Genome-Wide Identification and Co-expression network analysis of *Aux/IAA* Gene Family in *Salvia miltiorrhiza*

Bin Huang[1], Yuxin Qi[1], Xueshuang Huang[1*], Peng Yang[1*]

[1] Hunan Provincial Key Laboratory for Synthetic Biology of Traditional Chinese Medicine, School of Pharmaceutical Sciences, Hunan University of Medicine, Huaihua, Hunan, 418000, China

* Corresponding author: Peng Yang, yangpenghnum@163.com; Xueshuang Huang, xueshuanghuang@126.com

Abstract

The Auxin/Indole-3-Acetic Acid (Aux/IAA) gene family serves as a principal group of genes responsible for modulating plant growth and development through the auxin signaling pathway. Despite the significance of this gene family, the identification and characterization of members within the well-known Chinese medicinal herb *Salvia miltiorrhiza* (*S. miltiorrhiza*) have not been thoroughly investigated. In this study, we employed bioinformatics methods to identify 23 Aux/IAA genes within the genome of *S. miltiorrhiza*. These genes were classified into typical IAA and atypical IAA based on their domain structure. Our analysis of the promoter regions revealed that the expression of these genes is regulated not only by auxins, but also by other hormones and environmental factors. Furthermore, we found that the expression patterns of these genes varied across various tissues of *S. miltiorrhiza*. While our initial hypothesis suggested that the primary function of these genes was the interaction between *SmIAA* and *ARF*, gene co-expression network analysis revealed that they are also influenced by various other transcription factors, such as *WRKY* and *ERF*. The findings establish a sturdy basis for future investigations into the function of the Aux/IAA gene family and exhibit promising prospects for enhancing the genetics of this medicinal flora and its associated species.

Introduction

33 Auxin, the first identified plant hormone, is present ubiquitously in all plants and participates in
34 every stage of plant growth and development (Guilfoyle & Hagen 2007b). Indole-3-acetic acid
35 (IAA) is a common plant auxin (Luo et al. 2018; Wu et al. 2017) and plays a crucial role in many
36 processes of plant growth and development, such as cell division, embryogenesis,
37 morphogenesis, directional response, dormancy extension, apical dominance, and tissue
38 differentiation (Kepinski & Leyser 2005; Leyser et al. 1996; Leyser 2017; Luo et al. 2018;
39 Mironova et al. 2017).

40 The response and signal transduction of plant cells to auxin are primarily mediated through the
41 TIR1/AFB-Aux/IAA-ARF signaling pathway in the nucleus (Guilfoyle & Hagen 2007a). In this
42 signaling pathway, auxin signal transduction is primarily controlled by Auxin/Indole-3-acetic
43 acid (Aux/IAA) and auxin response factor (ARF) (Lavy & Estelle 2016). In an environment
44 with low concentrations of auxin, ARF and Aux/IAA protein form a heterodimer and bind to the
45 TOPLESS (TPL) protein and histone deacetylase (HDACs) to inhibit the response of ARF to
46 auxin (Luo et al. 2018; Parry et al. 2009; Winkler et al. 2017). However, when the concentration
47 of auxin is high in the environment, auxin binds to the TIR1 receptor, and the resulting SCF^{TIR1}
48 complex breaks down the dimer formed by Aux/IAA protein and ARF protein through ubiquitin
49 (Szemenyei et al. 2008; Winkler et al. 2017). At the same time, proteases degrade Aux/IAA
50 protein to restore ARF's activity (Luo et al. 2018; Parry et al. 2009; Winkler et al. 2017).

51 Aux/IAA protein is a crucial gene involved in the early response to auxin, and plays a vital role
52 in regulating plant growth and development (Luo et al. 2018). The protein typically consists of
53 four distinct domains, as described in previous studies (Dreher et al. 2006; Guilfoyle & Hagen
54 2007b). Domain I, acting as a transcriptional repressor, can be repressed by the adjacent
55 promoter region. It contains an EAR motif associated with the ethylene response factor (ERF)
56 (Szemenyei et al. 2008; Tiwari et al. 2003). The LxLxL motif in the EAR motif of domain I may
57 attract TPLs, which, when combined with the conserved N-terminal TPD domain of TPLs, create
58 a common repressor, thus serving as transcriptional suppressors of downstream auxin regulating
59 genes. Domain II serves as the degradation determinant, featuring the conserved TIR1/AFB
60 recognition sequence GWPPV, which plays a vital role in the stability of the Aux/IAA protein
61 (Shimizu-Mitao & Kakimoto 2014). The GWPPV motif is a combination of five amino acids,
62 namely Gly-Trp-Pro-Pro-Val. Mutations in the GWPPV recognition sequence may delay the
63 degradation of Aux/IAA and disrupt the auxin signal transduction pathway (Shimizu-Mitao &
64 Kakimoto 2014). Domain III, with its βαα-fold structure, bears structural and functional
65 resemblance to the DNA recognition motif in Arc and Met J repressors (Dreher et al. 2006; Liu
66 et al. 2011; Luo et al. 2018). The fourth and final domain, Domain IV, contains one acidic
67 domain and one SV40 type NLS (PKKKRKV) domain (Dreher et al. 2006; Liu et al. 2011; Luo
68 et al. 2018). Domains III and IV are identical to the C-terminal dimerization domain (CTD
69 domain) of ARF proteins, capable of generating type I / II PB1 domains that interact with

70 Aux/IAA and ARF family proteins to regulate ARF activity (Guilfoyle & Hagen 2007b; Luo et
71 al. 2018; Winkler et al. 2017).

72 Currently, through genome-wide analysis, the Aux/IAA gene family has been identified in over
73 30 plant species. For example, *Arabidopsis thaliana* (*A. thaliana*) contains 29 IAA proteins
74 (Liscum & Reed 2002), *Oryza sativa* (*O. sativa*) contains 31 members (Jain et al. 2006),
75 *Solanum tuberosum* (*S. tuberosum*) and *Vitis vinifera (V. vinifera)* all contain 26 members (Çakir
76 et al. 2013; Gao et al. 2016). Certain Aux/IAA genes have been shown to play a crucial role in
77 plant growth, among which the AtIAA7 protein may inhibit flowering time under short-day light
78 by negatively regulating the expression of *GA20ox1* and *GA20ox2* genes (Luo et al. 2018; Mai et
79 al. 2011). The AtIAA8 protein interacts with the AtARF6/8 protein and modulates the level of
80 jasmonic acid (JA) to influence flower organ development (Wang et al. 2013). The AtIAA17
81 protein plays an important role in regulating leaf senescence (Shi et al. 2015). Functional
82 analysis of the *AtIAA28* mutant revealed that *AtIAA28* promotes the transcription of lateral root
83 initiation genes in response to auxin signals (Rogg et al. 2001).

84 *Salvia miltiorrhiza (S. miltiorrhiza)*, a member of the genus *Salvia* within the *Labiatae* family, is
85 well known for its dried roots and rhizomes, which serve as a traditional Chinese medicine (Xu
86 et al. 2016). Due to its advantageous growth characteristics, including a simple growth
87 environment, short reproductive cycle, small genome size, and ease of tissue culture, *S.
88 miltiorrhiza* has been established as a model system for the study of medicinal plants (Wang et
89 al. 2021; Xu et al. 2016). The Aux/IAA gene family plays a crucial role in regulating plant
90 growth and development (Guilfoyle & Hagen 2007b)., making the identification of these genes
91 in *S. miltiorrhiza* of paramount importance for understanding the developmental process and
92 cellular response of auxin in this medicinal plant. This information could have important
93 implications for the development of new plant varieties with improved growth and stress
94 tolerance, as well as for the utilization of *S. miltiorrhiza* as a valuable medicinal resource.
95 However, in *S. miltiorrhiza*, the use of the whole genome to study the Aux/IAA gene family has
96 not been reported.

97 In the present investigation, we executed a comprehensive analysis of the Aux/IAA gene family
98 in *S. miltiorrhiza* through a genome-wide approach. As a result, we successfully identified a total
99 of 23 members of the *S. miltiorrhiza* Aux/IAA gene family (designated as *SmIAA*).
100 Subsequently, we performed bioinformatic analyses to evaluate the physicochemical properties
101 and sequence features of these genes. In addition, we utilized RNA-seq data to investigate the
102 tissue-specific expression profiles of the *SmIAAs*, and further established a co-expression
103 network to uncover the co-expression relationships between the *SmIAAs* and *transcription
104 factors*.

105

Materials & Methods

Identification and Characterization of the Aux/IAA Gene Family in *S. miltiorrhiza*

The Aux/IAA protein sequences of *A. thaliana* and *O. sativa* were downloaded from The Arabidopsis Information Resources (TAIR) database and Rice Genome Annotation Project (RGAP) database respectively. The genomic data of *S. miltiorrhiza* was downloaded from National Genomics Data Center (NGDC) database (The project number is GWHAOSJ00000000).

We employed two methods to identify potential members of the *SmIAAs*. First, we utilized the *A. thaliana* Aux/IAA protein sequence as a seed sequence for BLASTP homologous alignment in the *S. miltiorrhiza* local database. We then validated the domain structure of the resulting sequences using the Pfam database (El-Gebali et al. 2018).and retained only those sequences that contained the AUX_IAA domain as candidate members of the *SmIAAs*. Second, we utilized the iTAK database to comprehensively evaluate the entire *S. miltiorrhiza* protein sequence and identified additional members of the candidate *SmIAAs* (Zheng et al. 2016). By combining the candidate gene family members obtained by these two methods, we determined the final set of members comprising the *SmIAAs*.

To explore the evolutionary relationships among *SmIAAs*, we utilized the Neighbor-Joining (NJ) approach implemented in MEGA X software (Kumar et al. 2018), employing the following parameters: 50% partial deletion and 500 bootstrap replications. Additionally, we analyzed the basic physicochemical properties of the protein sequences using the Expasy online website (http://web.expasy.org/protparam/) and predicted their subcellular localization using the WoLF PSORT software (https://wolfpsort.hgc.jp/). To further investigate the chromosomal distribution of SmIAAs, we obtained relevant information from the S. miltiorrhiza genome annotation file and visualized it using the MG2C software (http://mg2c.iask.in/mg2c_v2.0/).

Genetic Analysis of Aux/IAA Gene Sequence and Promoter Cis-Acting Elements in *S. miltiorrhiza*

The structure of *SmIAA* gene sequence was analyzed by TBtools software according to the GFF3 file and CDS file of *S. miltiorrhiza* (Chen et al. 2020). The protein sequence domain of SmIAA proteins were mined by CD-search online database (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi). Finally, Tbtools software was used to visually analyze the results (Chen et al. 2020). Furthermore, GO enrichment analysis was performed using R software. In addition, the promoter sequence of 2000 bp upstream of the start codon of the *SmIAAs* was extracted from the *S. miltiorrhiza* genome. Subsequently, the promoter cis-acting elements were analyzed using the online website PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), with the results of this analysis being visualized through TBtools software, as detailed in S_Table1.

Comparative Analysis of Aux/IAA Gene Family Across Different Species

143 In order to explore the genetic and evolutionary relationships among members of the Aux/IAA
144 gene family in various species, a multiple sequence alignment of 23 SmIAA proteins, 29 AtIAA
145 proteins, and 31 OsIAA proteins was performed using the online program Clustal Omega,
146 followed by MEGAX to create the phylogenetic tree (S_Table2), with the parameters used in
147 line with previous analyses (Kumar et al. 2018; Sievers & Higgins 2014). Furthermore, we
148 utilized the Maximum Likelihood (ML) method to construct an additional evolutionary tree to
149 compare with the results obtained using the NJ method (S_Figure 4). This was done with the
150 implementation of specific parameters, including a 50% partial deletion criterion and 500
151 bootstrap replications. The phylogenetic tree was shown using the online software Evolview
152 (https://www.evolgenius.info/evolview/) (Subramanian et al. 2019). MEME, an online program
153 (https://meme-suite.org/meme/meme_5.3.3/tools/meme), was used to assess conserved motifs.

154 **Collinearity Analysis and Selection Pressure Analysis**

155 The GFF annotation file of the genome was utilized to acquire the chromosomal location data for
156 *SmIAAs*. The MCScanX software, developed by Wang et al. (2012), was subsequently employed
157 to compute the collinear block and tandem repeat gene information. The TBtools software (Chen
158 et al. 2020) was utilized to visualize the chromosomal distribution and collinear relationships of
159 the *SmIAAs*. For each identified homologous gene pair by MCScanX, the synonymous
160 substitution rate (Ks) and non-synonymous substitution rate (Ka) were computed using
161 KaKs_Calculator software (Wang et al. 2010)for each synonymous site. The Ka/Ks ratio was
162 utilized to evaluate the rate of protein evolution and determine the selection pressure of the genes
163 during evolution.

164 **RNA-seq Analysis of *SmIAA* Expression in *S. miltiorrhiza***

165 The RNA-seq data of *S. miltiorrhiza* were obtained from the NCBI database
166 (PRJNA744957&PRJNA798876). The received Sequence Read Archive (SRA) files were first
167 converted to the fastq format using the fastq-dump program. The downloaded raw data was then
168 trimmed and preprocessed using TrimGalore (https://github.com/FelixKrueger/TrimGalore), and
169 the resulting clean data was utilized for downstream analysis. To construct a genome index with
170 default settings, the Hisat2-build command in Hisat2 (version 2.3.0) (Zhang et al. 2021) was
171 used. The featureCounts software was then employed to calculate the raw counts of each gene in
172 each sample. The standardized transcriptional abundance was determined by computing the
173 Transcripts Per Million (TPM) value. The raw counts of each sample were used to generate a
174 gene expression matrix of 21 transcriptome samples (S_Table3). Finally, the expression data for
175 the *SmIAAs* was extracted and a corresponding heat map was generated using the TBtools
176 software (Chen et al. 2020).

177 **Gene Co-expression Network Analysis in *S. miltiorrhiza***

178 Transcription factors were identified in the entire protein sequence of Salvia miltiorrhiza using
179 the iTAK online software developed by Zheng et al. (2016). Gene co-expression network
180 analysis was conducted using R software version 3.9.4 and the Weighted Gene Co-expression

181   Network Analysis (WGCNA) package developed by Langfelder & Horvath (2008). First, the
182   goodSamplesGenes function within the WGCNA package was utilized to identify and filter out
183   any genes with missing values. The soft threshold power of 14 was then selected to construct the
184   weighted gene network, with parameter settings of minModuleSize=30, mergeCutHeight=0.25,
185   and maxBlockSize=6,000. Additionally, to ensure the biological significance of each co-
186   expression module, GO enrichment analysis was performed for each module using R software.

187   **Gene Expression Analysis of *S. miltiorrhiza* using qPCR and Transcriptomics Data**
188   **Integration**

189   Two-year-old *S. miltiorrhiza* seedlings were collected and frozen in a -80°C refrigerator. The
190   roots, stems, leaves, and flowers were used for specific expression detection. Fluorescent
191   quantitative specific primers (S_Table 4) were designed using Primer 5.0, and *SmUBQ* was used
192   as an internal reference gene. Total RNA was extracted using the RNAprep Pure Plant Plus Kit
193   (Tiangen, Beijing) and reverse transcribed using the PrimeScriptTM reverse transcriptase kit
194   (Takara Tokyo, Japan). Real-time PCR reactions were performed using the PerfectStart Green
195   qPCR SuperMix Kit (Transgene, Beijing), and fluorescence signals were collected using a
196   qTOWER 3G PCR instrument. Each sample was analyzed in triplicate, and relative expression
197   was calculated using the $2^{-\triangle\triangle CT}$ method (S_Table 5). The correlation between transcriptome data
198   and qPCR data was calculated and visualized by Sangerbox website (http://sangerbox.com/)
199   (Shen et al. 2022).
200

201   # Result

202   **Identification of *Aux/IAA* Gene Family *members in S. miltiorrhiza***

203   In this study, two distinct methods were employed to identify the Aux/IAA gene family across
204   the entire genome of *S. miltiorrhiza*. The initial approach involved using *Arabidopsis* Aux/IAA
205   protein sequences to conduct a BLASTp search on the proteomic data of *S. miltiorrhiza*.
206   Subsequently, the results were subjected to domain detection via submission to the Pfam
207   database (El-Gebali et al. 2018), and the undesired outcomes were removed. This led to the
208   identification of a total of 23 potential Aux/IAA gene family members. The second method
209   involved employing the iTAK database to analyze the complete protein sequences of *S.
210   miltiorrhiza,* which also resulted in the identification of 23 candidate genes. By comparing the
211   outcomes of the two methods, it was established that the *S. miltiorrhiza* genome contained 23
212   members of the Aux/IAA gene family. In order to simplify future research, the nomenclature of
213   *Arabidopsis genes* was adopted, and the genes were labeled as SmIAA1-SmIAA23 based on
214   their distribution across the evolutionary tree (S_Figure 1).

215   A comprehensive analysis was conducted to investigate the physicochemical properties of the
216   SmIAA protein sequence. The results revealed that the number of amino acids in the SmIAA
217   protein ranged from 87 to 371 aa, with an average of 211.61 aa. The molecular weight of the

218 SmIAA protein was found to be between 10,364.81 and 40,197.45 Da, while the theoretical
219 isoelectric point ranged from 4.82 to 9.81. Moreover, the subcellular localization analysis
220 indicated that some proteins, namely SmIAA1, SmIAA4, SmIAA5, SmIAA8, SmIAA10,
221 SmIAA11, SmIAA14, SmIAA16, SmIAA18, SmIAA22, and SmIAA23, were located in the
222 cytoplasm, whereas the remaining proteins were all localized in the nucleus (Table 1).

223 Furthermore, we referred to the genome annotation information of *S. miltiorrhiza* to identify the
224 distribution patterns of *SmIAAs*, with the exception of *SmIAA23* on contig685. Our analysis
225 revealed that these genes were situated on chromosomes other than chromosome 7, with most
226 genes located at the upper and lower ends of the chromosome. The highest frequency of gene
227 distribution was observed on chromosome 5 (eight genes), while the lowest was on
228 chromosomes 6 and 8 (one gene). Moreover, the genes on chromosomes 2 and 5 were more
229 densely distributed compared to those on other chromosomes, which were more sparsely and
230 randomly distributed (S_Figure 2).

231 **Analysis of gene structure and protein sequence**

232 Previous studies have demonstrated that gene structural diversity is a significant driver for the
233 evolution of many gene families (Singh & Jain 2015). To further explore the gene structure of
234 *SmIAAs* (Figure 1A), we conducted an analysis which revealed variable splicing in all genes. The
235 number of introns ranged from 1 to 6. Notably, some genes did not have UTRs. *SmIAA2,*
236 *SmIAA3, SmIAA4, SmIAA6, SmIAA13, SmIAA17,* and *SmIAA23* contained only CDS and no
237 UTR.

238 To gain more insights into the properties of SmIAA proteins, we performed domain analysis
239 using the Pfam database. The results showed that all proteins contained Aux/IAA domains,
240 although their placement and size varied (Figure 1B). We also conducted multiple sequence
241 alignment to evaluate the conserved amino acid sites in the SmIAA proteins. The domains of
242 SmIAA proteins were further divided into Domain I (EAR-motif), Domain II (Degron motif),
243 Domain III (Basic motif), and Domain IV (OPCA-like motif) (Figure 2 and S_Figure 3). Some
244 SmIAA proteins were found to be atypical Aux/IAA proteins, lacking at least one conserved
245 domain. This structural diversity may provide these proteins with diverse functions, although
246 further research is required.

247 **Phylogenetic analysis of SmIAA proteins**

248 To investigate the potential relationship between typical and atypical AUX/IAA proteins, we
249 identified and selected the Aux/IAA proteins from two model organisms, *A. thaliana* and *O.*
250 *sativa*, as well as *S. miltiorrhiza* for further study. We utilized MEGA X to construct
251 phylogenetic trees of the Aux/IAA proteins in three species. The results obtained from the ML
252 method and NJ method were largely consistent, indicating the reliability of our findings.
253 Additionally, we employed the MEME database to analyze the motifs present in these proteins.
254 The resulting evolutionary tree was primarily divided into two groups, I and II, respectively. As
255 illustrated in Figure 3, the evolutionary tree and conserved motifs were consistent with previous

256 studies on the Aux/IAA family of plants. The majority of the members were classified in Group
257 I, with nearly all of them being canonical Aux/IAA proteins possessing four canonical domains
258 (motif1, 2, 3, and 4). On the other hand, nearly all the members of Group II were found to be
259 atypical Aux/IAA proteins, lacking at least one conventional domain. Atypical Aux/IAA proteins
260 have been found to be prevalent in the Aux/IAA gene family of various plants, where they play a
261 significant role in plant adaptability to changing environmental conditions, although their
262 specific functions remain largely unknown (Luo et al. 2018).

263 **Analysis of collinearity and selection pressure**

264 We conducted a collinear analysis of *SmIAAs* to elucidate their relationships further. The results
265 showed that nine lineal homologous gene pairs were identified among all *SmIAAs*, suggesting
266 that chromosomal replication processes play a crucial role in the amplification of the *S.*
267 *miltiorrhiza* Aux/IAA gene family members (Figure 4).

268 To gain a better understanding of the evolutionary processes of genes and proteins, studying
269 selective interaction patterns can be a useful approach. Selection has been shown to be a valuable
270 technique for speculating about gene function (Wang et al. 2009). The Ka/Ks value, which is the
271 ratio of the two protein-coding genes' non-synonymous substitution rates (Ka) and synonymous
272 substitution rates (Ks), can be used to identify whether the gene is under selection pressure.
273 When Ka/Ks > 1, the gene is thought to be positively selected; when Ka/Ks = 1, the gene is
274 thought to be purified and selected; and when Ka/Ks < 1, the gene is thought to be neutrally
275 selected. Except for *SmIAA8/SmIAA14* and *SmIAA8/SmIAA19*, the Ka/Ks value of the other *S.*
276 *miltiorrhiza* lineal homologous gene pairs is less than 1, which means that these *SmIAAs* evolved
277 under purification selection, indicating that the sequence of the Aux/IAA gene is relatively
278 conservative. These findings suggest that the Aux/IAA gene plays a significant role in *S.*
279 *miltiorrhiza* and attains an optimal state, which is consistent with the conservative function of
280 several *SmIAAs*.

281 **Analysis of GO enrichment and *cis*-acting element**

282 The GO categorization system is a widely used method for characterizing genes and proteins in
283 various organisms. The system is organized into three distinct categories, including molecular
284 function, biological process, and cellular component, with each category corresponding to
285 specific functions or properties. To investigate the potential functions of the *SmIAAs*, we utilized
286 TBtools to examine the GO enrichment of *SmIAA* and identified the top ten most significantly
287 enriched GO terms across the three categories (if less than 10, all).

288 Figure 5A demonstrates that the *SmIAAs* are associated with 10, 5, and 3 GO terms related to
289 biological process, cellular component, and molecular function, respectively. The majority of
290 *SmIAAs* are believed to play a role in various biological activities. Their molecular functions
291 include transcription regulator activity, DNA-binding transcription factor activity, and identical
292 protein binding. The cell composition mainly includes the nucleus, nuclear, intracellular
293 membrane-bounded organelle, membrane-bounded organelle, and other cellular components.

294   Furthermore, *SmIAAs* are primarily involved in biological processes related to the response to
295   auxin, organic substance, endogenous stimulus, and hormone, among others.

296   To gain a more comprehensive understanding of the role of *SmIAAs* in the growth and
297   development of *S. miltiorrhiza*, we analyzed the cis-acting elements in the promoter region of the
298   related genes, which is 2000 bp upstream of the initiation codon (Figure 5B and S_Table 1). Our
299   analysis revealed the presence of numerous cis-acting elements associated with auxin, salicylic
300   acid, abscisic acid, methyl jasmonate, and gibberellin in the promoter region of *SmIAAs*, in
301   addition to key cis-acting elements associated with transcription and light response. The presence
302   of these cis-regulatory components suggests that hormones have the capacity to regulate *SmIAAs*.
303   Overall, our analysis of cis-acting elements reveals that different family members contain a
304   variety of response motifs, implying that the SmIAA gene family may participate in various
305   physiological processes of *S. miltiorrhiza* by responding to various response elements such as
306   hormone, low temperature, light, and so on.

Analysis of gene expression pattern

308   The temporal and spatial expression of a gene may provide insights into its functional features
309   (Wang et al. 2021). In this study, transcriptome data retrieved from NCBI was used to extract the
310   TPM values of root, stem, leaf, and flower tissues of *S. miltiorrhiza*, and subsequently analyze
311   the expression pattern of the *SmIAAs*. The results of gene expression patterns based on
312   transcriptome data indicated considerable variability in the expression of several *SmIAAs* across
313   different tissues of *S. miltiorrhiza*, as well as in the expression trends of individual genes. Based
314   on the results of expression cluster analysis, *SmIAAs* could be categorized into six groups (A~F)
315   (Figure 6). Group A consisted of two genes (*SmIAA4* and *SmIAA6*) that were primarily expressed
316   in the leaves of *S. miltiorrhiza*. Group B included four genes (*SmIAA19, SmIAA16, SmIAA10,*
317   and *SmIAA11*) that showed similar expression levels in the leaves, stems, and flowers of *S.
318   miltiorrhiza*, and relatively low expression levels in the roots. Group C comprised five genes
319   (*SmIAA14, SmIAA1, SmIAA5, SmIAA23,* and *SmIAA22*) that were predominantly expressed in the
320   flowers and roots of *S. miltiorrhiza*. Group D encompassed six genes (*SmIAA12, SmIAA18,*
321   *SmIAA21, SmIAA7, SmIAA17,* and *SmIAA20*) that showed the highest expression levels in the
322   stems, roots, and flowers of *S. miltiorrhiza*, but minimal expression in the leaves. Group E
323   consisted of three genes (*SmIAA13, SmIAA15,* and *SmIAA8*) that were primarily expressed in the
324   stems and roots of *S. miltiorrhiza*. However, the three genes in group E (*SmIAA2, SmIAA9,* and
325   *SmIAA3*) were significantly expressed only in the stem of *S. miltiorrhiza*. In summary, *SmIAAs*
326   were expressed in clusters in various tissues of *S. miltiorrhiza*, with the majority of *SmIAAs*
327   exhibiting high expression levels in roots, stems, and particularly flowers.

Weighted correlation network analysis

329   Although the Aux/IAA gene family is crucial for auxin signal transduction, prior studies have
330   suggested that its members may interact with other transcription factors (TFs) besides *ARFs* (Luo
331   et al. 2018). To explore these interactions, we employed the iTAK tool to identify all TFs in the
332   *S. miltiorrhiza* genome and utilized Weighted Correlation Network Analysis (WGCNA) to

333   examine additional transcriptome data from *S. miltiorrhiza* flower tissue at different stages
334   (withering, full blooming, and bud). Our analysis identified 68 TFs, including *ARF*, *ERF*, *MYB*,
335   and others (Table 2 and S_Table 6). We used a soft threshold of β=14 to divide the 1814 TFs
336   into 10 modules, with gene numbers ranging from 45 to 490 in each module (S_Table 7). The
337   outcomes of the GO enrichment analysis of distinct modules revealed that each of these modules
338   has distinct biological significance (S_Table 8). Furthermore, utilizing the Pearson correlation
339   data, we observed a significant association between *SmIAAs* and the modules, suggesting that
340   *SmIAAs* may be linked with the module hub genes (Figure 7). We selected the red and cyan
341   modules for further investigation, as they were highly associated with *SmIAA14* and negatively
342   related to *SmIAA22*, respectively. We identified the top 10 hub genes in each module and
343   annotated their homologs in the iTAK database (Table 3). Surprisingly, only one *ARF* was
344   among the 20 key genes identified across various modules (EVM0019866), suggesting that
345   *SmIAAs* may have other important functions beyond their role in the auxin pathway. However,
346   more research is needed to confirm this hypothesis.

347   **The Validation of qRT-PCR**

348   To confirm the RNA-Seq results, we employed the quantitative reverse transcription polymerase
349   chain reaction (qRT-PCR) technique to evaluate the expression levels of 12 genes, comprising of
350   4 *SmIAAs*, 4 hub genes selected by co-expression modules, and 4 transcription factors chosen
351   randomly (Figure 8). The findings demonstrated that the alteration in gene expression detected
352   by qRT-PCR and RNA-Seq techniques was similar, and the outcomes of the two datasets were
353   positively correlated, suggesting that the co-expression network established in this study is
354   reliable.

355

356   # Discussion

357   The Aux/IAA gene family is a large and diverse family of transcriptional regulators that plays a
358   critical role in the response pathway of plant auxin. The family has been identified and studied in
359   a wide range of plant species, from soybean to wheat, tomato, and March plum (Gao et al. 2016;
360   Jain et al. 2006; Liscum & Reed 2002; Luo et al. 2018). The family members share a common
361   structure that includes four conserved domains, and their expression is regulated by the presence
362   and concentration of auxin in plant cells. This study focuses on the identification, structural, and
363   phylogenetic analysis of the Aux/IAA gene family in *S. miltiorrhiza*. On the whole, the number
364   of *SmIAAs* is similar to that of Aux/IAA gene family members in *A. thaliana* (Liscum & Reed
365   2002). Most of the *SmIAAs* are located on chromosomes other than chromosome 7, and many of
366   them are situated at the ends of the chromosomes. This suggests that the expansion of the
367   *SmIAAs* may have been caused by tandem repeats.

368   Protein domains are important functional units, and conserved motifs are subunits of domains
369   that contribute to the diverse biological functions of the domains. In the proteins encoded by the

370 *SmIAAs*, 17 members contained four typical conserved domains of the family, while the
371 remaining six members had deletions of these typical domains. These atypical Aux/IAA proteins
372 are mainly missing Domain I (EAR-motif), so it is inferred that the transcriptional inhibitory
373 function of these six members of *S. miltiorrhiza* on downstream ARF may be missing (Luo et al.
374 2018). It is reported that atypical Aux/IAA proteins are ubiquitous in various plant *Aux/IAA* gene
375 families and play an important role in plant adaptation to changing environmental conditions, but
376 more specific functions still need to be further studied (Li et al. 2017; Luo et al. 2018).

377 In addition to 4 typical conserved motifs, there are 6 conserved motifs in Aux/IAA protein. The
378 overlap and difference of conserved motifs among different members reflect its functional
379 characteristics to some extent. Other different conserved motifs may enable members to
380 participate in functions other than auxin signal pathway, and the different expression patterns of
381 members in different tissues of *S. miltiorrhiza* indicate that these different conserved motifs play
382 different roles to a certain extent, but the specific functions of their effects need to be further
383 studied. Moreover, the promoter analysis of *SmIAAs* suggests that they may respond to various
384 stimuli, including auxin, gibberellin, abscisic acid, salicylic acid, methyl jasmonate, drought,
385 salt, heat stress, and low temperature. The presence of these cis-acting elements may explain
386 why *SmIAAs* have a wide range of functions in *S. miltiorrhiza*, including tissue-specific roles.
387 However, further experimental research is required to confirm this. Overall, the physicochemical
388 properties of SmIAA proteins contribute to their regulation of gene expression and involvement
389 in plant adaptation to changing environmental conditions.

390 Transcriptome analysis showed that the expression of different Aux/IAA family genes was tissue
391 specific, but Aux/IAA family genes were involved in each different tissue of *S. miltiorrhiza*.
392 Previous reports have pointed out that *Aux/IAA* gene not only interacts with ARF, but also is
393 regulated by other transcription factors (Guilfoyle & Hagen 2007b; Luo et al. 2018). In this
394 study, co-expression network analysis was performed using transcriptome data from *S.*
395 *miltiorrhiza* to identify transcription factors that may interact with the *SmIAAs*. This analysis
396 identified 10 modules, and further correlation analysis revealed that each module was negatively
397 correlated with at least one *SmIAA*. Analysis of the specific expression patterns of different
398 *SmIAAs* in different tissues suggests that the different modules may play different roles in
399 different tissues of *S. miltiorrhiza* (Table 4). Further analysis of the black and red modules
400 revealed an interesting finding: although the Aux/IAA family is known to play an important role
401 in auxin signaling, the hub genes in these two modules had only a small number of *ARFs* that
402 were significantly associated with *SmIAAs*. This suggests that the *SmIAAs* may have additional,
403 unknown functions in *S. miltiorrhiza*, but further research is needed to confirm this.

404

Conclusions

406 In this study, a comprehensive analysis was conducted on the Aux/IAA gene family in *S.*
407 *miltiorrhiza*, which included whole genome identification, physicochemical property analysis,

408 and expression analysis. Moreover, a gene co-expression network was constructed to investigate
409 the interactions of *SmIAAs* with transcription factors. The findings revealed that this gene family
410 comprised numerous members with diverse structures and functions. Notably, it was observed
411 that these genes are strongly correlated with transcription factors other than ARF. For instance,
412 *SmIAA9* was found to cooperate with the *WRKY (EVM0007532)* to regulate stem growth, while
413 *SmIAA22* and the *DBB (EVM0017797)* were observed to antagonize each other and regulate root
414 development. These results suggest that *SmIAAs* perform various unknown functions in different
415 tissues and during different growth and development processes of *S. miltiorrhiza*. The outcomes
416 of this research provide a foundation for further elucidation of the function of the *SmIAAs* and
417 offer valuable resources for improving the characteristics of *S. miltiorrhiza* and breeding new *S.*
418 *miltiorrhiza* varieties through genetic engineering.

419

420 **Conflicts of Interest:** The authors declare no conflict of interest.

421

422 # Reference

423 Çakir B, Kiliçkaya O, and Olcay AC. 2013. Genome-wide analysis of Aux/IAA
424   genes in Vitis vinifera: Cloning and expression profiling of a grape
425   Aux/IAA gene in response to phytohormone and abiotic stresses. *Acta*
426   *Physiologiae Plantarum* 35:365-377. 10.1007/s11738-012-1079-7
427 Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, and Xia R. 2020. TBtools:
428   An Integrative Toolkit Developed for Interactive Analyses of Big
429   Biological Data. *Molecular Plant* 13:1194-1202.
430   https://doi.org/10.1016/j.molp.2020.06.009
431 Dreher KA, Brown J, Saw RE, and Callis J. 2006. The Arabidopsis Aux/IAA
432   protein family has diversified in degradation and auxin responsiveness.
433   *Plant Cell* 18:699-714. 10.1105/tpc.105.039172
434 El-Gebali S, Mistry J, Bateman A, Eddy S, Luciani A, Potter S, Qureshi M,
435   Richardson L, Salazar G, Smart A, Sonnhammer E, Hirsh L, Paladin L,
436   Piovesan D, Tosatto S, and Finn R. 2018. The Pfam protein families
437   database in 2019. *Nucleic acids research* 47. 10.1093/nar/gky995
438 Gao J, Cao X, Shi S, Ma Y, Wang K, Liu S, Chen D, Chen Q, and Ma H. 2016.
439   Genome-wide survey of Aux/IAA gene family members in potato (Solanum
440   tuberosum): Identification, expression analysis, and evaluation of their
441   roles in tuber development. *Biochemical and Biophysical Research*
442   *Communications* 471:320-327. https://doi.org/10.1016/j.bbrc.2016.02.013
443 Guilfoyle TJ, and Hagen G. 2007a. Auxin response factors. *Current Opinion in*
444   *Plant Biology* 10:453-460. 10.1016/j.pbi.2007.08.014
445 Guilfoyle TJ, and Hagen G. 2007b. Auxin response factors. Current Opinion in
446   Plant Biology. p 453-460.

447   Jain M, Kaur N, Garg R, Thakur JK, Tyagi AK, and Khurana JP. 2006. Structure
448   and expression analysis of early auxin-responsive Aux/IAA gene family in
449   rice (Oryza sativa). *Functional and Integrative Genomics* 6:47-59.
450   10.1007/s10142-005-0005-0
451   Kepinski S, and Leyser O. 2005. The Arabidopsis F-box protein TIR1 is an auxin
452   receptor. *Nature* 435:446-451. 10.1038/nature03542
453   Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: molecular
454   evolutionary genetics analysis across computing platforms. *Molecular*
455   *Biology and Evolution* 35:1547.
456   Langfelder P, and Horvath S. 2008. WGCNA: An R package for weighted
457   correlation network analysis. *BMC Bioinformatics* 9. 10.1186/1471-2105-9-
458   559
459   Lavy M, and Estelle M. 2016. Mechanisms of auxin signaling. *Development*
460   *(Cambridge)* 143:3226-3229. 10.1242/dev.131870
461   Leyser HMO, Pickett FB, Dharmasiri S, and Estelle M. 1996. Mutations in the
462   AXR3 gene of Arabidopsis result in altered auxin response including
463   ectopic expression from the SAUR-AC1 promoter. *Plant Journal* 10:403-413.
464   10.1046/j.1365-313x.1996.10030403.x
465   Leyser O. 2017. Auxin signaling. *Plant Physiol*.
466   Li H, Wang B, Zhang Q, Wang J, King GJ, and Liu K. 2017. Genome-wide analysis
467   of the auxin/indoleacetic acid (Aux/IAA) gene family in allotetraploid
468   rapeseed (Brassica napus L.). *BMC Plant Biology* 17. 10.1186/s12870-017-
469   1165-5
470   Liscum E, and Reed JW. 2002. Genetics of Aux/IAA and ARF action in plant
471   growth and development. *Plant Molecular Biology* 49:387-400.
472   10.1023/A:1015255030047
473   Liu Y, Jiang HY, Chen W, Qian Y, Ma Q, Cheng B, and Zhu S. 2011. Genome-wide
474   analysis of the auxin response factor (ARF) gene family in maize (Zea
475   mays). *Plant Growth Regulation* 63:225-234. 10.1007/s10725-010-9519-0
476   Luo J, Zhou J-J, and Zhang J-Z. 2018. Aux/IAA Gene Family in Plants: Molecular
477   Structure, Regulation, and Function. *International journal of molecular*
478   *sciences* 19. 10.3390/ijms19010259
479   Mai YX, Wang L, and Yang HQ. 2011. A Gain-of-Function Mutation in IAA7/AXR2
480   Confers Late Flowering under Short-day Light in Arabidopsis F. *Journal*
481   *of Integrative Plant Biology* 53:480-492.
482   Mironova V, Teale W, Shahriari M, Dawson J, and Palme K. 2017. The Systems
483   Biology of Auxin in Developing Embryos. *Trends in Plant Science* 22:225-
484   235. 10.1016/j.tplants.2016.11.010
485   Parry G, Calderon-Villalobos LI, Prigge M, Peret B, Dharmasiri S, Itoh H,
486   Lechner E, Gray WM, Bennett M, and Estelle M. 2009. Complex regulation
487   of the TIR1/AFB family of auxin receptors. *Proceedings of the National*

488   *Academy of Sciences of the United States of America* 106:22540-22545.
489   10.1073/pnas.0911967106
490 Rogg LE, Lasswell J, and Bartel B. 2001. A gain-of-function mutation in IAA28
491   suppresses lateral root development. *Plant Cell* 13:465-480.
492   10.1105/tpc.13.3.465
493 Shen W, Song Z, Zhong X, Huang M, Shen D, Gao P, Qian X, Wang M, He X, Wang T,
494   Li S, and Song X. 2022. Sangerbox: A comprehensive, interaction-friendly
495   clinical bioinformatics analysis platform. *iMeta* 1:e36. 10.1002/imt2.36
496 Shi H, Reiter RJ, Tan DX, and Chan Z. 2015. INDOLE-3-ACETIC ACID INDUCIBLE
497   17 positively modulates natural leaf senescence through
498   melatonin-mediated pathway in Arabidopsis. *Journal of Pineal Research*
499   58:26-33.
500 Shimizu-Mitao Y, and Kakimoto T. 2014. Auxin Sensitivities of All Arabidopsis
501   Aux/IAAs for Degradation in the Presence of Every TIR1/AFB. *Plant and*
502   *Cell Physiology* 55:1450-1459. 10.1093/pcp/pcu077
503 Sievers F, and Higgins DG. 2014. Clustal omega. *Current protocols in*
504   *bioinformatics* 48:3.13. 11-13.13. 16.
505 Singh VK, and Jain M. 2015. Genome-wide survey and comprehensive expression
506   profiling of Aux/IAA gene family in chickpea and soybean. *Frontiers in*
507   *Plant Science* 6. 10.3389/fpls.2015.00918
508 Subramanian B, Gao S, Lercher MJ, Hu S, and Chen W-H. 2019. Evolview v3: a
509   webserver for visualization, annotation, and management of phylogenetic
510   trees. *Nucleic acids research* 47:W270-W275. 10.1093/nar/gkz357
511 Szemenyei H, Hannon M, and Long JA. 2008. TOPLESS mediates auxin-dependent
512   transcriptional repression during Arabidopsis embryogenesis. *Science*
513   319:1384-1386. 10.1126/science.1151461
514 Tiwari SB, Hagen G, and Guilfoyle T. 2003. The roles of auxin response factor
515   domains in auxin-responsive transcription. *Plant Cell* 15:533-543.
516   10.1105/tpc.008417
517 Wang D, Zhang S, He F, Zhu J, Hu S, and Yu J. 2009. How do variable
518   substitution rates influence Ka and Ks calculations? *Genomics Proteomics*
519   *Bioinformatics* 7:116-127. 10.1016/s1672-0229(08)60040-6
520 Wang D, Zhang Y, Zhang Z, Zhu J, and Yu J. 2010. KaKs_Calculator 2.0: a
521   toolkit incorporating gamma-series methods and sliding window
522   strategies. *Genomics Proteomics Bioinformatics* 8:77-80.
523 Wang J, Yan DW, Yuan TT, Gao X, and Lu YT. 2013. A gain-of-function mutation
524   in IAA8 alters Arabidopsis floral organ development by change of
525   jasmonic acid level. *Plant Molecular Biology* 82:71-83. 10.1007/s11103-
526   013-0039-y

527 Wang K, Cheng Y, Yi L, He H, Zhan S, and Yang P. 2021. Genome-wide
528   identification of the Tubby-Like Protein (TLPs) family in medicinal
529   model plant Salvia miltiorrhiza. *PeerJ* 9:e11403. 10.7717/peerj.11403
530 Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo
531   H, Kissinger JC, and Paterson AH. 2012. MCScanX: A toolkit for detection
532   and evolutionary analysis of gene synteny and collinearity. *Nucleic*
533   *acids research* 40. 10.1093/nar/gkr1293
534 Winkler M, Niemeyer M, Hellmuth A, Janitza P, Christ G, Samodelov SL, Wilde V,
535   Majovsky P, Trujillo M, Zurbriggen MD, Hoehenwarter W, Quint M, and
536   Villalobos LIAC. 2017. Variation in auxin sensing guides AUX/IAA
537   transcriptional repressor ubiquitylation and destruction. *Nature*
538   *Communications* 8. 10.1038/ncomms15706
539 Wu W, Liu Y, Wang Y, Li H, Liu J, Tan J, He J, Bai J, and Ma H. 2017.
540   Evolution analysis of the Aux/IAA gene family in plants shows dual
541   origins and variable nuclear localization signals. *International journal*
542   *of molecular sciences* 18. 10.3390/ijms18102107
543 Xu H, Song J, Luo H, Zhang Y, Li Q, Zhu Y, Xu J, Li Y, Song C, Wang B, Sun W,
544   Shen G, Zhang X, Qian J, Ji A, Xu Z, Luo X, He L, Li C, Sun C, Yan H,
545   Cui G, Li X, Li Xe, Wei J, Liu J, Wang Y, Hayward A, Nelson D, Ning Z,
546   Peters Reuben J, Qi X, and Chen S. 2016. Analysis of the Genome Sequence
547   of the Medicinal Plant Salvia miltiorrhiza. *Molecular Plant*
548   9:949-952. 10.1016/j.molp.2016.03.010
549 Zhang Y, Park C, Bennett C, Thornton M, and Kim D. 2021. Rapid and accurate
550   alignment of nucleotide conversion sequencing reads with HISAT-3N.
551   *Genome Research* 31:1290-1295.
552 Zheng Y, Jiao C, Sun H, Rosli Hernan G, Pombo Marina A, Zhang P, Banf M, Dai
553   X, Martin Gregory B, Giovannoni James J, Zhao Patrick X, Rhee Seung Y,
554   and Fei Z. 2016. iTAK: A Program for Genome-wide Prediction
555   and Classification of Plant Transcription Factors, Transcriptional
556   Regulators, and Protein Kinases. *Molecular Plant* 9:1667-1670.
557   10.1016/j.molp.2016.09.014
558

Figure 1

Figure 1. The gene structure and protein structure analysis of *SmIAAs*.

(A) The upper panel shows the gene structure of the *SmIAAs*. The CDS regions are represented by yellow boxes, respectively. The untranslated regions (UTRs) are represented by green boxes. The scale bar indicates the length of the gene structure in base pairs. (B) The protein structures of SmIAA proteins were predicted using the Pfam database. The AUX_IAA domain contained in SmIAA proteins is represented by a graded blue box. The evolutionary tree was constructed using the neighbor-joining (NJ) method, based on the multiple sequence alignment of SmIAA protein sequences. The evolutionary distances were calculated using the Poisson correction method. The whole figure was created using Tbtools, a software tool for the visualization and annotation of biological data.
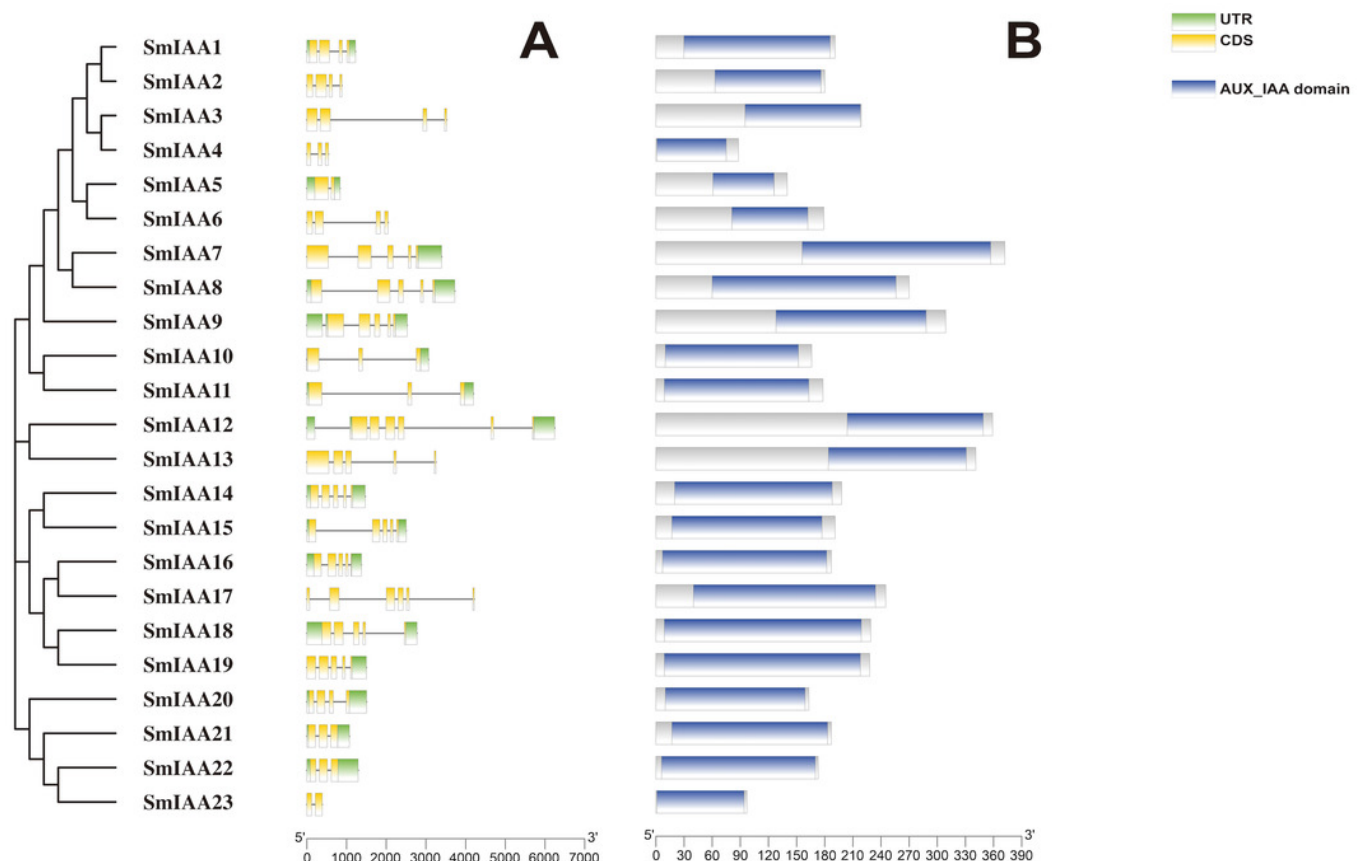
Figure 2

Figure 2. Multiple sequence alignment analysis of SmIAA family proteins.

Figure 2. Multiple sequence alignment analysis of SmIAA family proteins. The figure shows the results of a multiple sequence alignment analysis of SmIAA proteins. The alignment was performed using the Clustal Omega program, and the results were visualized using the Jalview program. The different regions of the proteins are color-coded, with the conserved domains indicated by different colored boxes. The numbers above the alignment indicate the position of each residue in the protein sequence. The alignment reveals that the SmIAA proteins share a high degree of sequence conservation in the conserved domains, including the AUX/IAA domain and the domain II/III. The variable regions of the proteins are less conserved, reflecting their divergent functions. The alignment also reveals the presence of several conserved motifs, which may be involved in protein-protein interactions or other functions.
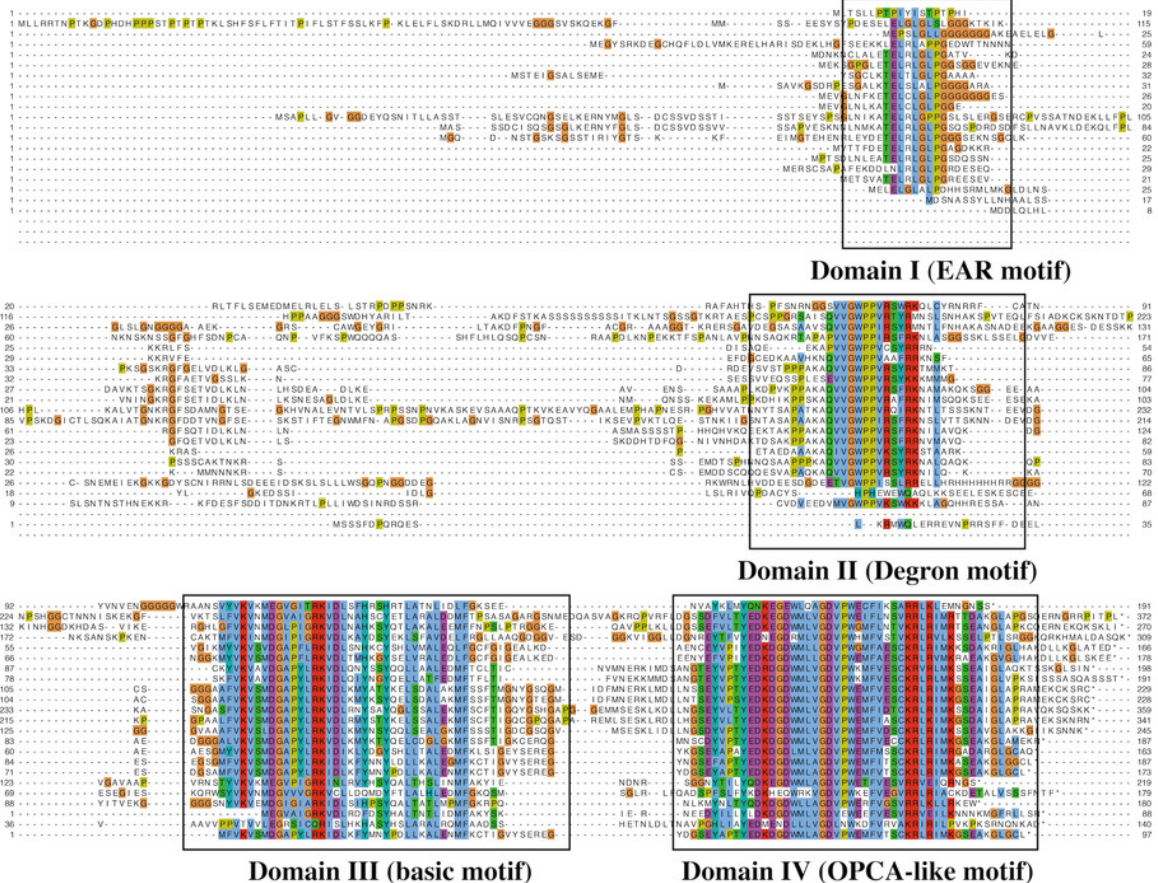
Figure 3

Figure 3. Phylogenetic tree and conserved motif analysis of Aux/IAA family genes in *A.thaliana, O.sativa, and S.miltiorrhiza.*

Figure 3. Phylogenetic tree and conserved motif analysis of Aux/IAA family genes in *A.thaliana, O.sativa, and S.miltiorrhiza.* (A) The evolutionary tree was constructed using the neighbor-joining (NJ) method, based on the multiple sequence alignment of Aux/IAA family protein sequences from Arabidopsis thaliana, Oryza sativa, and Salvia miltiorrhiza. The evolutionary distances were calculated using the Poisson correction method, and the numbers on the branches represent the bootstrap values (in percentage) from 500 replicates. The different clades are labeled with different colors, corresponding to the different species shown in the figure. All the sequences used to construct the evolutionary tree can be obtained from Supplementary Table 2. (B) The conserved motifs of the Aux/IAA proteins were identified using the MEME Suite software. The different motifs are represented by different colored boxes. The Motifs that are specially marked are related to the unique domain of Aux/IAA.

A

B

Salvia miltiorrhiza

Arabidopsis thaliana

Oryza sativa

Motif 4 (Domain I)
Motif 6
Motif 3 (Domain II)
Motif 2 (Domain III)
Motif 7
Motif 1 (Domain IV)
Motif 5
Motif 8
Motif 9
Motif 10
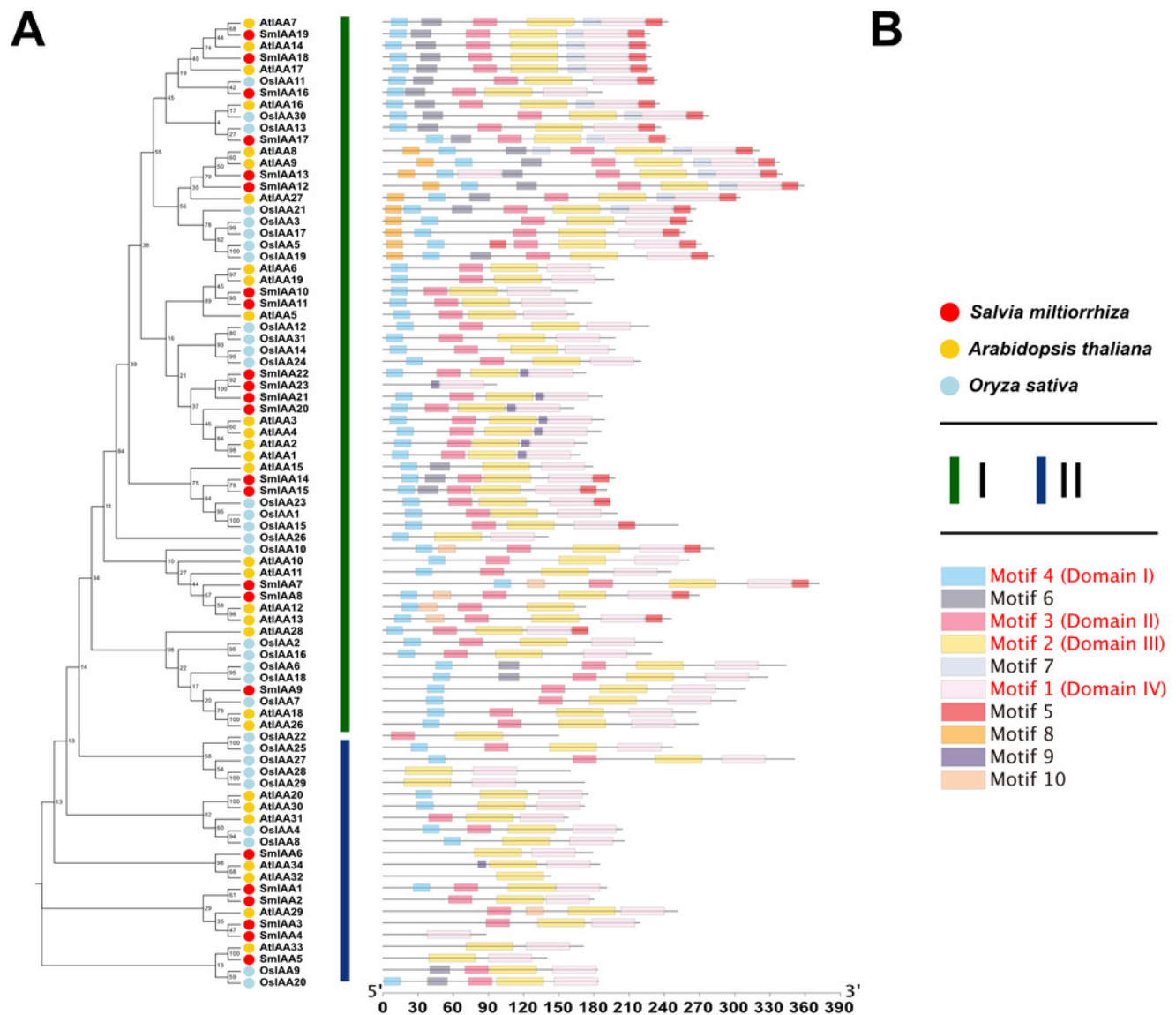
Figure 4

Figure 4. Collinearity and selection pressure analysis of *SmIAAs*.

Figure 4. Collinearity and selection pressure analysis of *SmIAAs*. The collinearity analysis of *SmIAAs* was performed using the MCScanX software, and the results are shown as syntenic blocks. The lines connecting the genes represent the syntenic relationships between them. The collinearity analysis revealed several syntenic gene pairs, indicating the occurrence of gene duplication events during the evolution of *SmIAAs*. The Ka/Ks ratio analysis was used to measure the selection pressure acting on each syntenic gene pair. The Ka/Ks ratios were calculated using the KaKs_Calculator software.

Ka/Ks

SmIAA2-SmIAA1: 0.2803
SmIAA3-SmIAA4: 0.1566
SmIAA3-SmIAA11: 0.3445
SmIAA13-SmIAA12: 0.1097
SmIAA8-SmIAA14: NaN
SmIAA8-SmIAA19: NaN
SmIAA8-SmIAA16: 0.2829
SmIAA8-SmIAA21: 0.1796
SmIAA16-SmIAA21: 0.2513
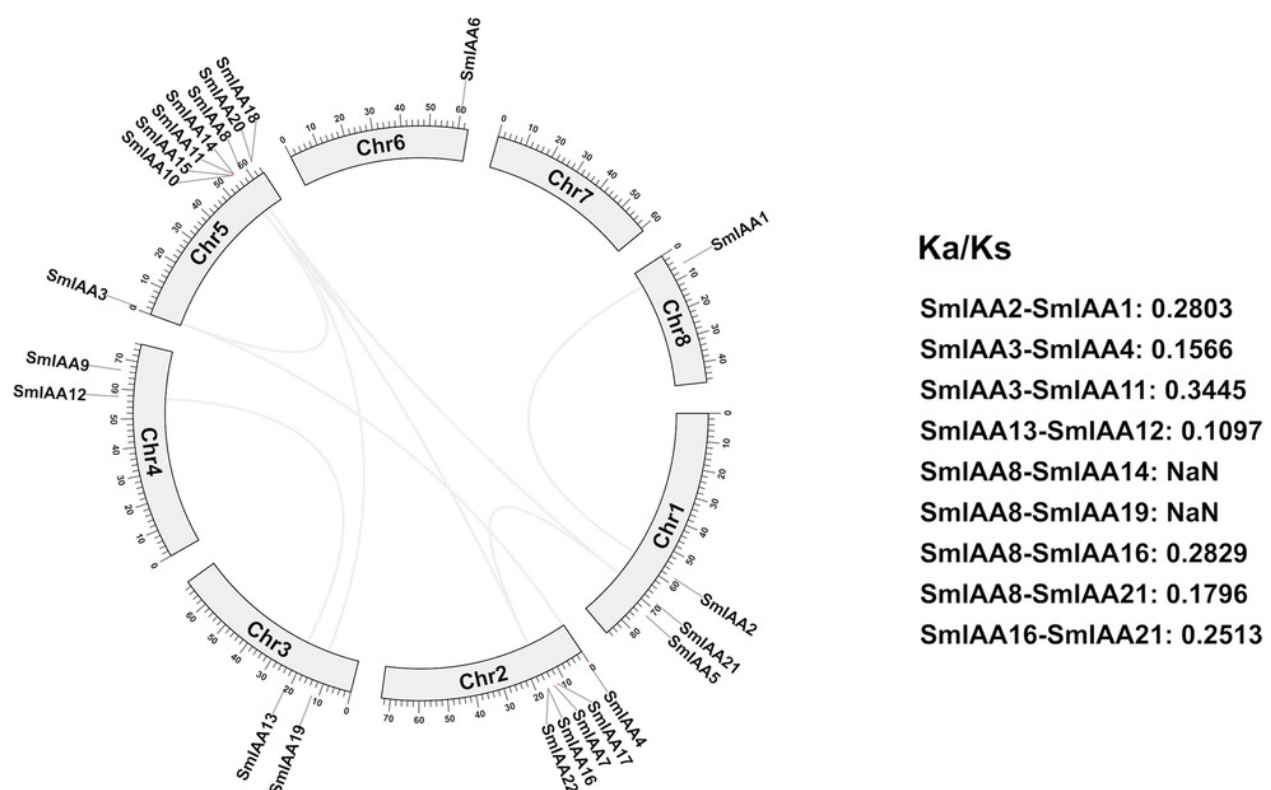
Figure 5

Figure 5. GO enrichment and various cis-element analysis of *SmIAAs.*

Figure 5. GO enrichment and various cis-element analysis of *SmIAAs.* (A) GO enrichment analysis was performed to investigate the biological functions of the *SmIAAs*. The GO terms are grouped into three main categories: biological process, cellular component, and molecular function. (B) The presence of various *cis*-elements in the promoter regions of the SmIAA genes was analyzed using the PlantCARE database. The different colored bars represent different *cis*-elements. The cis-elements are grouped into different categories, including hormone-related, stress-related, and development-related cis-elements.

A

GO Term

transcription regulator activity
DNA-binding transcription factor activity
identical protein binding
nucleus
intracellular membrane-bounded organelle
membrane-bounded organelle
intracellular organelle
organelle
response to auxin
response to organic substance
response to endogenous stimulus
response to hormone
regulation of DNA-templated transcription
regulation of nucleic acid-templated transcription
regulation of RNA biosynthetic process
response to chemical
DNA-templated transcription
nucleic acid-templated transcription

0.00 3.00 6.00 9.00 12.00 15.00 18.00
-log10(P-value)

- Molecular function
- Cellular component
- Biological process

B

SmIAA1
SmIAA2
SmIAA3
SmIAA4
SmIAA5
SmIAA6
SmIAA7
SmIAA8
SmIAA9
SmIAA10
SmIAA11
SmIAA12
SmIAA13
SmIAA14
SmIAA15
SmIAA16
SmIAA17
SmIAA18
SmIAA19
SmIAA20
SmIAA21
SmIAA22
SmIAA23

- Defense And Stress Responsiveness
- Meja-Responsiveness
- GA-Responsiveness
- Anaerobic Induction
- Circadian Control
- Light Responsiveness
- Auxin-Responsiveness
- SA-Responsiveness
- ABA-Responsiveness
- Cold-Responsiveness
- Zein Metabolism Regulation
- Meristem Expression
- Endosperm Expression
- Seed-Specific Regulation
- Cell Cycle Regulation

5' 0 200 400 600 800 1000 1200 1400 1600 1800 2000 3'

Figure 6. Expression profiles of *SmIAAs* in different tissues of *S. miltiorrhiza*.

Figure 6. Expression profiles of *SmIAAs* in different tissues of *S. miltiorrhiza*. The expression patterns of *SmIAAs* in different tissues of *S. miltiorrhiza* were analyzed using transcriptome data. Heat map visualization of the expression profiles was generated using TPM values, and row clustering was performed to group genes with similar expression patterns together. The heat map shows the expression levels of the *SmIAAs* across different tissues of *S. miltiorrhiza*, including root, stem, leaf, and flower. The color scale represents the relative expression level of each gene, with red indicating high expression and blue indicating low expression. The clustering of the genes reveals several groups with similar expression patterns, indicating potential functional similarities among these genes.
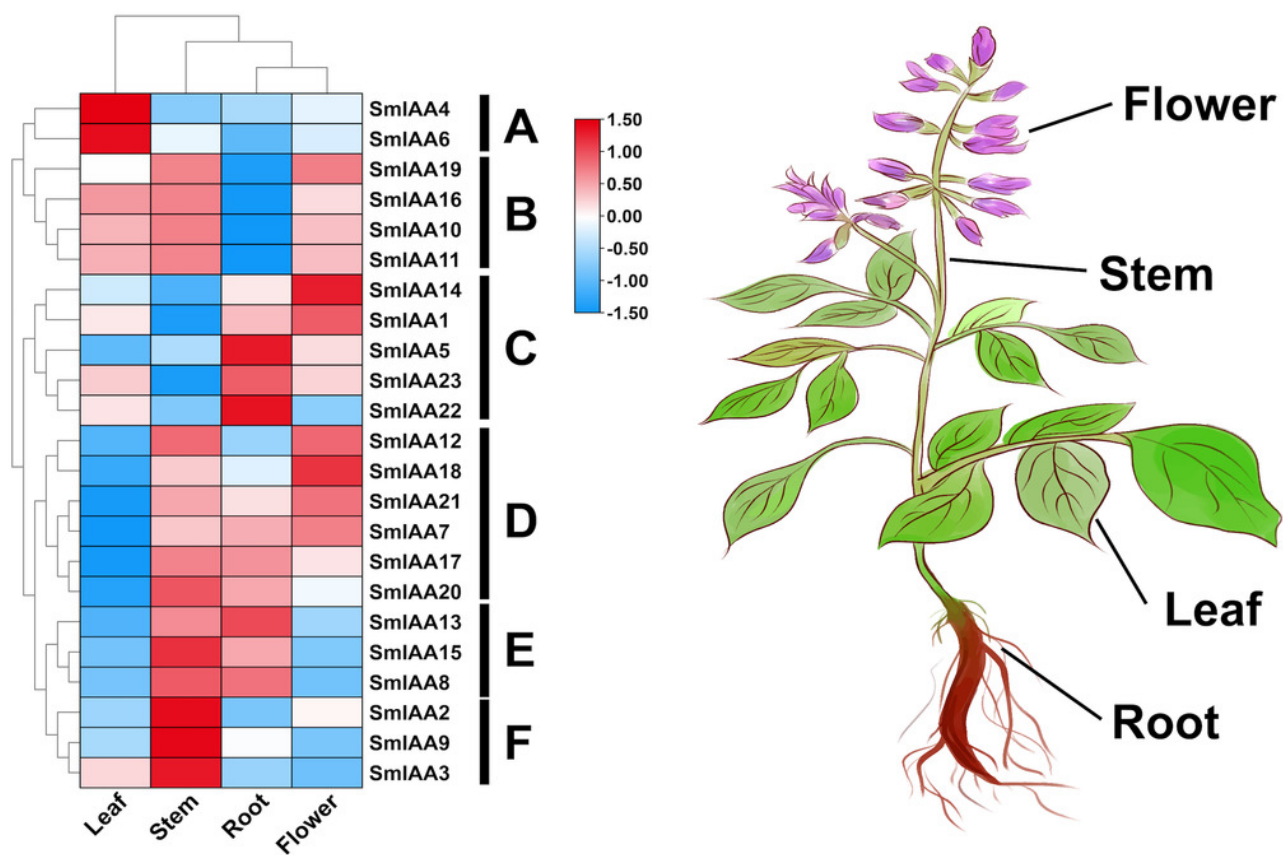
Figure 7

Figure 7. The correlation diagram between Module and SmIAAs calculated by WGCNA method.

Figure 7. The correlation diagram between Module and SmIAAs calculated by WGCNA method. To investigate the correlation between SmIAA genes and different gene modules, a Weighted Gene Co-expression Network Analysis (WGCNA) was performed based on transcriptome data from different tissues of *S.miltiorrhiza*.
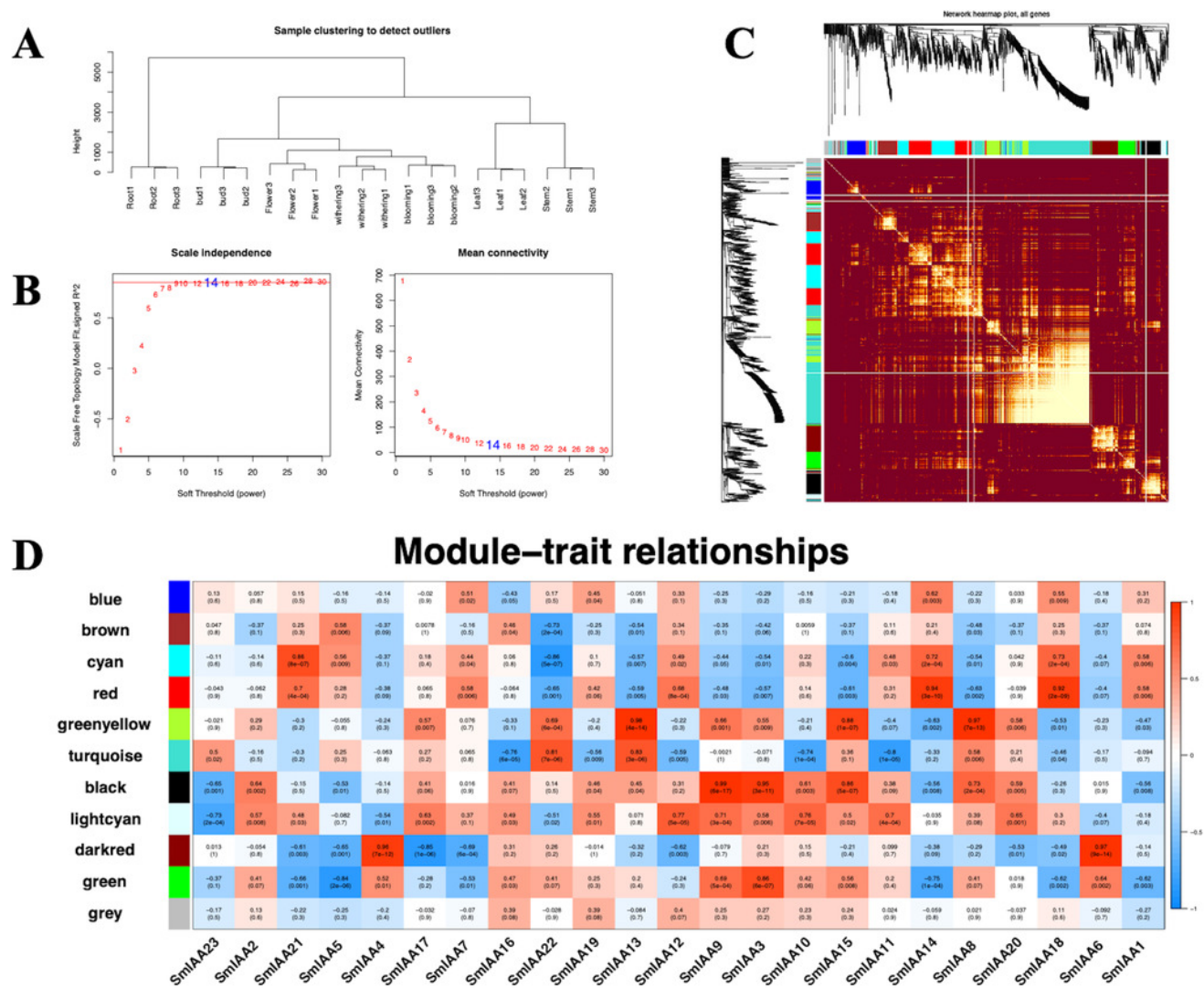
Figure 8

Figure 8. Correlation analysis between RNA-seq and qPCR data of 12 randomly selected *SmIAAs*.

Figure 8. Correlation analysis between RNA-seq and qPCR data of 12 randomly selected *SmIAAs*. To validate the RNA-seq data, quantitative PCR (qPCR) was performed on 12 randomly selected *SmIAAs*. The expression levels of these genes were quantified by both RNA-seq and qPCR, and the correlation coefficient (R^2) was calculated to compare the results. The results presented in this figure demonstrate a strong positive correlation ($R^2 > 0.8$) between the RNA-seq and qPCR data for all 12 genes, indicating that the RNA-seq data is reliable and accurate.
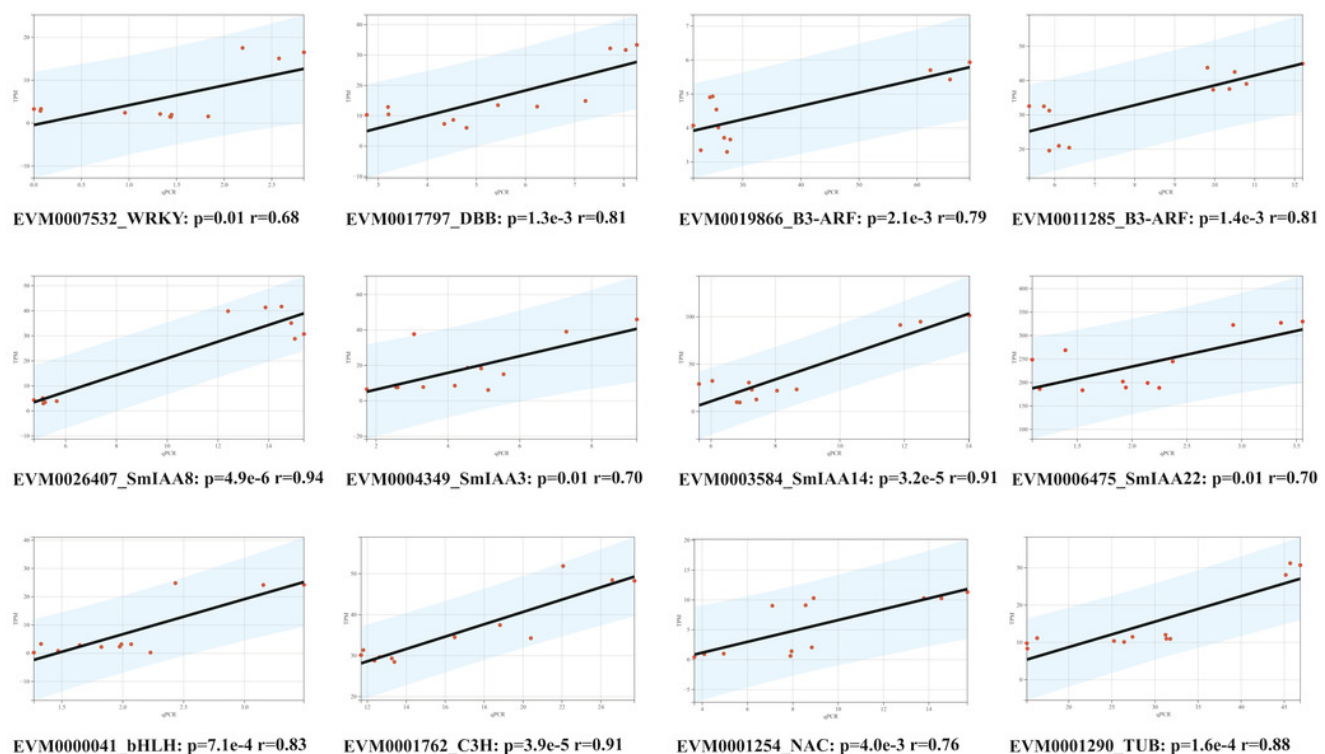
Table 1. Summarizes the physiological and biochemical features of SmIAAs.

The table includes information about the Aux/IAA gene family in *S. miltiorrhiza*, such as the number of members, molecular weight, isoelectric point, and the presence of conserved domains.

2 **Table 1. Summarizes the physiological and biochemical features of SmIAAs.**

Name	Gene ID	Accession ID	Chromosome Localization		Mw (Da)	pI	CDS (bp)	Length (aa)	Subcellular Localization	Homologs in Arabidopsis
SmIAA1	EVM0000296.1	GWHTAOSJ027615	Chr8	5920653	5921878	21688.8	9.81	573	190	cytoplasm	AT3G62100
SmIAA2	EVM0009395.1	GWHTAOSJ003786	Chr1	57848442	57849325	20565.38	8.58	540	179	nucleus	AT1G04550
SmIAA3	EVM0004349.1	GWHTAOSJ017483	Chr5	872484	876005	24827.71	5.94	657	218	nucleus	AT3G62100
SmIAA4	EVM0022630.1	GWHTAOSJ006473	Chr2	327881	328425	10364.81	4.82	264	87	cytoplasm	AT1G04550
SmIAA5	EVM0009745.1	GWHTAOSJ005170	Chr1	73323215	73324050	16049.36	9.25	420	139	cytoplasm	AT5G57420
SmIAA6	EVM0021654.1	GWHTAOSJ023551	Chr6	60358018	60360065	20499.12	5.16	537	178	nucleus	AT1G04550
SmIAA7	EVM0001033.1	GWHTAOSJ007657	Chr2	13722326	13725721	40197.45	9.38	1116	371	nucleus	AT2G46990
SmIAA8	EVM0026407.1	GWHTAOSJ019954	Chr5	58348218	58351945	28501.2	8.64	810	269	cytoplasm	AT2G33310
SmIAA9	EVM0002334.1	GWHTAOSJ016330	Chr4	66298635	66301163	34131.41	9	927	308	nucleus	AT5G25890
SmIAA10	EVM0010171.1	GWHTAOSJ019616	Chr5	55028101	55031170	18580.66	8.17	498	165	cytoplasm	AT1G15580
SmIAA11	EVM0027652.1	GWHTAOSJ019676	Chr5	55711914	55716112	19848.71	5.41	534	177	cytoplasm	AT1G15580
SmIAA12	EVM0010443.1	GWHTAOSJ015659	Chr4	57556179	57562422	38312.09	7.6	1077	358	nucleus	AT3G23050
SmIAA13	EVM0025242.1	GWHTAOSJ011536	Chr3	21196572	21199823	36647.47	8.6	1023	340	nucleus	AT3G04730
SmIAA14	EVM0003584.1	GWHTAOSJ019677	Chr5	55738962	55740433	21585.91	8.15	594	197	cytoplasm	AT1G80390
SmIAA15	EVM0019034.1	GWHTAOSJ019617	Chr5	55050465	55052967	20718.73	7.72	573	190	nucleus	AT5G43700
SmIAA16	EVM0009041.1	GWHTAOSJ007792	Chr2	15980426	15981800	20799.85	6.73	561	186	cytoplasm	AT4G14550
SmIAA17	EVM0020724.1	GWHTAOSJ007565	Chr2	12458073	12462291	26535.11	8.96	735	244	nucleus	AT3G04730
SmIAA18	EVM0011993.1	GWHTAOSJ020429	Chr5	62919943	62922717	24660.27	6.02	687	228	cytoplasm	AT4G14550
SmIAA19	EVM0021151.1	GWHTAOSJ010857	Chr3	12283059	12284558	25314.25	7.54	684	227	nucleus	AT4G14550
SmIAA20	EVM0010879.1	GWHTAOSJ020426	Chr5	62894169	62895673	17843.23	5.31	489	162	nucleus	AT5G43700
SmIAA21	EVM0004282.1	GWHTAOSJ004659	Chr1	68141757	68142829	20661.38	6.61	561	186	nucleus	AT5G43700
SmIAA22	EVM0006475.1	GWHTAOSJ007793	Chr2	16036879	16038173	19453.26	5.39	519	172	cytoplasm	AT5G43700
SmIAA23	EVM0007611.1	GWHTAOSJ000893	contig685	54663	55050	10916.64	5.09	291	96	cytoplasm	AT4G14560

Table 2. Different transcription factors in the *S. miltiorrhiza* genome

Table 2. Different transcription factors in the *S. miltiorrhiza* genome This table presents the different transcription factors identified in the genome of *S. miltiorrhiza* using iTAK tool.

2 **Table 2. Different transcription factors in the *S. miltiorrhiza* genome**

3 This table presents the different transcription factors identified in the genome of *S. miltiorrhiza*
4 using iTAK tool.

Total			1814				
Class	**Num**	**Class**	**Num**	**Class**	**Num**	**Class**	**Num**
Alfin-like	7	CAMTA	4	HB-other	12	OFP	25
AP2/ERF-AP2	19	CPP	8	HB-PHD	2	PLATZ	12
AP2/ERF-ERF	143	CSD	4	HB-WOX	15	RWP-RK	17
AP2/ERF-RAV	3	DBB	5	HRT	1	S1Fa-like	2
B3	44	DBP	2	HSF	34	SAP	2
B3-ARF	26	DDT	8	LFY	1	SBP	17
BBR-BPC	6	E2F-DP	8	LIM	10	SRS	9
BES1	9	EIL	7	LOB	50	STAT	1
bHLH	134	FAR1	43	MADS-M-type	25	TCP	29
bZIP	67	GARP-ARR-B	12	MADS-MIKC	46	Tify	11
C2C2-CO-like	14	GARP-G2-like	48	MYB	145	Trihelix	31
C2C2-Dof	34	GeBP	10	MYB-related	65	TUB	12
C2C2-GATA	28	GRAS	62	NAC	84	ULT	1
C2C2-LSD	5	GRF	12	NF-X1	3	VOZ	2
C2C2-YABBY	9	HB-BELL	15	NF-YA	8	Whirly	2
C2H2	110	HB-HD-ZIP	43	NF-YB	21	WRKY	78
C3H	56	HB-KNOX	11	NF-YC	9	zf-HD	16

Table 3. Top 10 Hub genes in the black and cyan modules

This table presents the top 10 hub genes in the black and cyan modules. The hub genes were identified using the weighted gene co-expression network analysis (WGCNA) method. The table includes information about the module color, gene name, gene annotation, correlation, p-value, and the K-MEANS-based connectivity (KME).

Table 3. Top 10 Hub genes in the black and cyan modules

Group	Gene	Annotation	R	P	KME
black	EVM0007532	WRKY	0.9931	3.72E-19	0.9931
black	EVM0000041	bHLH	0.9929	5.06E-19	0.9929
black	EVM0019585	zf-HD	0.9865	2.14E-16	0.9865
black	EVM0012769	MYB	0.9857	3.78E-16	0.9857
black	EVM0010344	bHLH	0.9851	5.29E-16	0.9851
black	EVM0018729	HB-WOX	0.9840	1.05E-15	0.9840
black	EVM0021887	AP2/ERF-ERF	0.9800	8.72E-15	0.9800
black	EVM0023258	SBP	0.9797	9.96E-15	0.9797
black	EVM0015830	zf-HD	0.9791	1.32E-14	0.9791
black	EVM0011877	bHLH	0.9737	1.15E-13	0.9737
cyan	EVM0017797	DBB	0.9906	7.35E-18	0.9906
cyan	EVM0019866	B3-ARF	0.9763	4.30E-14	0.9763
cyan	EVM0001762	C3H	0.9744	9.03E-14	0.9744
cyan	EVM0004701	MYB	0.9731	1.45E-13	0.9731
cyan	EVM0002416	HSF	0.9697	4.34E-13	0.9697
cyan	EVM0011962	bZIP	0.9662	1.23E-12	0.9662
cyan	EVM0004337	MYB-related	0.9645	1.91E-12	0.9645
cyan	EVM0016656	C2H2	0.9607	5.02E-12	0.9607
cyan	EVM0025053	MYB-related	0.9584	8.45E-12	0.9584
cyan	EVM0000238	C2C2-Dof	0.9583	8.73E-12	0.9583

Table 4. The relationship between different modules and different tissues of *S. miltiorrhiza*

This table shows the relationship between different modules and different tissues of *S. miltiorrhiza.* The modules were identified using the weighted gene co-expression network analysis (WGCNA) method. The table includes information about the module color, *SmIAAs* in each module, and the tissues in which the module was highly expressed.

Table 4. The relationship between different modules and different tissues of _Salvia_
miltiorrhiza

Module	R	Gene	Tissue
green	-0.62	SmIAA1	Flower
darkred	-0.69	SmIAA7	Flower
lightcyan	0.77	SmIAA12	Flower
red	0.94	SmIAA14	Flower
red	0.92	SmIAA18	Flower
cyan	0.86	SmIAA21	Flower
darkred	0.96	SmIAA4	Leaf
darkred	0.97	SmIAA6	Leaf
green	-0.84	SmIAA5	Root
cyan	-0.86	SmIAA22	Root
lightcyan	-0.73	SmIAA23	Root
black	0.64	SmIAA2	Stem
black	0.95	SmIAA3	Stem
black	0.99	SmIAA9	Stem
lightcyan	0.76	SmIAA10	Stem&Flower
turquoise	-0.8	SmIAA11	Stem&Flower
turquoise	-0.76	SmIAA16	Stem&Flower
turquoise	-0.56	SmIAA19	Stem&Flower
greenyellow	0.97	SmIAA8	Stem&Root
greenyellow	0.98	SmIAA13	Stem&Root
greenyellow	0.88	SmIAA15	Stem&Root
darkred	-0.85	SmIAA17	Stem&Root
lightcyan	0.65	SmIAA20	Stem&Root