BMC Genomics

RESEARCH

Open Access

A transcription factor of SHI family AaSHI1 activates artemisinin biosynthesis genes in *Artemisia annua*

Yinkai Yang[1†], Yongpeng Li[1†], Li Jin[1†], Pengyang Li[1], Qin Zhou[1], Miaomiao Sheng[1], Xiaojing Ma[2], Tsubasa Shoji[3], Xiaolong Hao[1*] and Guoyin Kai[1*]

Abstract

Background Transcription factors (TFs) of plant-specific SHORT INTERNODES (SHI) family play a significant role in regulating development and metabolism in plants. In *Artemisia annua*, various TFs from different families have been discovered to regulate the accumulation of artemisinin. However, specific members of the SHI family in *A. annua* (AaSHIs) have not been identified to regulate the biosynthesis of artemisinin.

Results We found five *AaSHI* genes (*AaSHI1* to *AaSHI5*) in the *A. annua* genome. The expression levels of *AaSHI1*, *AaSHI2*, *AaSHI3* and *AaSHI4* genes were higher in trichomes and young leaves, also induced by light and decreased when the plants were subjected to dark treatment. The expression pattern of these four *AaSHI* genes was consistent with the expression pattern of four structural genes of artemisinin biosynthesis and their specific regulatory factors. Dual-luciferase reporter assays, yeast one-hybrid assays, and transient transformation in *A. annua* provided the evidence that AaSHI1 could directly bind to the promoters of structural genes *AaADS* and *AaCYP71AV1*, and positively regulate their expressions. This study has presented candidate genes, with AaSHI1 in particular, that can be considered for the metabolic engineering of artemisinin biosynthesis in *A. annua*.

Conclusions Overall, a genome-wide analysis of the AaSHI TF family of *A. annua* was conducted. Five *AaSHIs* were identified in *A. annua* genome. Among the identified AaSHIs, AaSHI1 was found to be localized to the nucleus and activate the expression of structural genes of artemisinin biosynthesis including *AaADS* and *AaCYP71AV1*. These results indicated that AaSHI1 had positive roles in modulating artemisinin biosynthesis, providing candidate genes for obtaining high-quality new *A. annua* germplasms.

Keywords *Artemisia annua*, Artemisinin biosynthesis, Transcriptional regulation, SHI family, Transcription factor

†Yinkai Yang, Yongpeng Li and Li Jin contributed equally to this work.

*Correspondence:
Xiaolong Hao
xiaolong19890217@126.com
Guoyin Kai
guoyinkai1@126.com
[1] Zhejiang Provincial TCM Key Laboratory of Chinese Medicine Resource Innovation and Transformation, Zhejiang International Science and Technology Cooperation Base for Active Ingredients of Medicinal and Edible Plants and Health, Jinhua Academy, School of Pharmaceutical Sciences, Academy of Chinese Medical Sciences, Zhejiang Chinese Medical University, Hangzhou 310053, China
[2] State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-Di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China
[3] Institute of Natural Medicine, University of Toyama, Toyama 930-0194, Japan

Background

Malaria is a parasitic disease caused by *Plasmodium* infection and results in over 200 million cases globally every year [1]. Artemisinin, a sesquiterpene lactone that contains a peroxy bridge, has demonstrated remarkable effectiveness in the treatment of malignant malaria. Chinese scientist Youyou Tu was awarded the Nobel Prize in Physiology or Medicine in 2015 for her discovery of the anti-malarial property of artemisinin. Although artemisinic acid, the precursor of artemisinin, can be synthesized in yeast by synthetic biology technology, its production cost remains high and cannot meet the increasingly market demand [2]. Currently, the primary source of the artemisinin is still the Compositae family plant *Artemisia annua*. In *A. annua*, artemisinin is mainly synthesized and stored in glandular secretory trichomes (GSTs), which are specialized ten-cell morphological structure found in young leaves and flower buds. As a typical sesquiterpene compound, a series of metabolic and regulatory genes in artemisinin biosynthetic pathway have been thoroughly characterized [3, 4]. Within GSTs, farnesyl pyrophosphate (FPP), a common precursor for sesquiterpene synthesis, is converted to dihydroartemisinic acid (DHAA) through the catalytic action of four enzymes that are specifically localized in GSTs (AaADS, AaCYP71AV1, AaDBR2 and AaALDH1) [5] (Fig. 1). In the sub-epidermal space of GSTs, DHAA can be turned into artemisinin [6] (Fig. 1).

Currently, various metabolic engineering strategies have been demonstrated to increase the content of artemisinin in *A. annua*, including enhancing the metabolic flux in artemisinin biosynthetic pathway, using transcription factors (TFs), applying exogenous factors and so on [7]. Among them, TFs could simultaneously regulate several structural genes of artemisinin biosynthesis. It has been reported that many TFs from different families are involved in light-mediated regulation of artemisinin biosynthesis. AaHY5 from bZIP family was a core TF induced by light and positively regulates artemisinin biosynthesis by controlling the activity of AaGSW1 or activates AaGSW1 by modulating the expression of AaWRKY9 [8, 9]. Meanwhile, AaHY5 could regulate AaWRKY14 to activate *AaCYP71AV1* expression [10]. A R2R3-MYB TF, AaMYB15 was induced by dark and inhibits artemisinin biosynthesis by directly binding to the *AaORA* promoter [11]. WRKY-family AaGSW1 and AP2/ERF-family AaORA, mentioned above, are two TFs specifically expressed in the GSTs, and promote artemisinin biosynthesis [12, 13].

Our analysis predicted presence of SHI binding in the promoter regions of the structural genes of the artemisinin pathway. Therefore, in the present study, we have carried out genome-wide analysis of SHI gene family. SHI

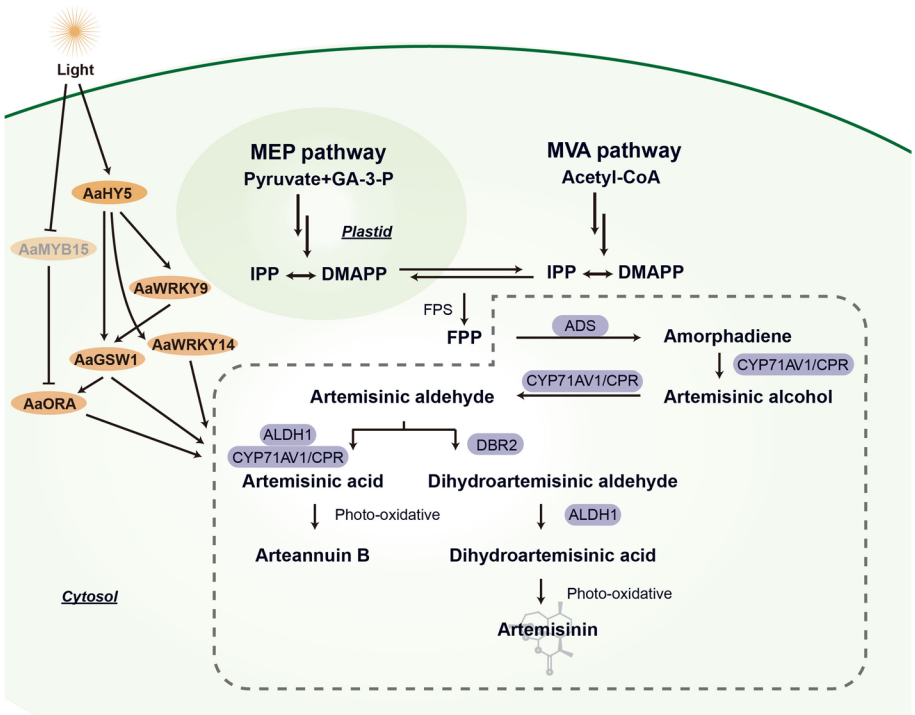

Fig. 1 The artemisinin biosynthetic pathway and its regulatory network in light conditions

family TFs are unique to plants and exhibit functional diversity, encompassing the regulation of growth, development, metabolism, and response to stresses [14]. SHI family proteins are characterized by the presence of two functional domains: a circular zinc finger domain at the N-terminal and an IGGH domain at the C-terminal [15]. The circular zinc finger domain consists of two fingers, which adapt a transverse palm arrangement, similar to the DNA-binding domain. The IGGH domain contains certain acidic residues that play a role in mediating both homologous and heterogeneous dimerization among the SHI proteins [15]. There are ten SHI family proteins in *Arabidopsis thaliana*, of which AtSRS8 is reported to be a pseudogene [16]. STYLISH1, which is related to the development of shoot apical meristems, can directly bind to the YUCCA4 promoter to regulate auxin synthesis [15]. SRS5 negatively regulates lateral root formation through inhibiting the expression of *LBD16* and *LBD29* genes [17]. In addition, light-induced SRS5 can also be ubiquitinated and degraded by COP1 protein and positively modulate the photomorphogenesis in seedlings by directly activating target genes expression [18]. Moreover, the maize LRP1 is auxin-responsive and associated with the initiation of lateral and seminal roots in maize [19]. In rice, OsSHI1 regulates tillering and panicle branching depending on the physical interaction with IPA1 [20]. A trichome-specific SHI TF SlEOT1 in tomato has been shown to regulate TPS expression [21]. However, there have been no reports on the AaSHI TFs in *A. annua*, and the regulatory function of AaSHIs in relation to artemisinin biosynthesis remains largely unexplored.

In this study, we identified five *AaSHI* genes in the *A. annua* genome. Expression analysis conducted in various tissues and under different light conditions revealed a positive correlation among *AaSHI1*, *AaSHI2* and *AaSHI4*, the four structural genes of artemisinin biosynthesis, and two GST-specific TF genes, *AaGSW1* and *AaORA*. Yeast one-hybrid assays and dual-luciferase (dual-LUC) reporter assays showed that AaSHI1 could directly activate the expression of *AaADS* and *AaCYP71AV1*. Transient transformation in *A. annua* leaves confirmed that AaSHI1 positively regulated artemisinin biosynthesis. The data presented this study demonstrated that AaSHI1 was a positive regulator of artemisinin biosynthetic pathway. Taken together, the present study identifies candidate transcription factors, which could be used as target for genetic manipulation of artemisinin biosynthesis in *A. annua*.

Results

Identification of AaSHI members in the *A. annua* genome

To obtain AaSHI homologues, nine amino acid sequences of AtSHI family members from *A. thaliana* were used to blast with four sets of haplotype chromosome genomes of *A. annua* (Supplementary Table 1). Multiple sequence alignments and protein integrity alignment analysis found that unctg_3838g01590241 (LQ-9_phase0) was half the length of the unctg_3207g01555681 and lacked the zinc finger domains, while chr5g00243221 (LQ-9_phase1) also lacked the zinc finger domains (Supplementary Figue 1). As a result, five *AaSHI* genes in the HAN1_phase0 genome were selected for subsequent analyses, and named as *AaSHI1* (*chr1g04025621*), *AaSHI2* (*chr3g03131551*), *AaSHI3* (*chr5g03332101*), *AaSHI4* (*chr6g00096641*), *AaSHI5* (*chr6g00640341*), respectively. Then we analyzed the physical and chemical properties of these identified genes. These genes encoded 276–390 amino acids with the molecular weight (MW) ranged from 68.19 kDa to 95.67 kDa, and their isoelectric points ranged from 5.05 to 5.13 (Table 1).

Phylogenetic tree construction, multiple sequence alignment and gene structure analysis

To further investigate the evolutionary relationships and their evolutionary conservation among individual members of the AaSHI family, a phylogenetic tree of 42 SHI members from *A. annua* and 5 other species (*A. thaliana, Zea mays, Vitis vinifera, Solanum lycopersicum, Oryza sativa*) was constructed using MEGAX software (Fig. 2A). Phylogenetic tree showed that these amino acid sequences formed three branches, and the members of AaSHI were distributed in two branches. AaSHI members were most closely related to the dicotyledonous plant *A. thaliana*, followed by *V. vinifera*, then *Z. mays* and *O. sativa*. Amino acid sequence alignment and conserved domain analysis indicated that all five AaSHI proteins contained the circular zinc finger domains and IGGH domains (Fig. 2B). Meanwhile, MEME online software was used to predict conserved structural domains (Fig. 2C, Supplementary Table 3), and five motifs were identified. Among them, motif 2 and motif 4 were distributed in five AaSHI proteins, which just corresponded to the circular zinc finger domain and IGGH domain.

Table 1 Detailed information for five AaSHI members in the *A. annua* genome

NAME	Gene ID	ORF length (bp)	Protein length (aa)	Mw (kDa)	pI
AaSHI1	*chr1g04025621*	1,005	334	83.54	5.08
AaSHI2	*chr3g03131551*	1,041	346	85.71	5.07
AaSHI3	*chr5g03332101*	972	323	79.64	5.11
AaSHI4	*chr6g00096641*	1,173	390	95.67	5.05
AaSHI5	*chr6g00640341*	831	276	68.19	5.13

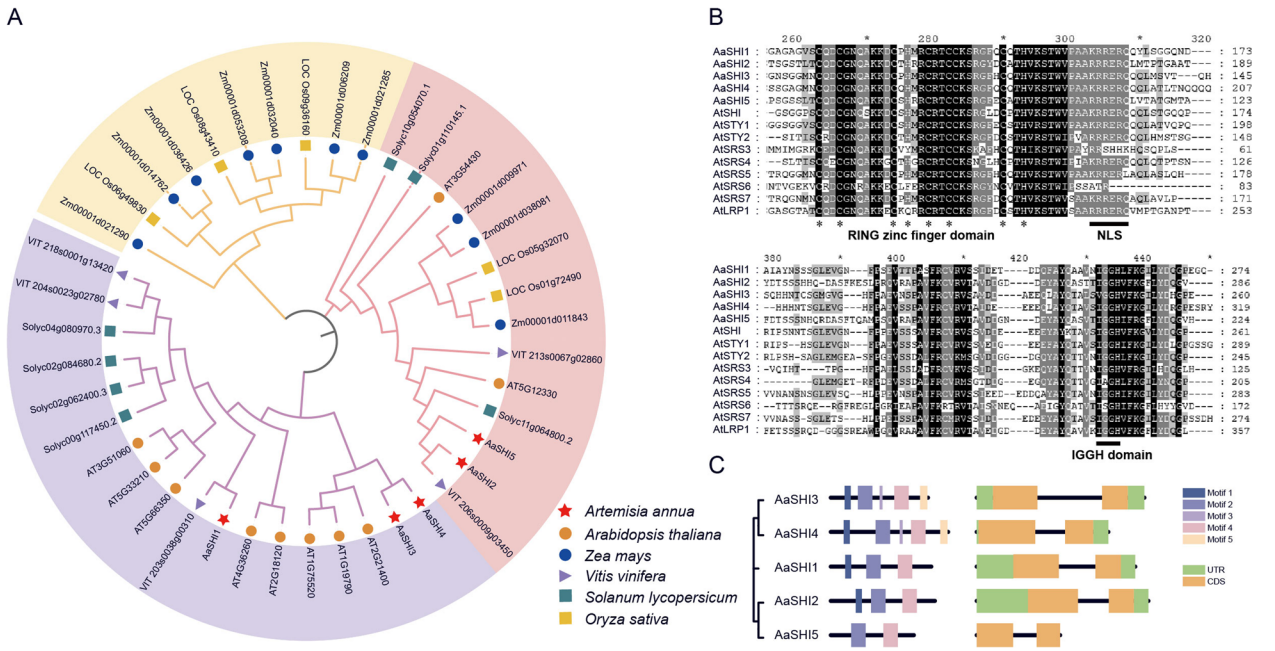

Fig. 2 Phylogenetic analysis and Amino acid sequence alignment of AaSHI members. **A** Using six species (*A. annua, A. thaliana, Z. mays, V. vinifera, S. lycopersicum, O. sativa*) for phylogenetic tree analysis, different species were labeled with different symbols. **B** Protein sequence alignment of AaSHIs and AtSHI/STY family members. The range of nuclear localization signal (NLS) and IGGH domain were marked with horizontal lines and RING zinc finger domain was marked with asterisks. **C** The distribution of conserved motifs and gene structure of the *AaSHI* gene. Five different motifs are marked with different colors. In the gene structure map, the green part was the UTR region, the yellow part was the CDS region, and the intron was the vacant part

Gene structure analysis showed that they all have two exons and one intron, proving that the five *AaSHI* genes share the same structure (Fig. 2C).

Chromosome localization and synteny analysis

Based on the genome information of *A. annua*, the chromosome location analysis of *AaSHI* genes showed that *AaSHI1*, *AaSHI2* and *AaSHI3* were distributed on chr1, chr3 and chr5 respectively, while *AaSHI4* and *AaSHI5* were distributed on chromosome 6, and there were no tandem duplication events among members of *A. annua SHI* gene family during the evolution (Fig. 3A). Synteny analysis detected that two *AaSHI* genes, *AaSHI3* and *AaSHI4*, participated in a segmental duplication event. Large-scale comparative synteny maps of *AaSHI* and *AtSHI*, *SlSHI*, *VvSHI* genes showed that *A. annua* and *V. vinifera* had the highest synteny with eight pairs of genes, but only six pairs of synteny genes with *S. lycopersicum* (Fig. 3B, Supplementary Table 4).

Analysis of *cis*-acting elements of *AaSHI* genes promoter

To understand the biological processes in which AaSHIs may be involved, *cis*-acting element analysis of the promoter region located 3.0 kb upstream of the start codon of the *AaSHIs* was performed via the PlantCARE online website (Supplementary Table 5). As shown in Fig. 4A, the promoter regions of *AaSHIs* mainly contained four major types of *cis*-acting elements: Plant growth and development, phytohormone responsive, light responsive, abiotic and biotic stress. MYB binding sites were the most common in the *AaSHI* genes promoter, followed by MYC binding elements were also more distributed except for *AaSHI5*, which indicated that AaSHIs may act as a potential regulatory target for MYB and MYC TFs to regulate the growth of *A. annua*. Interestingly, the ABRE element and W-box in the *AaSHI5* promoter were significantly more than the other four *AaSHI* genes, suggesting that its function may differ from the others. In addition, some jasmonic acid response elements could also be found in the all *AaSHI* genes promoter. To intuitively illustrate the distribution of *cis*-acting elements, we calculated the number of elements of each type and represented them as a bar chart (Fig. 4B). These data indicated that the AaSHI genes potentially have diverse roles in the phytohormone and environmental response of *A. annua.*

Expression analysis of *AaSHI* genes using RNA-Seq data

Previous studies have provided some RNA-seq databases to analyze the gene expression profile. Different tissues including young leaf (top leaves of plants), old

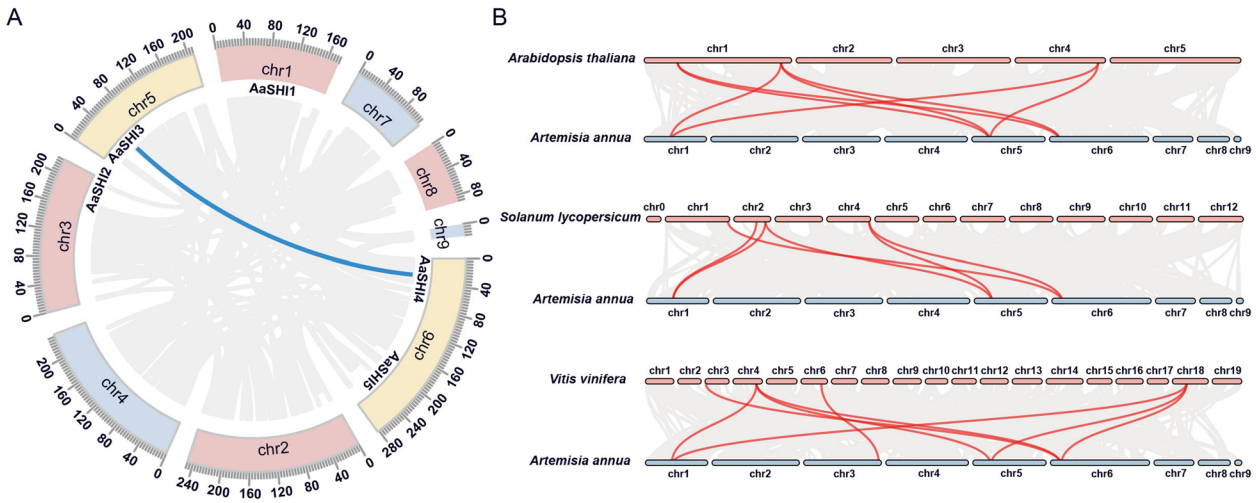

Fig. 3 The chromosomal location and synteny analysis of *AaSHI* genes. **A** Circos diagram illustrated the chromosomal locations of *AaSHI* genes and their synteny. The blue line indicated the presence of replication events between two genes. **B** Synteny analysis of *AaSHI* genes between *A. annua* and *A. thaliana*, *S. lycopersicum* and *V. vinifera*, respectively. The red line represented the synteny of *SHI* genes between two species

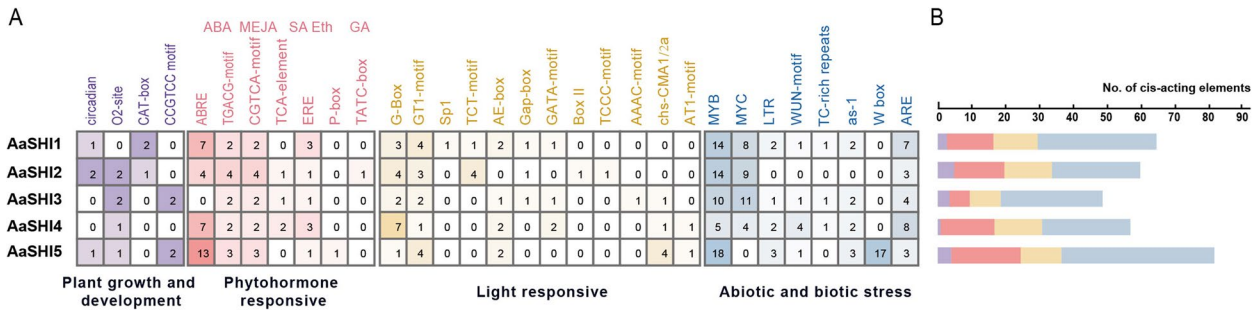

Fig. 4 Promoter *cis*-regulatory elements analysis of the *AaSHI* genes. **A** Classified and statistically analyzed the data exported from plantcare, and visualize it in the form of heat maps. The color depth was consistent with the number of *cis*-acting elements. **B** The sum of *cis*-acting elements of each major type. The color of the column corresponds to the chart A

leaf (bottom leaves of plants), bud, flower, stem, seed, and root from two-week-old *A. annua* have been collected to establish tissues transcriptome [3]. The leaves have been collected at light and dark treatment to perform transcriptome sequencing [22]. The apical meristematic of *A. annua* was named Leaf 0 and the first leaf below meristem was named Leaf 1, samples have been collected from top to bottom in distinct leaf positions to establish phyllotaxy transcriptome [23]. Evaluate the expression level of AaSHI gene family with four structural genes and GST-specific TFs by TPM (Transcript Per Million), visual mapping was performed after normalization using the TBtools software [24]. The results showed that *AaSHI1*, *AaSHI2*, *AaSHI3* and *AaSHI4* were all specifically expressed in the trichome, which was consistent with the expression pattern of artemisinin, while *AaSHI5* was specifically expressed in the stem (Fig. 5A). Previous studies have shown that DHAA content decreases in phyllotaxy, with very low DHAA concentration in old leaves [25]. The results of gene expression in distinct leaf positions indicated that except for *AaCYP71AV1*, other genes showed significant leaf order, that is, they were highly expressed in tender leaves and low or even not expressed in old leaves (Fig. 5B). Meanwhile, *AaSHI* genes with structural genes and GST-specific TFs were all induced by light, *AaSHI5* was not detected due to its low transcript level (Fig. 5C).

The expression pattern of AaSHIs

The expressions of *AaSHI* genes were evaluated by qRT-PCR. As shown in Fig. 6A-D, *AaSHI1-4* were highly expressed in the young leaves but weakly expressed in root, which is in line with the structural genes. *AaSHI5* showed the highest expression in the stem (Fig. 6E). In addition, the expression levels of artemisinin biosynthetic pathway genes were highest in the tender leaves

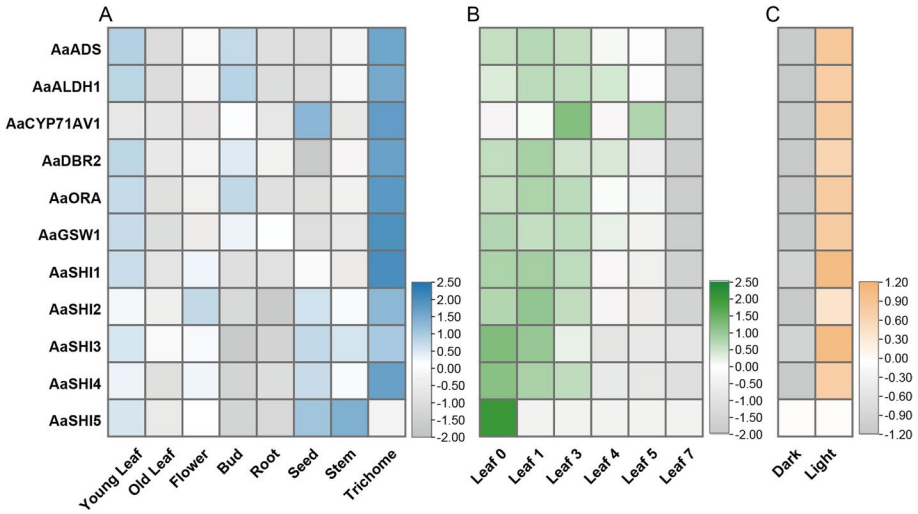

Fig. 5 Expression profile of *AaSHI* genes, structural genes of artemisinin biosynthesis and GST-specific TFs by RNA-seq databases. The transcriptional expression levels of eleven genes in seven tissues including root, stem, old leaf, young leaf, bud, flower and trichome (**A**), leaf 0, leaf 1, leaf 3, leaf 4, leaf 5 and leaf 7 in *A. annua* plants from top to bottom (**B**), and treated with darkness and light (**C**). The TPM of each gene was subjected to log2 transformation, gray indicated lower transcript abundance, while corresponding colors indicated higher transcript abundance

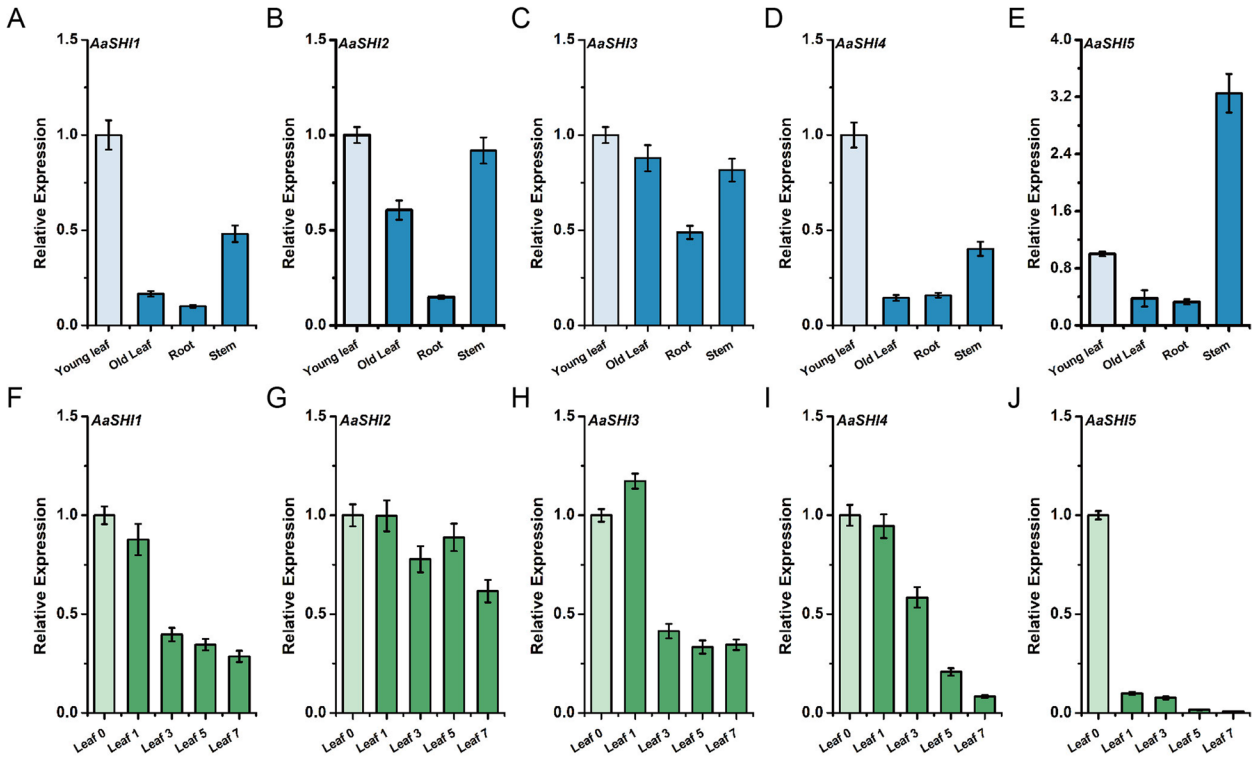

Fig. 6 qRT-PCR validation of five *AaSHI* genes. **A**-**E** The relative expression levels of AaSHI genes in young leaf, old leaf, stem and root were detected by qRT-PCR. The relative expression in young leaf of each *AaSHI* gene was set to 1. **F**-**J** The relative expression levels of AaSHI genes of leaf 0, leaf 1, leaf 3, leaf 5 and leaf 7 in *A. annua* plants from top to bottom were detected by qRT-PCR. The relative expression of each *AaSHI* genein leaf 0 was set to 1

and gradually decreased as leaf age. Similar expression pattern was observed in *AaSHI1*, *AaSHI2*, *AaSHI4*, and *AaSHI5* (Fig. 6F-J). Considering that *AaSHI5* was preferentially expressed in stem rather than young leaves, we speculated that the AaSHI1, AaSHI2 and AaSHI4 are related to the biosynthesis of artemisinin.

Correlation analysis with structural genes of artemisinin biosynthesis

Next, we conducted co-expression analysis of *AaSHIs*, structural genes of artemisinin biosynthesis, and two key regulators of artemisinin biosynthetic pathway *AaGSW1* and *AaORA*, based on their tissue/organ and light-treated transcriptome data. The Pearson coefficient was used to calculate the correlation between the AaSHI TFs and structural genes. Pairs of genes that meet the criteria will be screened and then visualized using Cytoscape software. TPM of *AaSHIs* in GST was indicated by a color gradient. Correlation analysis showed that three genes (*AaSHI1*, *AaSHI2* and *AaSHI4*) had a positive correlation with all four structural genes as well as GST-specific TFs. However, *AaSHI3* and *AaSHI5* were not correlated with these key genes (Fig. 7, Supplementary Table 6). Therefore, we chose AaSHI1, AaSHI2 and AaSHI4 for further functional study.

Subcellular localization of the AaSHI proteins

The subcellular localization prediction was conducted on the Plant mPLoc website, and the all three TFs were predicted to be localized to the nucleus. Subsequently, the C-ternimal of *AaSHIs* were fused with YFP and expressed in *N. benthamiana* leaves to experimentally verify the subcellular localization. As shown in Fig. 8, pHB-AaSHI1/2/4-YFP were detected in the nucleus exclusively, while the control YFP (Yellow Fluorescent Protein) displayed in both nucleus and cytoplasm. This is consistent with website predictions and their potential functional localization as a TF.

AaSHIs transactivated the expression of structural genes of artemisinin biosynthesis

To verify the activation effect of AaSHIs on four structural genes of artemisinin biosynthesis, we used tobacco leaves for transient transformation. The reporter vector was obtained by inserting the successfully cloned promoter into the *pGreenII 0800-LUC* vector with homologous recombination method and

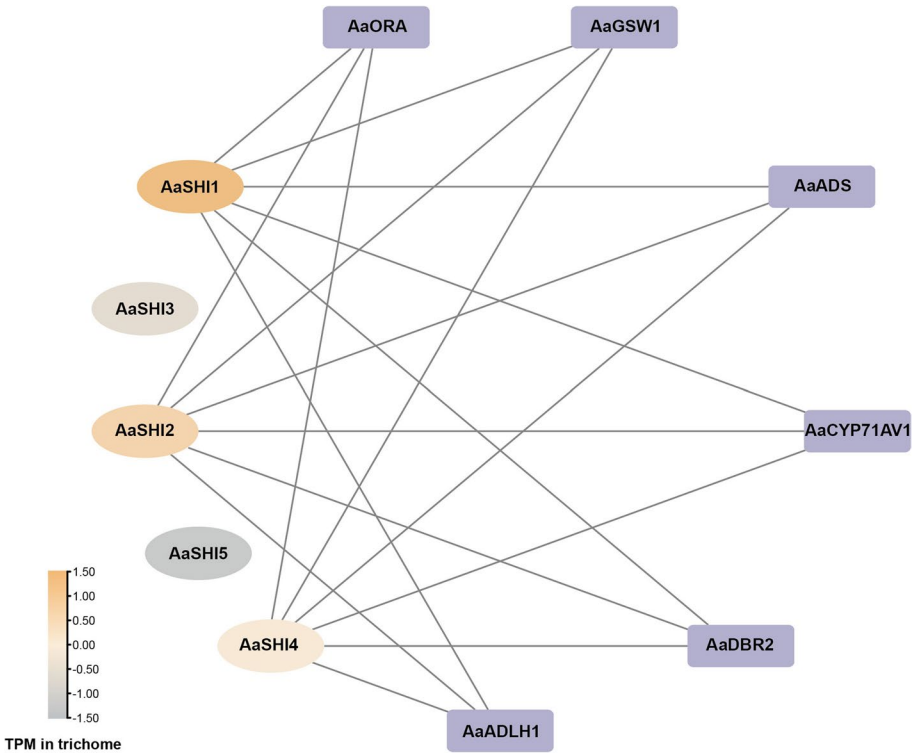

Fig. 7 Correlation analysis of *AaSHI* gene family with structural genes of artemisinin biosynthesis and GST-specific TFs. Five *AaSHIs* were labeled with ellipses, six genes that have been reported to regulate artemisinin biosynthesis were labeled with rectangles. Different color gradients were assigned to *AaSHIs* based on their expression levels in GST. Regulatory relationships were all represented by solid line due to their positive regulation

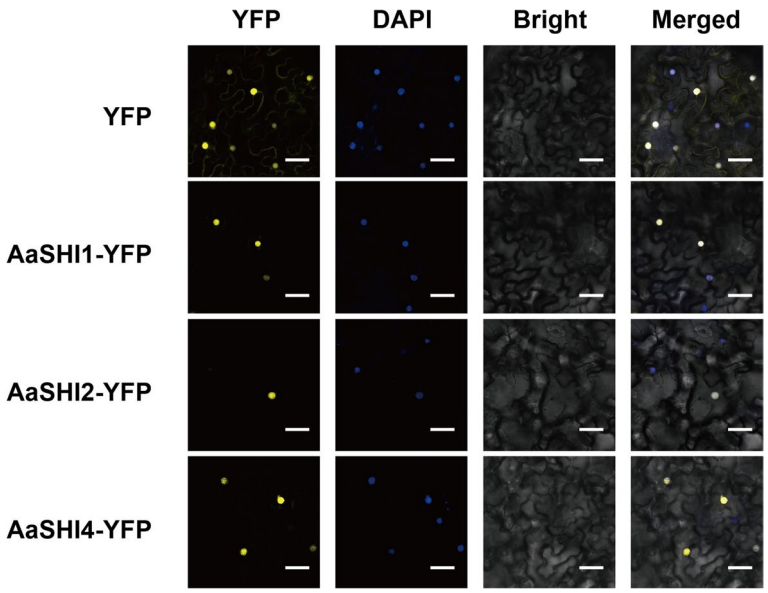

Fig. 8 The subcellular localization of AaSHI1, AaSHI2, AaSHI4 in *N. benthamiana* leaves. YFP was excited at 488 nm. Determined the nucleus by DAPI staining and excited at 405 nm. YFP is used as a negative control. Scale bars: 20 μm

transferred into *A. tumefaciens* GV3101 (pSoup) for dual-LUC assay. Meanwhile, *AaSHIs* were inserted in *pHB* vector driven by the *35S* promoter (Fig. 9A). The results indicated that AaSHI1 has significant activation effect on the promoters of *AaADS* and *AaCYP71AV1*, while AaSHI2 could solely activate the expression of *AaADS* gene (Fig. 9B-E). In contrast, AaSHI4 had no activation effect on these artemisinin pathway genes. Furthermore, we analyzed the SHI-binding sites in the 3,000 bp upstream promoter regions of these four structural genes. It was reported that the binding sites for SHI proteins were ACTCTAC, ACTCCAT, ACT CAAC and ACTCTAA [15, 17, 18, 20], so we deduced that the possible binding site for SHI TFs was ACTCnAn. The analysis results showed that there were six potential binding sites on the promoters of *AaADS* and three on *AaCYP71AV1* (Supplementary Fig. 2). Yeast one-hybrid assays were performed to further examine the binding ability. The ORF of *AaSHI1* and *AaSHI2* were inserted into the *pB42AD* effector vector. Each predicted binding site motif along with the four nucleotide sequences on both sides was artificially synthesized into three repeat fragment and inserted into the *pLacZ* reporter vector. As shown in Fig. 9F and G, AaSHI1 could directly bind to *AaADS* and *AaCYP71AV1* promotors while AaSHI2 could only bind to *AaADS* promotor. These data suggested AaSHI1 and AaSHI2 served as direct positive regulators of structural genes of artemisinin biosynthesis.

AaSHI1/AaSHI2 promotes the expression of structural genes of artemisinin biosynthesis

We have previously demonstrated that AaSHI1 and AaSHI2 have transcriptional activation effects on structural genes of artemisinin biosynthesis, and identified corresponding targets by yeast one-hybrid assay. To further verify the function of AaSHIs, transient transformation was conducted to verify their effects on structural genes in vivo. *pHB-AaSHI1/2-YFP* constructs in *Agrobacterium* strain cells GV3101 were injected into the back of the first pair of true leaves of *A. annua* (Fig. 10A). After 48 h of cultivation, samples were taken and qRT-PCR was used to detect the expression level of structural genes. The results showed that *AaSHI1* and *AaSHI2* genes were successfully expressed at high levels in *A. annua* leaves (Fig. 10B, C). And AaSHI1 and AaSHI2 significantly promoted *AaADS* expression in *A. annua*, about 8.6-fold and 3.2-fold of that of the empty control, respectively. AaSHI1 also significantly activated the expression of *AaCYP71AV1*. (Fig. 10D, E). These results indicated that AaSHI1 and AaSHI2 had positive roles in regulating artemisinin biosynthetic pathway.

Discussion

Malaria remains a great threat to global security and caused about 247 million infections and 619,000 deaths worldwide in 2021 (World Malaria Report 2022). Artemisinin and its derivatives which show potent antimalarial activity have been widely used for the treatment

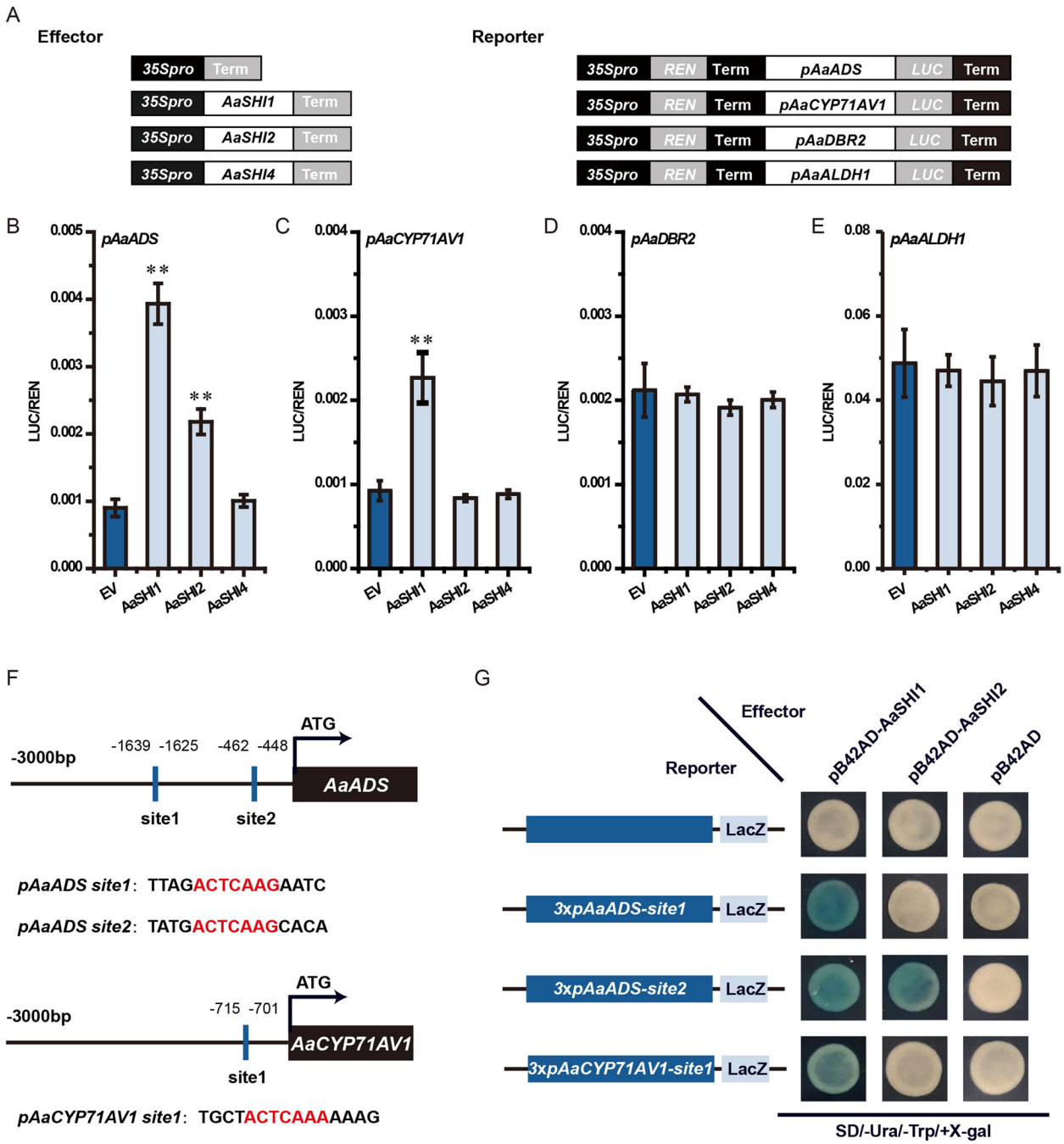

Fig. 9 Dual-LUC assays verified the activation effect of AaSHI proteins on artemisinin biosynthesis. **A** Schematic diagram of vetor construction of and pHB-AaSHI1/2/4, pGREEN0800-promoter-LUC. **B**-**E** The results of dual-LUC assay in tobacco leaf cells. Error bars indicate the mean ± standard deviation (SD). Student's t-test was used to evaluate the significant difference of AaSHIs activation on structural genes. **, $p < 0.01$. **F** Schematic diagram of the binding sites on the promoter and the nucleotide sequence of the binding sites. **G** Yeast one-hybrid assay for interaction between AaSHI1/2 protein with binding site motifs, the triple fragments used were presented in F

of malaria and have significantly reduced its fatality. Artemisinin is originally isolated and purified from *A. annua*, a Chinese medicinal plant. Currently, because of the low artemisinin production by using heterologous systems, such as tobacco and *Physcomitrella patens*, the main source of artemisinin was still the cultivated *A. annua* [26, 27]. It is essential to increase the artemisinin production in *A. annua*, thereby meeting the large-scale global

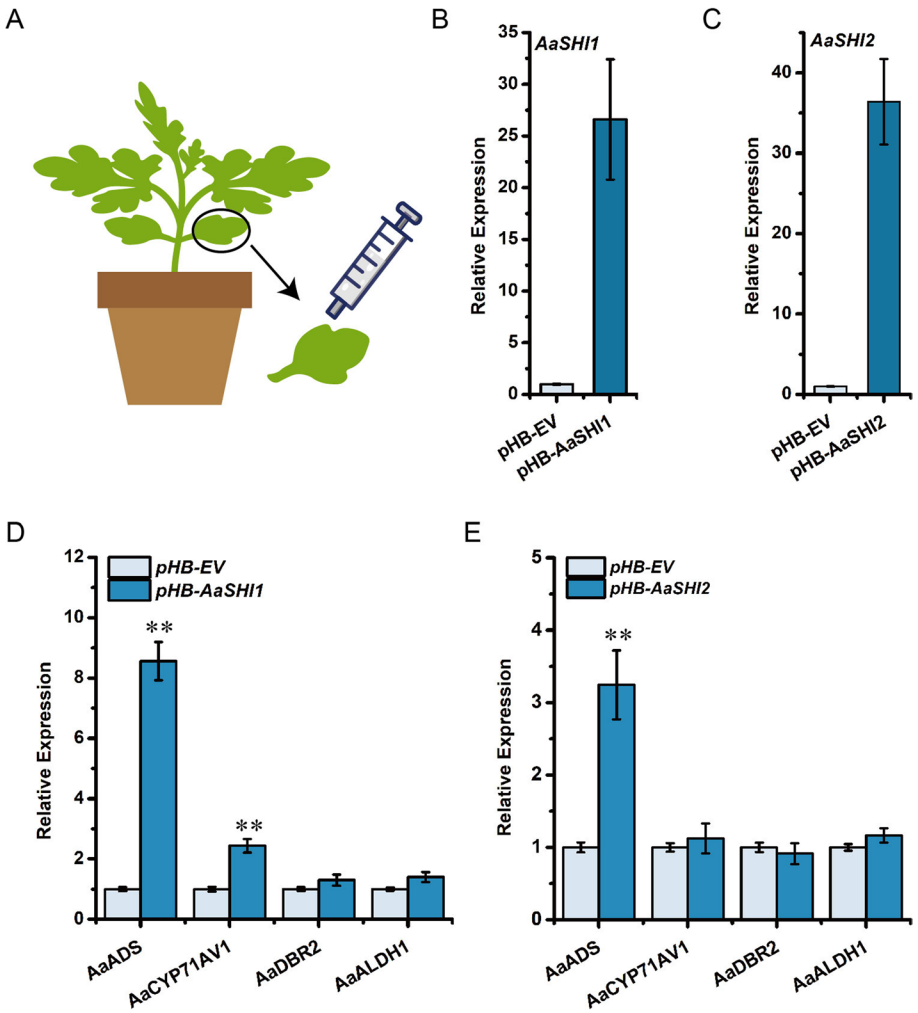

Fig. 10 Transient expression assays indicated that AaSHI1 promoted the expression level of structural genes of artemisinin biosynthesis. **A** Schematic diagram for the first pair of true leaves of *A. annua* injected with *Agrobacterium* strain. **B**, **C** The relative expression of *AaSHI* and *AaSHI2* after transient transformation in *A. annua* leaves. **D**, **E** The relative expression of four structural genes in *A. annua* leaves with transient transformation of *AaSHI1* and *AaSHI2* genes. Error bars indicate the mean ± SD. Student's t-test was used to evaluate the significant difference of AaSHI1and AaSHI2 activation on structural genes. **, $p < 0.01$

demand. Increasing the GST density and enhancing the expression levels of structural genes of artemisinin biosynthesis are the most effective strategies to improve the artemisinin content in *A. annua* [28]. Recently, many TFs have been confirmed to have important roles in regulating GST formation and artemisinin biosynthesis. For example, MYB TF family members such as AaMIXTA1, AaMYB5, AaMYB16, AaMYB17, AaTLR1 and AaTLR2 are related to GST initiation, while AaMYB15 and AaMYB108 are involved in the regulation of artemisinin biosynthesis [11, 29–33]. Similarly, several members from WRKY TF family have proved to be involved in GST formation or artemisinin biosynthesis. AaGSW1, AaWRKY9, AaWRKY14 and AaWRKY17 were found to modulate artemisinin biosynthesis while AaGSW2 acts as a key regulator of GST initiation [9, 10, 12, 34, 35]. In addition, two MADS-box members AaSEP1 and AaSEP4 were reported to regulate GST formation and artemisinin biosynthesis, respectively [36, 37]. These findings demonstrated there were functional differences among the members of one TF family. Given the advantages of *A. annua* genome, WRKY, bHLH and B-box TF family members were genome-wide characterized [38–40]. However, the regulatory roles of SHI TF family in artemisinin production in *A. annua* remain largely unknown.

The SHI TF family is an ancient plant gene family, in which AtLRP1 is the first cloned SHI/STY gene, named LRP (Lateral Root Primordia) due to its activation and

expression during the formation of lateral root primordia [41]. It serves as a molecular marker gene for studying the early stages of lateral root primordia formation and development. At present, there have been some studies on the function of the SHI TF family in *A. thaliana*, and in other species such as *Z. mays*, *O. sativa*, *Hordeum vulgare*, and *Glycine max* [19, 20, 42, 43]. It is proved that SHI TF family members can participate in the formation of *A. thaliana* roots and the development of organs such as leaves and flowers as well as regulate hormone biosynthesis and signal transduction [44, 45], while the regulation of secondary metabolism by the SHI family has not been reported. To investigate whether AaSHI members could regulate the synthesis of artemisinin, a genome-wide analysis of AaSHI in *A. annua* was conducted. After a series of sequence analysis, five TFs of the AaSHI family were ultimately screened and named AaSHI1-5 according to their sequence numbers (Table 1). The number of SHI family TFs varies from 0 to 28 in different species [14, 45], indicating that AaSHI was a small family within this range. The phylogenetic tree analysis divided AaSHIs into two groups and was recently related to AtSHIs and VvSHIs, while the remaining group was the SHI proteins of monocot *Z. mays*, *O. sativa*. The circular zinc finger domain and IGGH domain are typical conserved domains in the SHI family [46, 47], both of which are present in AaSHIs, demonstrating the conservation of the *AaSHI* gene in *A. annua* (Fig. 2). Chromosome mapping and intra-species synteny analysis showed that five SHI genes were distributed on four chromosomes, and no tandem repeats were found, but there was a fragment replication event (Fig. 3). Intraspecies synteny analysis found that there was one fragment replication event in the AaSHI family. Unlike some *AtSRSs* and *OsSHI1*, which are mainly expressed in roots and flowers and almost not expressed in leaves [16, 48], *AaSHIs* expressed in leaves, and the expression level of *AaSHI2* in leaves was higher than that in flowers and roots (Fig. 5A). Moreover, the expression of *AaSHIs* in leaves also exhibits leaf order, with the highest expression level in youngest leaves and decreasing with leaf order (Fig. 5B). We speculated that AaSHI may have functional differences from AtSHI of *A. thaliana*, or may have other potential regulatory functions. Analysis of *cis*-acting elements on the *AaSHIs* promoter revealed its possible involvement in plant growth and development as well as multiple signaling pathways (Fig. 4). The co-expression network revealed the correlation between *AaSHIs* with four structural genes of artemisinin biosynthesis and GST-specific TFs, which indicated that AaSHI1, AaSHI2, AaSHI4 were positively correlated with these genes and could be used as candidate genes for further functional studies (Fig. 7). Subcellular localization experiments revealed that all these three candidate genes localize in the nucleus, consistent with their function as TFs (Fig. 8). Dual-LUC assays (Fig. 9B-E) and yeast one-hybrid assays (Fig. 9F, G) indicated AaSHI1 and AaSHI2 had direct transcriptional activation effects on structural gene *AaADS*. In addition, AaSHI1 could directly activate the expression of *AaCYP71AV1*. Accordingly, transient expression assays in *A. annua* further demonstrated that AaSHI1 could significantly upregulate the expression of *AaADS* and *AaCYP71AV1* and AaSHI2 upregulated the expression of *AaADS* in vivo (Fig. 10). Taken together, we concluded that AaSHI1 was the most functional positive regulator of artemisinin biosynthesis by activating structural genes *AaADS* and *AaCYP71AV1*.

Numerous studies have demonstrated that TFs increase the artemisinin yield via activating the expression level of four structural genes of artemisinin biosynthesis. For example, AaWRKY1 could activate *AaADS* and *AaCYP71AV1* expression and enhance the artemisinin production [49]. AaMYC2 has the ability to bind to the G-box motifs on the promoters of *AaCYP71AV1* and *AaDBR2*, and overexpression of AaMYC2 leads to an increase in artemisinin production [50]. AaTCP15, as a TF capable of responding to both JA and ABA signals, can directly bind and activate the promoter of *AaDBR2* and *AaALDH1*. Meanwhile, AaORA, a positive regulatory factor, can interact with and activate the transcriptional activity of AaTCP15 by forming an AaORA-AaTCP15 module to synergistically activate *AaDBR2* [51]. In this study, we demonstrated the strong activation effects of AaSHI1 on *AaADS* and *AaCYP71AV1*.

Materials and methods

Plant materials

The *Artemisia annua* L. seeds used in this study was the variety "Huhao 1" [3], which was obtained from Chongqing, China and has been screened for multiple generations in Shanghai. Plant seeds were provided and authorized by Professor Tang Kexuan and his team from Shanghai Jiao Tong University. *A. annua* seeds were used in qRT-PCR analysis and transient transformation experiments. Tobacco (*Nicotiana benthamiana*) plants used for subcellular localization analysis and dual-LUC assay were stored in our laboratory (Zhejiang Provincial TCM Key Laboratory of Chinese Medicine Resource Innovation and Transformation, Zhejiang Chinese Medical University). All plant materials were planted at a constant temperature of 26 degrees Celsius, with 10 h of light and 14 h of darkness.

Identification of SHI-family TFs of *A. annua*

Four sets of haplotype chromosome genomes of *A. annua* were downloaded [52]. Nine sequences of SHI

gene family in *A. thaliana* were used as baits to align four sets of haplotype chromosomal genomes of *A. annua* using the blastp program of TBtools software [24], and four sets of AaSHI gene family sequences of *A. annua* were obtained. Five *AaSHI* gene sequences in the HAN 1_phase0 genome were finally selected for subsequent analyses after performing the protein alignment in the NCBI database. The physiochemical properties of the AaSHI proteins were analyzed using the online tool ExPASy program [53].

Phylogenetic tree construction, protein alignment and gene structure analysis

A phylogenetic tree of the SHI-family proteins from *A. annua*, *A. thaliana*, *O. sativa*, *S. lycopersicum*, *V. vinifera* and *Z. mays*, was constructed using the Neighbor-Joining method. The Bootstrap value was set to 1,000, and the other parameters were maintained to their default values. The sequences used in the phylogenetic tree were downloaded from Phytozome13. The alignment of AtSHIs with AaSHIs was performed and visualized by ClustalW and Genedoc software, respectively. The MEME online tool was employed to predict the conserved motifs of AaSHIs with the number of the motifs set to five. The conserved motifs and structure of the genes were visualized using TBtools.

Chromosomal localization, intraspecies and interspecies synteny analysis

Chromosomal locations and replication events of *AaSHI* genes as well as self-alignment of the whole genome sequence were visualized by TBtools. The whole genome sequence file and gene structure annotation file were obtained from EnsemblPlants database. The synonymous relationship between *AaSHIs* and *SHIs* from other three species (*A. thaliana*, *V. vinifera*, *S. lycopersicum*) was conducted with MCScanX program and TBtools software to perform synteny analysis [54].

Analysis of *cis*-regulatory elements

The *cis*-regulatory elements were predicted in promoter sequences (3,000 bp upstream of first ATG) of *AaSHI* genes and four structural genes of artemisinin biosynthesis using the PlantCARE online website [55]. Subsequently, the obtained *cis*-acting elements on the promoter of *AaSHIs* were classified and sorted. Due to the limited coverage of SHI TFs at present, online prediction cannot obtain SHI binding sites (SBS) on the promoter four structural genes. Therefore, we manually searched based on existing research.

Correlation analysis of *AaSHI* family genes with structural and regulatory genes of artemisinin biosynthesis

We retrieved the transcriptome data of five *AaSHI* genes, structural genes of artemisinin biosynthesis and two GST-specific TF genes. The correlation between five *AaSHI* family genes with four structural genes and two GST-specific artemisinin regulatory genes was calculated by Pearson coefficient. The correlation coefficients R and *p*-value would be obtained. The positive and negative correlation coefficients R represent the promoting and inhibitory effects of AaSHI TFs on structural and regulatory genes, respectively. The screening threshold was set as follows: the absolute value of correlation coefficient $R > 0.8$ and the p-value < 0.05 [56, 57]. Finally, the network diagram of the co-expression of five *AaSHI* family genes, four structural genes and two GST-specific artemisinin regulatory genes was visualized by Cytoscape_v3.7.2 software. Different presentation effects can be achieved by modifying parameters in the software.

qRT-PCR analysis

Total RNAs of *A. annua* were isolated from various tissues of the plants using the RNApure Plant Kit (Tiangen, China). cDNA synthesis was carried out using the HiScript III 1st Strand cDNA Synthesis Kit with gDNA Wiper (Vazyme, China). qPCR amplification was performed as previously reported [58]. *Actin* (EU531837.1) was used as an internal control. Each sample has three biological replicates.

Subcellular localization analysis

The Plant-mPLoc website was used to predict subcellular localization of AaSHI proteins. To further analyze the subcellular localization, high-Fidelity DNA polymerase KOD-Plus (Toyobo, Japan) was used to clone *AaSHI* genes. And then the full-length coding sequences of candidate *AaSHIs* (*AaSHI1*, *AaSHI2* and *AaSHI4*) were inserted into the plant expression *pHB-YFP* vector. The constructed plasmids *pHB-AaSHIs-YFP* and *pHB-YFP* (empty vector) were transformed into the *Agrobacterium tumefaciens* strain GV3101 to transient infect the 5-week-old *N. benthamiana* leaves. The DAPI signal and the YFP signal were observed using confocal microscopy after culturing *N. benthamiana* plants in dark for 24 h and then in light condition for 24 h at a constant temperature of 25 °C [59].

Dual-LUC assay

To generate reporter constructs used in the dual-LUC assays, the promoters of four structural genes were cloned and constructed into pGreenII0800 plasmid. The *pHB-AaSHIs-YFP* constructs were considered as effectors

and *pHB-YFP* construct was considered as control. Effectors and reporters were transformed into GV3101. Reporter strains and effector strains were mixed with ratio of 1:1 and transiently transformed the *N. benthamiana* leaves. The culture condition of tobacco used for dual-LUC assays was the same as those for subcellular localization analysis [60]. After two days, the samples were harvested to analyze the LUC/REN ratio [61]. Three biological replicates were performed with the dual-LUC reporter assay system (Promega, USA).

Yeast one-hybrid assay

The ORF of *AaSHI1* and *AaSHI2* were inserted into the *pB42AD* effector vector. The sequences containing predicted binding site motif along with the four nucleotide sequences on both sides were inserted into the *pLacZ* reporter vector. Effector vector and reporter vector were cotransferred into the yeast EGY48 strain by LiAc mediated method [59]. The positively transformed clones were grown on SD/-Ura-Trp medium with X-gal at 30 °C. The discoloration of yeast plaque was observed after 24 h.

Transient expression in the leaves of *A. annua*

The leaves of 2-week-old *A. annua* seedlings mentioned earlier were used for transient transformation [62]. *Agrobacterium* strain cells GV3101 containing *pHB-AaSHI1-YFP* constructs were injected into the back of the first pair of true leaves. The injected leaves were dried with absorbent paper and covered with a transparent plastic lid to maintain humidity. The seedings were cultivated in the dark for 24 h and then transfer to light conditions for another 24 h. The samples were collected for qRT-PCR.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10683-7.

Additional file 1: Supplementary Table 1. Blast results of four sets of haplotype chromosome genomes of *Artemisia annua*. Supplementary Table 2. Primer used in this study. Supplementary Table 3. Amino acid sequences of five motifs predicted from AaSHI protein sequences. Supplementary Table 4. Synteny analysis of SHI genes between *A. annua* and three other species. Supplementary Table 5. Specific statistical information on the *cis*-acting elements of the *AaSHI* genes. Supplementary Table 6. Correlation analysis with artemisinin structural genes and GST-specific TFs in different tissues of *A. annua*. Supplementary Fig. 1. Multiple sequence alignments of AaSHIs with four sets of haplotype chromosome genomes. Supplementary Fig. 2. Potential binding sites of SHI in the four artemisinin structural genes promoter region.

Acknowledgements

We appreciate the experimental support from the Public Platform of Pharmaceutical Research Center, Academy of Chinese Medical Sciences, Zhejiang Chinese Medical University.

Authors' contributions

XLH and GYK conceived and designed the project. YKY, PYL performed the experiments. YKY, YPL, QZ, MMS and XJM analyzed the data. YKY, YPL, LJ and XLH wrote the manuscript. TS, XLH and GYK revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by National Key Research and Development Program of China (2023YFC3503900), National Natural Science Foundation of China (82003889, 82304651), Zhejiang Provincial Natural Science Foundation of China (LQ21H280004), China-Japan Youth Exchange Program in Science, Technology and Humanities Seminar on "Twinning Short-Term Exchange Project", Key project at central government level: The ability establishment of sustainable use for valuable Chinese medicine resources (2060302), Research Project of Zhejiang Chinese Medical University (2022RCZXZK23, 2024JKZ-KTS08), China Postdoctoral Science Foundation (2022M722851) and Opening Project of Zhejiang Provincial Preponderant and Characteristic Subject of Key University (Traditional Chinese Pharmacology), Zhejiang Chinese Medical University (No. ZYAOXZD2019006).

Availability of data and materials

The datasets generated and analyzed during the current study are included in this article. The sequencing data that support the findings of this study are openly available in the global pharmacopoeia genome database (http://www.gpgenome.com/species/92.). Raw reads for RNA-Seq were downloaded from the NCBI database with accessionnumber SRP129502 (https://www.ncbi.nlm.nih.gov/sra/?term=SRP129502) and SRP092562 (https://www.ncbi.nlm.nih.gov/sra/SRP092562).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 22 February 2024 Accepted: 1 August 2024
Published online: 09 August 2024

References

1. White NJ. Qinghaosu (Artemisinin). The price of success. Science. 2008;320(5874):330–4.
2. Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai A, Main A, Eng D, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature. 2013;496(7446):528.
3. Shen Q, Zhang LD, Liao ZH, Wang SY, Yan TX, Shi P, Liu M, Fu XQ, Pan QF, Wang YL, et al. The Genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis. Mol Plant. 2018;11(6):776–88.
4. Tang KX, Shen Q, Yan TX, Fu XQ. Transgenic approach to increase artemisinin content in *Artemisia annua* L. Plant Cell Re. 2014;33(4):605–15.
5. Brown GD, Sy LK. In vivo transformations of dihydroartemisinic acid in *Artemisia annua* plants. Tetrahedron. 2004;60(5):1139–59.
6. Schramek N, Wang HH, Romisch-Margl W, Keil B, Radykewicz T, Winzenhorlein B, Beerhues L, Bacher A, Rohdich F, Gershenzon J, et al. Artemisinin biosynthesis in growing plants of *Artemisiaannua*. A (CO2)-C-13 study. Phytochemistry. 2010;71(2–3):179–87.
7. Li YP, Yang YK, Li L, Tang KX, Hao XL, Kai GY. Advanced metabolic engineering strategies for increasing artemisinin yield in Artemisia annua L. Hortic Res. 2024;11(2):uhad292.
8. Hao XL, Zhong YJ, Nutzmann HW, Fu XQ, Yan TX, Shen Q, Chen MH, Ma YA, Zhao JY, Osbourn A, Li L, Tang KX. Light-induced artemisinin

biosynthesis is regulated by the bZIP transcription factor AaHY5 in Artemisia annual. Plant Cell Physiol. 2019;60(8):1747–60.
9. Fu XQ, Peng BW, Hassani D, Xie LH, Liu H, Li YP, Chen TT, Liu P, Tang YL, Li L, et al. AaWRKY9 contributes to light- and jasmonate-mediated to regulate the biosynthesis of artemisinin in *Artemisia annua*. New Phytol. 2021;231(5):1858–74.
10. Zhou LM, Huang YZ, Wang Q, Guo DJ. AaHY5 ChIP-seq based on transient expression system reveals the role of AaWRKY14 in artemisinin biosynthetic gene regulation. Plant Physiol Bioch. 2021;168:321–8.
11. Wu ZKY, Li L, Liu H, Yan X, Ma YN, Li YP, Chen TT, Wang C, Xie LH, Hao XL, et al. AaMYB15, an R2R3-MYB TF in *Artemisia annua*, acts as a negative regulator of artemisinin biosynthesis. Plant Sci. 2021;308:110920.
12. Chen MH, Yan TX, Shen Q, Lu X, Pan QF, Huang YR, Tang YL, Fu XQ, Liu M, Jiang WM, et al. GLANDULAR TRICHOME-SPECIFIC WRKY 1 promotes artemisinin biosynthesis in *Artemisia annua*. New Phytol. 2017;214(1):304–16.
13. Lu X, Zhang L, Zhang FY, Jiang WM, Shen Q, Zhang LD, Lv ZY, Wang GF, Tang KX. AaORA, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to Botrytis cinerea. New Phytol. 2013;198(4):1191–202.
14. Fang D, Zhang WM, Ye ZY, Hu F, Cheng XZ, Cao J. The plant specific SHORT INTERNODES/STYLISH (SHI/STY) proteins: Structure and functions. Plant Physiol Bioch. 2023;194:685–95.
15. Eklund DM, Staldal V, Valsecchi I, Cierlik I, Eriksson C, Hiratsu K, Ohme-Takagi M, Sundstrom JF, Thelander M, Ezcurra I, et al. The *Arabidopsis thaliana* STYLISH1 protein acts as a transcriptional activator regulating auxin biosynthesis. Plant Cell. 2010;22(2):349–63.
16. Kuusk S, Sohlberg JJ, Magnus Eklund D, Sundberg E. Functionally redundant SHI family genes regulate Arabidopsis gynoecium development in a dose-dependent manner. Plant J. 2006;47(1):99–111.
17. Yuan TT, Xu HH, Li J, Lu YT. Auxin abolishes SHI-RELATED SEQUENCE5-mediated inhibition of lateral root development in Arabidopsis. New Phytol. 2020;225(1):297–309.
18. Yuan TT, Xu HH, Zhang Q, Zhang LY, Lu YT. The COP1 Target SHI-RELATED SEQUENCE5 Directly Activates Photomorphogenesis-Promoting Genes. Plant Cell. 2018;30(10):2368–82.
19. Zhang YX, von Behrens I, Zimmermann R, Ludwig Y, Hey S, Hochholdinger F. LATERAL ROOT PRIMORDIA 1 of maize acts as a transcriptional activator in auxin signalling downstream of the Aux/IAA gene rootless with undetectable meristem 1. J Exp Bot. 2015;66(13):3855–63.
20. Duan EC, Wang YH, Li XH, Lin QB, Zhang T, Wang YP, Zhou CL, Zhang H, Jiang L, Wang JL, et al. OsSHI1 regulates plant architecture through modulating the transcriptional activity of IPA1 in Rice. Plant Cell. 2019;31(5):1026–42.
21. Spyropoulou EA, Haring MA, Schuurink RC. Expression of Terpenoids 1, a glandular trichome-specific transcription factor from tomato that activates the terpene synthase 5 promoter. Plant Mol Biol. 2014;84(3):345–57.
22. Hao XL, Zhong YJ, Fu XQ, Lv ZY, Shen Q, Yan TX, Shi P, Ma YN, Chen MH, Lv XY, et al. Transcriptome Analysis of Genes Associated with the Artemisinin Biosynthesis by Jasmonic Acid Treatment under the Light in *Artemisia annua*. Front Plant Sci. 2017;8:971.
23 He YL, Fu XQ, Li L, Sun XF, Tang KX, Zhao JY. AaSPL9 affects glandular trichomes initiation by positively regulating expression of AaHD1 in Artemisia annua L. Plant Science. 2022;317:111172.
24. Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13:1194–202.
25. Li YP, Chen TT, Liu H, Qin W, Yan X, Wu-Zhang K, Peng BW, Zhang YJ, Yao XH, Fu XQ, et al. The truncated AaActin1 promoter is a candidate tool for metabolic engineering of artemisinin biosynthesis in Artemisia annua L. J Plant Physiol. 2022;274:153712.
26. Farhi M, Marhevka E, Ben-Ari J, Algamas-Dimantov A, Liang ZB, Zeevi V, Edelbaum O, Spitzer-Rimon B, Abeliovich H, Schwartz B, et al. Generation of the potent anti-malarial drug artemisinin in tobacco. Nat Biotechnol. 2011;29(12):1072–4.
27. Ikram NKBK, Kashkooli AB, Peramuna AV, van der Krol AR, Bouwmeester H, Simonsen HT. Stable production of the antimalarial drug artemisinin in the moss Physcomitrella patens. Front Bioeng Biotech. 2017;5:47.
28. Li YP, Qin W, Liu H, Chen TT, Yan X, He WZ, Peng BW, Shao J, Fu XQ, Li L, Hao XL, Kai GY, Tang KX. Increased artemisinin production by promoting glandular secretory trichome formation and reconstructing artemisinin biosynthetic pathway in *Artemisia annua*. Hortic Res. 2023;10(5):uhad055.
29. Shi P, Fu XQ, Shen Q, Liu M, Pan QF, Tang YL, Jiang WM, Lv ZY, Yan TX, Ma YN, et al. The roles of AaMIXTA1 in regulating the initiation of glandular trichomes and cuticle biosynthesis in *Artemisia annua*. New Phytol. 2018;217(1):261–76.
30. Qin W, Xie LH, Li YP, Liu H, Fu XQ, Chen TT, Hassani D, Li L, Sun XF, Tang KX. An R2R3-MYB Transcription Factor Positively Regulates the Glandular Secretory Trichome Initiation in *Artemisia annua* L. Front Plant Sci. 2021;12:657156.
31. Xie LH, Yan TX, Li L, Chen MH, Hassani D, Li YP, Qin W, Liu H, Chen TT, Fu XQ, et al. An HD-ZIP-MYB complex regulates glandular secretory trichome initiation in *Artemisia annua*. New Phytol. 2021;231(5):2050–64.
32. Liu H, Li L, Fu XQ, Li YP, Chen TT, Qin W, Yan X, Wu ZKY, Xie LH, Kayani SL, et al. AaMYB108 is the core factor integrating light and jasmonic acid signaling to regulate artemisinin biosynthesis in *Artemisia annua*. New Phytol. 2022;237(6):2224–37.
33. Lv ZY, Li JX, Qiu S, Qi F, Su H, Bu QT, Jiang R, Tang KX, Zhang L, Chen WS. The transcription factors TLR1 and TLR2 negatively regulate trichome density and artemisinin levels in *Artemisia annua*. J Integr Plant Biol. 2022;64(6):1212–28.
34. Chen TT, Li YP, Xie LH, Hao XL, Liu H, Qin W, Wang C, Yan X, Wu-Zhang KY, Yao XH, et al. AaWRKY17, a positive regulator of artemisinin biosynthesis, is involved in resistance to Pseudomonas syringae in *Artemisia annua*. Hortic Res. 2021;8(1):217.
35. Xie LH, Yan TX, Li L, Chen MH, Ma YN, Hao XL, Fu XQ, Shen Q, Huang YW, Qin W, et al. The WRKY transcription factor AaGSW2 promotes glandular trichome initiation in *Artemisia annua*. J Exp Bot. 2021;72(5):1691–701.
36. Chen TT, Yao XH, Liu H, Li YP, Qin W, Yan X, Wang XY, Peng BW, Zhang YJ, Shao J, et al. MADS-box gene AaSEP4 promotes artemisinin biosynthesis in *Artemisia annua*. Front Plant Sci. 2022;13:98231.
37. Chen TT, Liu H, Li YP, Yao XH, Qin W, Yan X, Wang XY, Peng BW, Zhang YJ, Shao J, et al. AaSEPALLATA1 integrates jasmonate and light-regulated glandular secretory trichome initiation in *Artemisia annua*. Plant Physiol. 2023;192:1483–97.
38 De Paolis A, Caretto S, Quarta A, Di Sansebastiano GP, Sbrocca I, Mita G, Frugis G. Genome-wide identification of WRKY Genes in *Artemisia annua*: Characterization of a Putative Ortholog of AtWRKY40. Plants-Basel. 2020;9(12):1669.
39. Chang SW, Li Q, Huang BK, Chen WS, Tan HX. Genome-wide identification and characterisation of bHLH transcription factors in *Artemisia annua*. Bmc Plant Biol. 2023;23(1):63.
40. He W, Liu H, Li Y, Wu Z, Xie Y, Yan X, Wang X, Miao Q, Chen T, Rahman SU, et al. Genome-wide characterization of B-box gene family in *Artemisiaannua* L. and its potential role in the regulation of artemisinin biosynthesis. Ind Crop Prod. 2023;199:116736.
41. Smith DL, Fedoroff NV. LRP1, a gene expressed in lateral and adventitious root primordia of arabidopsis. Plant Cell. 1995;7(6):735–45.
42. Yuo T, Yamashita Y, Kanamori H, Matsumoto T, Lundqvist U, Sato K, Ichii M, Jobling SA, Taketa S. A SHORT INTERNODES (SHI) family transcription factor gene regulates awn elongation and pistil morphology in barley. J Exp Bot. 2012;63(14):5223–32.
43. Zhao SP, Song XY, Guo LL, Zhang XZ, Zheng WJ. Genome-wide analysis of the shi-related sequence family and functional Identification of GmSRS18 Involving in Drought and Salt Stresses in Soybean. Int J Mol Sci. 2020;21(5):1810.
44. Estornell LH, Landberg K, Cierlik I, Sundberg E. SHI/STY Genes Affect Pre- and Post-meiotic Anther Processes in Auxin Sensing Domains in Arabidopsis. Front Plant Sci. 2018;9:150.
45 Fang D, Zhang WM, Cheng XZ, Hu F, Ye ZY, Cao J. Molecular evolutionary analysis of the SHI/STY gene family in land plants: a focus on the Brassica species. Front Plant Sci. 2022;13:958964.
46. Fridborg I, Kuusk S, Moritz T, Sundberg E. The Arabidopsis dwarf mutant shi exhibits reduced gibberellin responses conferred by overexpression of a new putative zinc finger protein. Plant Cell. 1999;11(6):1019–32.
47. Fridborg I, Kuusk S, Robertson M, Sundberg E. The Arabidopsis protein SHI represses gibberellin responses in Arabidopsis and barley. Plant physiol. 2001;127(3):937–48.
48. Kim SG, Lee S, Kim YS, Yun DJ, Woo JC, Park CM. Activation tagging of an Arabidopsis SHI-RELATED SEQUENCE gene produces

abnormal anther dehiscence and floral development. Plant Mol Biol. 2010;74(4–5):337–51.
49. Jiang W, Fu X, Pan Q, Tang Y, Shen Q, Lv Z, Yan T, Shi P, Li L, Zhang L, et al. Overexpression of AaWRKY1 Leads to an Enhanced Content of Artemisinin in Artemisia annua. Biomed Res Int. 2016;2016:7314971.
50. Shen Q, Lu X, Yan T, Fu X, Lv Z, Zhang F, Pan Q, Wang G, Sun X, Tang K. The jasmonate-responsive AaMYC2 transcription factor positively regulates artemisinin biosynthesis in *Artemisia annua*. New Phytol. 2016;210(4):1269–81.
51. Ma YN, Xu DB, Yan X, Wu ZK, Kayani SI, Shen Q, Fu XQ, Xie LH, Hao XL, Hassani D, et al. Jasmonate- and abscisic acid-activated AaGSW1-AaTCP15/AaORA transcriptional cascade promotes artemisinin biosynthesis in Artemisia annua. Plant Biotechnol J. 2021;19(7):1412–28.
52. Liao BS, Shen XF, Xiang L, Guo S, Chen SY, Meng Y, Liang Y, Ding DD, Bai JQ, Zhang D, et al. Allele-aware chromosome-level genome assembly of *Artemisia annua* reveals the correlation between ADS expansion and artemisinin yield. Mol Plant. 2022;15(8):1310–28.
53 Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31:3784–8.
54. Wang YP, Tang HB, DeBarry JD, Tan X, Li JP, Wang XY, Lee TH, Jin HZ, Marler B, Guo H, Kissinger JC, Paterson AH. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.
55. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 2002;30(1):325–7.
56. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
57. Wang JY, Li YP, Yang YK, Xiao CY, Ruan QY, Li PY, Zhou Q, Sheng MM, Hao XL, Kai GY. Comprehensive analysis of OpHD-ZIP transcription factors related to the regulation of camptothecin biosynthesis in *Ophiorrhiza pumila*. Int J Biol Macromol. 2023;242(Pt 3):124910.
58. Yang YK, Xiao CY, Cai Y, Wang JY, Ruan QY, Sheng MM, Li L, Tang KX, Kai GY, Hao XL. Jasmonic acid responsive AaJRM1 transcription factor positively regulates the biosynthesis of anti-malarial drug artemisinin in *Artemisia annua*. Ind Crop Prod. 2023;199:116769.
59. Hao XL, Pu ZQ, Cao G, You DW, Zhou Y, Deng CP, Shi M, Nile SH, Wang Y, Zhou W, et al. Tanshinone and salvianolic acid biosynthesis are regulated by SmMYB98 in *Salvia miltiorrhiza* hairy roots. J Adv Res. 2020;23:1–12.
60. Hao X, Xie C, Ruan Q, Zhang X, Wu C, Han B, Qian J, Zhou W, Nutzmann HW, Kai G. The transcription factor OpWRKY2 positively regulates the biosynthesis of the anticancer drug camptothecin in *Ophiorrhiza pumila*. Hortic Res. 2021;8(1):7.
61. Hao X, Wang C, Zhou W, Ruan Q, Xie C, Yang Y, Xiao C, Cai Y, Wang J, Wang Y, et al. OpNAC1 transcription factor regulates the biosynthesis of the anticancer drug camptothecin by targeting loganic acid O-methyltransferase in *Ophiorrhiza pumila*. J Integr Plant Biol. 2023;65(1):133–49.
62. Li YP, Chen TT, Wang W, Liu H, Yan X, Wu-Zhang KY, Qin W, Xie LH, Zhang YJ, Peng BW, et al. A high-efficiency Agrobacterium-mediated transient expression system in the leaves of *Artemisia annua* L. Plant Methods. 2021;17(1):106.

Publisher's Note