Y. and Ouyang F. 2002 Effect of light irradiation on hairy root growth and artemisinin biosynthesis of *Artemisia annua* L. *Process Biochem.* **38**, 581–585.

Lopez-Ochoa L., Acevedo-Hernandez G., Martinez-Hernandez A., Arguello-Astorga G. and Herrera-Estrella L. 2007 Structural relationships between diverse cis-acting elements are critical for the functional properties of a rbcS minimal light regulatory unit. *J. Exp. Bot.* **58**, 4397–4406.

Ma D., Pu G., Lei C., Ma L., Wang H., Guo Y. *et al.* 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.

Mannan A., Ahmed I., Arshad W., Asim M F, Qureshi R A, Hussain I. and Mirza B. 2010 Survey of artemisinin production by diverse *Artemisia* species in northern Pakistan. *Malaria J.* **9**, 310.

Nosten F., Nicholas J. and White N. J. 2007 Artemisinin-based combination treatment of Falciparum malaria. *Am. J. Trop. Med. Hyg.* **77**, 181–192.

Pedersen A. G., Baldi P., Chauvin Y. and Brunak S. 1999 The biology of eukaryotic promoter prediction-a review. *Comput. Chem.* **23**, 191–207.

Pirttila A. M., Hirsikorpi M., Kamarainen T., Jaakola L. and Hohtola A. 2001 DNA isolation methods for medicinal and aromatic plants. *Plant Mol. Biol. Rep.* **19**, 273–273.

Ptashne M. and Gann A. 1997 Trascriptional activation by recruitment. *Nature* **386**, 569–577.

Pu G. B., Ma D. M., Chen J. L., Ma L. Q., Wang H., Li G. F. *et al.* 2009 Salicylic acid activates artemisinin biosynthesisin *Artemisia annua* L. *Plant Cell Rep.* **28**, 1127–1135.

Romero M., Serrano M. A., Vallejo M., Effeth T., Alvarez M. and Main J. J. 2006 Antiviral effect of artemisinin from *Artemisia annua* against a model member of the Flaviviridae family, the Bovine Viral Diarrhoea Virus (BVDV). *Planta Med.* **72**, 1169–1174.

Smale S. T. and Kadonaga J. T. 2003 The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **72**, 449–479.

Stewart J. J. and Stargell L. A. 2001 The stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element. *J. Biol. Chem.* **276**, 30078–30084.

Wang H., Olofsson L., Lundgren A. and Brodelius P. E. 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.

World Health Organization 2010 World Health Organization progress towards the health-related millennium development goals. *Fact Sheet*, 290.

Xiao S. H. 2005 Development of antischistosomal drugs in China, with particular consideration to praziquantel and the artemisinins. *Acta Trop.* **96**, 153–167.

Xie X., Wu S., Lam K. M. and Yan H. 2006 PromoterExplorer: an effective promoter identification method based on the AdaBoost algorithm. *Bioinformatics* **22**, 2722–2728.

Zhu Q., Dabi T. and Lamb C. 1995 TATA box and initiator functions in the accurate transcription of a plant minimal promoter in vitro. *The Plant Cell* **7**, 1681–1689.

Received 3 December 2013, in revised form 9 July 2014; accepted 22 July 2014
Unedited version published online: 1 August 2014
Final version published online: 23 December 2014