BMC Plant Biology

RESEARCH

Open Access

Genomic survey and expression analysis of *LcARFs* reveal multiple functions to somatic embryogenesis in *Liriodendron*

Lin Xu[1†], Ye Liu[1†], Jiaji Zhang[1], Weihuang Wu[1], Zhaodong Hao[1], Shichan He[1], Yiran Li[1], Jisen Shi[1*] and Jinhui Chen[1*]

Abstract

Background Auxin response factors (ARFs) are critical transcription factors that mediate the auxin signaling pathway and are essential for regulating plant growth. However, there is a lack of understanding regarding the *ARF* gene family in *Liriodendron chinense*, a vital species in landscaping and economics. Thus, further research is needed to explore the roles of *ARFs* in *L. chinense* and their potential applications in plant development.

Result In this study, we have identified 20 *LcARF* genes that belong to three subfamilies in the genome of *L. chinense*. The analysis of their conserved domains, gene structure, and phylogeny suggests that *LcARFs* may be evolutionarily conserved and functionally similar to other plant *ARFs*. The expression of *LcARFs* varies in different tissues. Additionally, they are also involved in different developmental stages of somatic embryogenesis. Overexpression of *LcARF1*, *LcARF2a*, and *LcARF5* led to increased activity within callus. Additionally, our promoter-GFP fusion study indicated that *LcARF1* may play a role in embryogenesis. Overall, this study provides insights into the functions of *LcARFs* in plant development and embryogenesis, which could facilitate the improvement of somatic embryogenesis in *L. chinense*.

Conclusion The research findings presented in this study shed light on the regulatory roles of *LcARFs* in somatic embryogenesis in *L. chinense* and may aid in accelerating the breeding process of this tree species. By identifying the specific *LcARFs* involved in different stages of somatic embryogenesis, this study provides a basis for developing targeted breeding strategies aimed at optimizing somatic embryogenesis in *L. chinense*, which holds great potential for improving the growth and productivity of this economically important species.

Keywords *ARF* genes, *Liriodendron chinense*, Somatic embryogenesis

Introduction

Liriodendron chinense is an angiosperm species with leaves resembling traditional Chinese garments and cup-shaped flowers, making it a significant ornamental, economic, and forestry tree [1–3]. Crossbreeding is a crucial method to improve the genetic diversity of *Liriodendron* and cultivate hybrids with superb forest characteristics [4]. However, its low natural seed setting rate, difficult rooting during cutting propagation, and low survival rate of grafting breeding impede this species' popularization and application. Somatic embryogenesis,

†Lin Xu and Ye Liu contributed equally.

*Correspondence:
Jisen Shi
jshi@njfu.edu.cn
Jinhui Chen
chenjh@njfu.edu.cn
[1] Key Laboratory of Forest Genetics and Biotechnology, Ministry of Education of China, Co-Innovation Center for the Sustainable Forestry in Southern China, Nanjing Forestry University, No.159 Longpan Road, Nanjing 210037, China

a cost-saving procedure, can be employed to propagate hybrid plants while maintaining their superior traits [5–7]. Somatic embryogenesis for *Liriodendron* involves a multi-step regeneration process, including embryonic callus induction, somatic embryo induction and maturation, and plantlet germination [5, 7]. In order to successfully generate somatic embryos, it is crucial to regulate the development of the embryo, particularly the transition of cell fate [8, 9].

Plant hormones have a significant influence on somatic embryogenesis. Among them, auxin plays the most crucial role in acquiring embryogenic potential [10]. It is achieved by controlling gene expression via *ARFs*, which bind to cis-elements in downstream genes by utilizing two distinct DNA-binding domains [11]. *ARF* genes encode auxin response transcription factors that bind to TGTCTC auxin responsive elements in the promoters of early auxin response genes. A typical ARF protein contains several functional domains: a plant-specific DNA-binding domain (DBD) at the N-terminus, transcriptional activation or repression domain in the middle region (MR), and a C-terminal dimerization domain (CTD) in most cases [12]. The DBD has a B3-type DBD followed by dimerization domains (DD). DD can cause *ARFs* to form dimers, essential for binding to target DNA [11]. The CTD is related to motifs III and IV found in Aux/IAA (Auxin /Indole-3-Acetic Acid) proteins [13]. *Aux / IAA* family members can form dimers with *ARFs* via the CTD domain [14]. *ARFs* bind to *Aux/IAA*, which inhibits its activation on auxin responsive genes at low auxin concentrations. Under high auxin levels, auxin acts as a molecular glue between the TIR1/AFB (Transport Inhibitor Resistant 1/Auxin Signaling F-box) receptor and Aux/IAA protein, leading to subsequent ubiquitination and degradation of Aux/IAA protein, releasing *ARFs* from inhibition. Therefore, *ARFs* are crucial in the auxin regulatory pathway by converting chemical signals into gene regulation [15].

ARF proteins play important roles in various aspects of plant growth and development, including the development of roots, flowers, and embryogenesis. In *Arabidopsis*, *AtARF7* and *AtARF19* control lateral root formation through interaction with three IAA proteins (IAA3, IAA14, IAA18) [16]. *AtARF6* and *AtARF8* regulate petal expansion, stamen filament elongation, anther dehiscence, and pistil maturation, which ensure that pollen released from anthers is deposited on the stigma of the recipient [17]. In the early stages of zygotic embryogenesis, auxin is transferred from basal cells to apical cells, which induces embryo development. In later stages, the direction of auxin flux is reversed, leading to auxin accumulation in the hypothesis and triggering the onset of root meristems [18]. Many *ARFs* are expressed and involved in zygotic embryogenesis. At the spherical embryo stage in *Arabidopsis*, *AtARF1* is ubiquitous at the globular stage, *AtARF5* marks the lower tier of the embryo, but *AtARF2/9* are expressed in all suspensor cells and the lower-tier of protoderm cells [19]. At the heart embryo stage, *AtARF1/2* are ubiquitously expressed, and *AtARF5* is active in subdomains of the vascular tissue. *AtARF9* is expressed in the presumptive root meristem [19]. *AtARF9*, redundantly with *AtARF13*, mediates the differentiation of suspensor cells and prevents its transition to the embryo [20]. *AtARF5/7* interact with each other physically and control both axis formation in the embryo and auxin-dependent cell expansion [21, 22].

Some researches reveal that *ARFs* also play the same role during somatic embryogenesis as in zygotic embryogenesis. Extensive expression of many *AtARFs* in *Arabidopsis* embryogenic culture suggests that auxin signaling may play an important role in somatic embryogenesis induction [23]. In rice, down-regulation of *OsARF5*, a positive regulator of stem cell maintenance and meristem development, may account for better somatic embryo regeneration and differentiation in japonica than in indica [24]. Some studies in *Arabidopsis* have indicated that *AtARFs* are involved in acquiring embryonic competence in somatic embryogenesis [23]. *Liriodendron* is a noteworthy species in the fields of horticulture and economics. However, traditional breeding methods have proven to be time-consuming and inadequate in meeting the demands of current production. To address this issue, we have employed somatic embryogenesis as a breeding strategy to achieve rapid plant propagation. In this process, auxin serves as a crucial factor in regulating the efficiency of somatic embryogenesis. Isolating and identifying relevant genes, as well as thoroughly investigating the expression patterns of target genes during somatic embryogenesis, will greatly enhance the efficiency of this process and expedite the breeding of *Liriodendron*. Therefore, our study aims to isolate an important auxin signal transcription factor - auxin response factor (*ARF*) - from the entire genome of *Liriodendron*. The isolation and identification of this gene family are crucial for elucidating the mechanism underlying somatic embryogenesis in *Liriodendron*.

Results

Genome-wide identification of *ARFs* in *L. chinense*

ARF protein sequences from *Arabidopsis* as reference, 20 candidate *ARFs* were identified in *L. chinense* genome by BLASTP and HMMER programs (see methods). The *LcARF* members were named based on the homology with *AtARFs* in this study (Fig. 1 and Table 1). The molecular weight of these LcARFs was between 47.46 kDa and

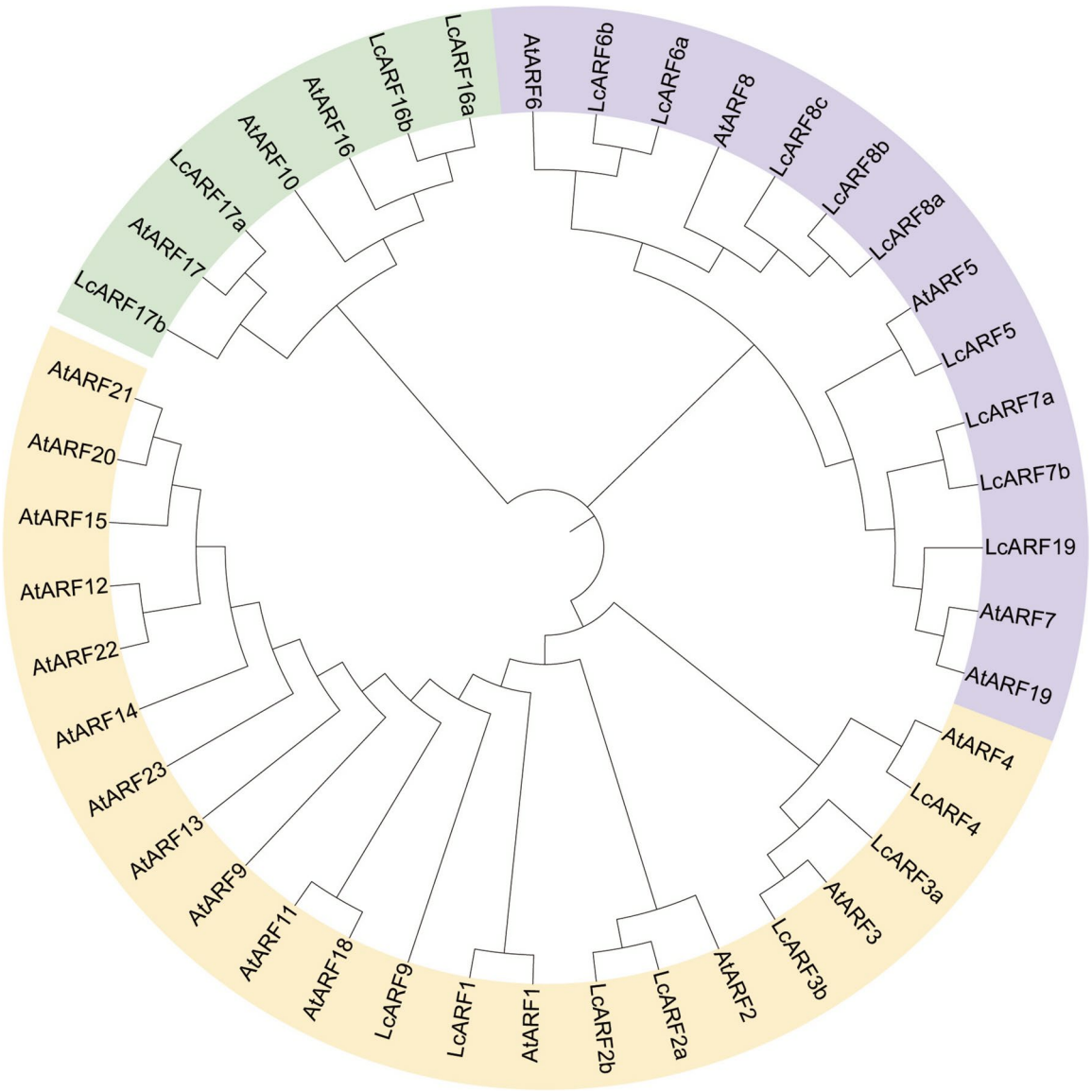

Fig. 1 Phylogenetic tree of ARF proteins in *Arabidopsis* and *L. chinense*. The ARFs can be classified into three major classes based on their phylogenetic relationship. The different-colored areas represent distinct classes within the ARF family

133.46 kDa. The theoretical PIs of these LcARFs were 4.6 ~ 9.13 (Table 1), indicating that most *LcARFs* encoded weakly acidic proteins. All *LcARFs* were predicted to localize in nucleus (Table 1).

19 *LcARFs* were distributed on 11 chromosomes of *L. chinense*, except *LcARF17a* was mapped to a separate scaffold not yet assembled into full chromosomes (Fig. S1). There were 4 *LcARFs* mapped on chromosome 3, which contains most of the *LcARFs*. The rest *LcARFs* per chromosome varied from 0 to 4, with no apparent correlation between chromosomal length and the number of *LcARFs* present.

One way for organisms to acquire new genes is gene duplication. We found that a few *LcARFs* had extremely high sequence similarities (Fig. 1). *LcARF2a-2b*, *LcARF3a-3b*, *LcARF6a-6b*, *LcARF7a-7b*, *LcARF16a-16b* and *LcARF17a-17b* form twin pairs, and *LcARF8a/b/c* from triplet pairs. Surprisingly, no tandem duplication events were found because most of the duplicated *LcARFs* were located on different chromosomes, or they are separated by at least a few million bases if on one chromosome (Table 1).

Table 1 Physicochemical properties and chromosomal details of LcARF proteins

Gene ID	Gene name	Number of amino acids	Molecular weight	Theoretical pI	Formula	Total number of atoms	Instability index	Aliphatic index	Grand average of hydropathicity	Predicted location(s)
Lchi01654	LcARF1	590	65956.41	5.69	$C_{2901}H_{4506}N_{814}O_{888}S_{30}$	9139	55.89	70.54	-0.469	Nucleus.
Lchi16424	LcARF2a	740	82677.75	5.64	$C_{3610}H_{5599}N_{1023}O_{1129}S_{40}$	11401	56.69	59.03	-0.716	Nucleus.
Lchi02396	LcARF2b	763	85164.6	6.03	$C_{3746}H_{5823}N_{1061}O_{1157}S_{29}$	11816	58.86	65.52	-0.64	Nucleus.
Lchi09892	LcARF3a	597	65843.48	4.86	$C_{2911}H_{4450}N_{790}O_{912}S_{23}$	9086	45.95	73.79	-0.394	Nucleus.
Lchi08511	LcARF3b	683	75266.67	6.32	$C_{3346}H_{5146}N_{930}O_{1008}S_{23}$	10453	57.66	71.79	-0.451	Nucleus.
Lchi19395	LcARF4	731	81623.34	4.6	$C_{3505}H_{5451}N_{997}O_{1170}S_{42}$	11165	47.58	66.94	-0.662	Nucleus.
Lchi29014	LcARF5	424	47755.46	8.96	$C_{2134}H_{3307}N_{599}O_{614}S_{18}$	6672	52.87	77.03	-0.333	Nucleus.
Lchi18918	LcARF6a	928	102934.55	6.14	$C_{4538}H_{7045}N_{1293}O_{1389}S_{31}$	14296	67.57	78.36	-0.419	Nucleus.
Lchi18919	LcARF6b	314	35268.28	8.84	$C_{1569}H_{2444}N_{446}O_{452}S_{15}$	4926	62.55	77.68	-0.321	Nucleus.
Lchi10785	LcARF7a	1193	133463.24	9.13	$C_{5837}H_{9272}N_{1722}O_{1770}S_{49}$	18650	64.17	77.42	-0.566	Nucleus.
Lchi05212	LcARF7b	902	98651	5.52	$C_{4315}H_{6760}N_{1216}O_{1352}S_{42}$	13685	56.4	75.04	-0.375	Nucleus.
Lchi12475	LcARF8a	601	66621.31	5.84	$C_{2966}H_{4598}N_{814}O_{893}S_{21}$	9292	56.29	80.42	-0.301	Nucleus.
Lchi12473	LcARF8b	601	66693.38	5.76	$C_{2969}H_{4602}N_{814}O_{895}S_{21}$	9301	56.95	80.42	-0.306	Nucleus.
Lchi28544	LcARF8c	452	49572.42	4.94	$C_{2176}H_{3359}N_{601}O_{687}S_{20}$	6843	49.87	72.08	-0.309	Nucleus.
Lchi09822	LcARF9	744	83171.29	7.96	$C_{3662}H_{5779}N_{1051}O_{1107}S_{29}$	11628	57.78	73.64	-0.512	Nucleus.
Lchi08159	LcARF16a	692	75991.06	7.52	$C_{3343}H_{5235}N_{947}O_{1015}S_{33}$	10573	48.01	73.11	-0.368	Nucleus.
Lchi05394	LcARF16b	582	64144.49	6.1	$C_{2816}H_{4412}N_{802}O_{861}S_{27}$	8918	51.13	74.4	-0.368	Nucleus.
Lchi33157	LcARF17a	572	62416.55	8.26	$C_{2780}H_{4282}N_{776}O_{823}S_{21}$	8682	47.11	68.99	-0.295	Nucleus.
Lchi20149	LcARF17b	445	48464.55	5.41	$C_{2137}H_{3314}N_{578}O_{668}S_{21}$	6718	45.92	67.87	-0.421	Nucleus.
Lchi11431	LcARF19	570	64248.68	6.34	$C_{2853}H_{4412}N_{796}O_{851}S_{24}$	8936	56.53	75.26	-0.394	Nucleus.

Phylogenetic analysis of *LcARFs*

We expanded the neighbor-joining analysis to include *ARF* sequences from other taxa to gain insight into the evolutionary relationship between *LcARFs* and their homologs in other plant species. Including typical species of algae, bryophytes, herbs, woody plants, monocots, dicots, etc.; 4 ARF sequences from the liverwort *Marchantia polymorpha*, 15 sequences from the moss *Physcomitrella patens*, 23 sequences from rice, 17 sequences from *Vitis vinifera*, 21 sequences from *Theobroma cacao*, 25 sequences from *Solanum Lycopersicum*, 13 sequences from *Amborella trichopoda* and 23 from *Arabidopsis* (Fig. 2). A B3 domain-containing sequence from green algae *Chlamydomonas reinhardtii* was used as an outgroup. In this analysis, 162 ARF sequences were included in a moderately well-supported phylogenetic tree. Cre13.g562400 from *C. reinhardtii* was the most divergent gene as it showed

Fig. 2 Unrooted Classification tree representing relationships among *ARF* genes of 10 species. A total of 162 ARF protein sequences from 10 species were selected to construct a Bayesian phylogenetic tree. Different color blocks represent different evolutionary branches

sequence similarity to *ARF* genes in the B3 DNA-binding region and lackeds other motifs. Consistent with analyses by Mutte et al. (2018) and Finet et al. (2013), the *ARFs* can be classified into three major clades on their phylogenetic relationship; class A, class B, and class C (Fig. 2).

We further analyzed the phylogeny of the *ARF* family in more detail (Fig. 2). Class A contained four subclades, and the guide branch percentages well supported the relationship among the subclades in clade A. ARF5 subclade, ARF6/8 subclade and ARF7 subclade are the A-ARFs in land plants, while contains the proto-A-ARF subclade in bryophytes. Node e is the foundation of the other four subsections, indicating that they may have very conserved domains representing the A-ARF class precursors in plants. Branch B can be further divided into 6 subclades, including node f (ARF11 / 18 subclade), g (ARF13 subclade), h (ARF1 subclade), i (ARF2 subclade), j (ARF3/4 subclade), and k (bryophytes B-ARF). Remarkably, node k was the foundation of the five subsections: f, g, h, i, and j, indicating that they may represent precursors of the B-ARF class in land plants. Node g subsection contains all members of the *Arabidopsis* ARF family, which indicates that ARF may be repeated and diversified in eukaryotic plants. Branch C consists of node l (ARF16 subclade), m (bryophytes C-ARF), and n (ARF17 subclade). As observed in clades A and B, we found that the node m daughter clade was only represented by bryophyte members, meaning they may be a group of proto-C-ARF.

Conserved domain analysis of LcARFs

Most ARF proteins contain a conserved N-terminal DNA-binding domain (DBD) composed of plant-specific B3-type and auxin responsive motifs, and a highly conserved C-terminal CTD domain corresponding to motif III and IV of the Aux/IAA proteins. To better understand the structural similarity of these ARFs, we constructed a neighbor-joining phylogenetic tree using the amino acids sequences of 20 LcARFs, resulting in 3 major *ARF* classes, subclade A-C (Figs. 3A, S2). The results demonstrated that LcARF4, LcARF8c, LcARF16b, and LcARF17b do not have a DBD domain. It is possible that the DNA-binding ability of these proteins could be impaired due to lack of a DNA-binding domain. The remaining LcARFs contained highly conserved and complete DBD domains. In addition, most LcARFs had conserved and intact CTD domains, except for LcARF3b, LcARF5, LcARF6b, LcARF16b and LcARF17a/b. Due to the lack of the AUX_IAA binding domain, the expression of these proteins may not be regulated by AUX/IAA.

Unlike the DBD and CTD regions are conserved, the protein sequences of the MRs (the middle region) are highly variable. Six LcARFs not only lacked AUX_IAA domain, but also l MR domain (Fig. 3A). It has been proposed that ARF proteins whose middle regions are enriched in Ser, Pro, Thr, and Gly, might act as transcriptional repressors, whereas those are enriched in Gln and Leu in MR might act as transcriptional activators. The MRs of LcARF2a, LcARF6a, and LcARF7a were abundant in Gln and Leu residues, suggesting they might be transcriptional activators (Fig. 3B). Ser, Pro, Thr, and Gly were abundant in the MRs of LcARF1, LcARF8a and

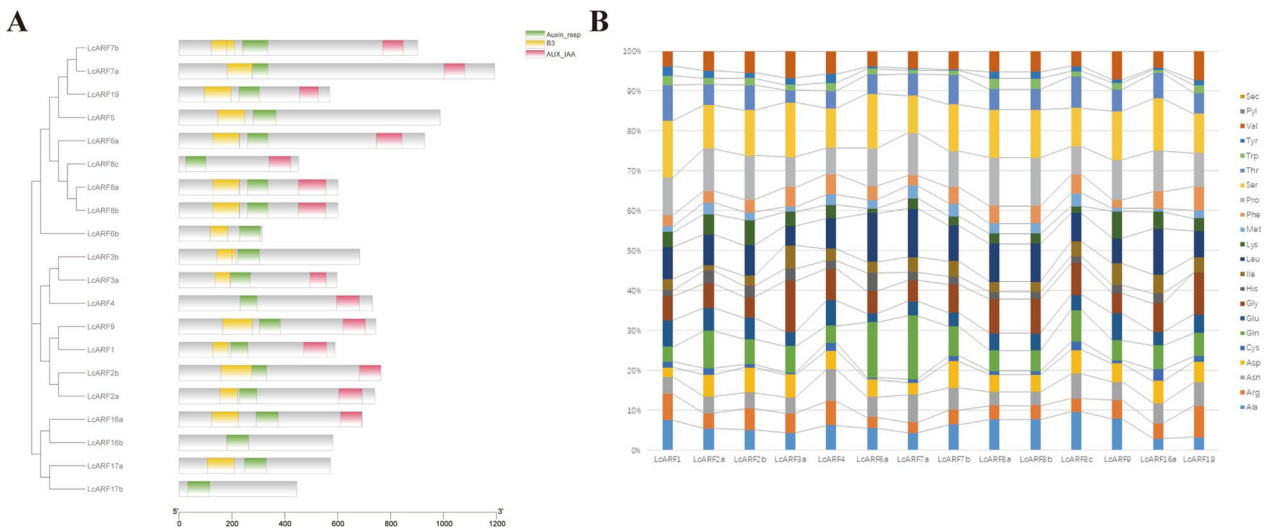

Fig. 3 Analysis of conserved domains of *LcARF* gene family. **A** Schematic organization of conserved domains in LcARF proteins. **B** Amino acid composition of MR domains in LcARF proteins, bars represent the percentage of different amino acid residues in MR domains of LcARFs

LcARF8b, suggesting that they might be transcriptional repressors (Fig. 3B). In summary, we speculated that the conserved domains of LcARFs were beneficial to ensuring the execution of its essential functions. In contrast, the diversity of MR domains helped LcARFs respond to different environments' adaptability.

Gene structural analysis of *LcARFs*

In addition to conserved protein domains, the pattern of intron-exon positions between ARF subgroups can also provide clues on evolutionary relationships. To identify the intron-exon structure of individual *LcARFs*, an alignment of the full-length cDNA sequences with the corresponding genomic DNA sequences was performed. For the members in the *LcARF1/2a/2b/5/6a/7a/7b/8a/8b/9* clade had 11~15 introns, members of *LcARF3a/3b/4/6b/8c/19* clade had 7~10 introns, and members of *LcARF16a/16b/17a/17b* clade had 1~4 introns (Fig. 4). Consistent with previous findings from *Arabidopsis* and rice, members of *LcARF16/17* clade had a relatively lower number of introns. Strikingly, the exon-intron structure was conserved among subclades of homologous genes and differed between subclades

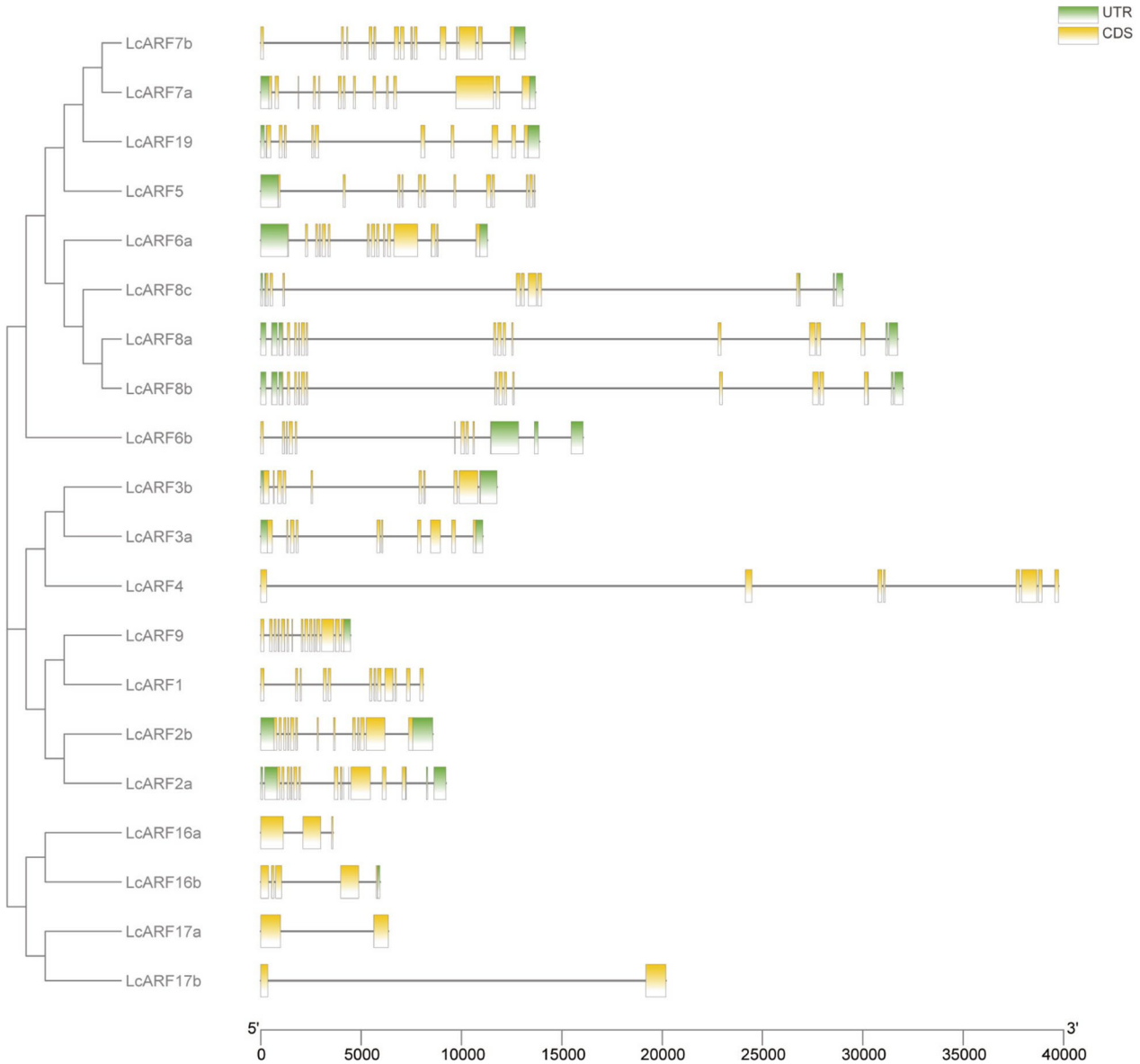

Fig. 4 Analysis of intron-exon organization of *LcARF* gene family. The intron-exon organization of *LcARF* genes was plotted using Tbtools (Version 1.09832)

of non-homologous genes. The phylogenetic analysis of LcARF proteins also supported this.

Numbers of cis-acting elements in *LcARFs*

To study the potential functions of *LcARFs* in *L. chinense*, we use the 2000 bp sequence upstream of ATG (start codon) to search cis-acting elements in the analysis of *LcARFs*. There were 11 types of cis-elements in these upstream regions, such as light response, temperature response, stress response, and others (Fig. 5). Almost all *LcARFs* had hormone response elements. *LcARF2a /7b/8a/8b/9/16b/17a/17b* had the most ABA-responsive elements. *LcARF4/6a/7a/7b/8a/8b/9/16b/17a* had most MeJA-responsive elements. *LcARF1* also had more drought response elements than other elements. Furthermore, *LcARF6a/6b/7a/7b/9* contained low-temperature response elements. It suggested that hormones or stress can influence the expression of *LcARFs*. All *LcARFs* had light response elements, revealing that these genes may play an important role in responding to light. Moreover, *LcARFs* also contained circadian elements related to plant growth and development, indicating that *LcARFs* might be involved in plant growth and development.

Interactions between LcARFs

In order to study the interaction between LcARFs to form homo/hetero-dimer, we predicted the protein-protein interaction network of LcARFs using *A. thaliana* as a reference (Fig. 6), and eight interacting links of LcARFs were found. It was conceivable that these interacting blocks of LcARFs in *L. chinense* may regulate different biological processes. Except for LcARF7b and LcARF19, the remaining 6 LcARFs interact with each other.

Fig. 5 Analysis of cis-acting elements of *LcARF* gene family. The 2000-bp regulatory region upstream of ATG was analyzed with the PlantCARE software

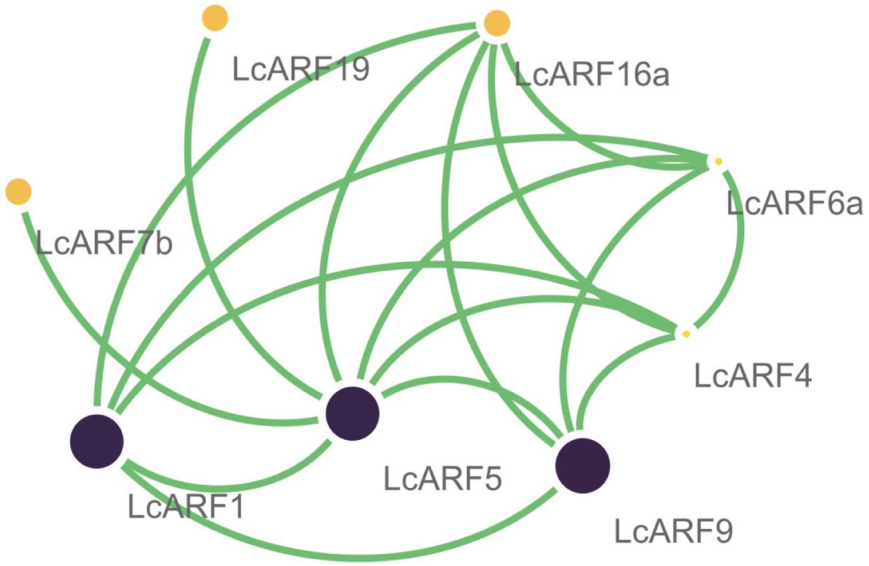

Fig. 6 Prediction of protein interaction. Protein-protein interaction network of ARFs in *L. chinense*, the results were based on an *Arabidopsis* association model

Subcellular localization of LcARFs

The location of the 20 LcARFs was all predicted to be nucleus localization. To verify the results, we select six *ARFs* were selected for the transient expression in protoplasts from the callus of *L. chinense*. We used a 35S promoter to drive coding sequences of *ARFs* fused with GFP (*35S::ARF::GFP*). GFP was only detected in the nucleus, suggesting that six ARFs were located in the nucleus (Fig. 7), consistent with the predicted results.

LcARFs tissue-specific expression

To understand how *LcARFs* might act during *L. chinense*' growth development, we profiled 7 transcriptomes of different organs/tissues (shoot, leaf, bud, stigma, stamen, sepal, petal) using RNA-seq to quantify the expression patterns of all 20 *LcARFs* (Fig. 8A). We found that all 20 *LcARFs* showed tissue-specific expression. Class A and B genes were highly expressed in stigma, sepal, and bud. All members of class C were expressed explicitly in the stamen. These results indicated that *LcARFs* were mainly expressed in young tissues, and their functions may be mainly related to flower development (such as stigma, stamen, or bud).

To validate transcriptome data from different tissues, we performed qRT-PCR experiments on three selected genes (Fig. 8B). Three genes had a high stigma and sepal expression levels with the same expression patterns. This result was consistent with RNA-seq (Fig. 8A). Overall, the expression of *LcARFs* was tissue-specific, and the specific expression in sexual organs and buds revealed that this gene family had essential functions in flower and bud development.

LcARFs universally expressed at different stages of somatic embryogenesis

To explore *LcARFs* expression during somatic embryogenesis, we analyzed in 6 successive stages (callus, globular embryo, heart-shaped embryo, torpedo embryo, early cotyledon embryo, cotyledon embryo) of *L. chinense* somatic embryogenesis by RNA-seq to quantify the expression patterns of all 20 *LcARF* genes (Fig. 9). All 20 *LcARFs* were expressed during somatic embryogenesis. Most members of class A were mainly expressed at the cotyledon stage. Members of class B were mainly expressed in torpedo, early cotyledon, and cotyledon stages. Members of class C were highly expressed in callus. Therefore, we speculated that members of class A and class B may be mainly involved in late somatic embryogenesis. The gene of class C may be involved in the maintenance of stem cells and early somatic embryogenesis.

Next, we selected four genes of class A member *LcARF6b* and class B member (*LcARF2b/3a/4*) for qRT-PCR verification (Fig. 9G). The results of qRT-PCR proved the reliability of transcriptome data, that is, class A members were mainly highly expressed in cotyledon embryo, and class B members were mainly highly expressed from torpedo embryo to cotyledon embryo. It also laid a theoretical foundation for improving the efficiency of somatic embryogenesis and propagating excellent *L. chinense* species.

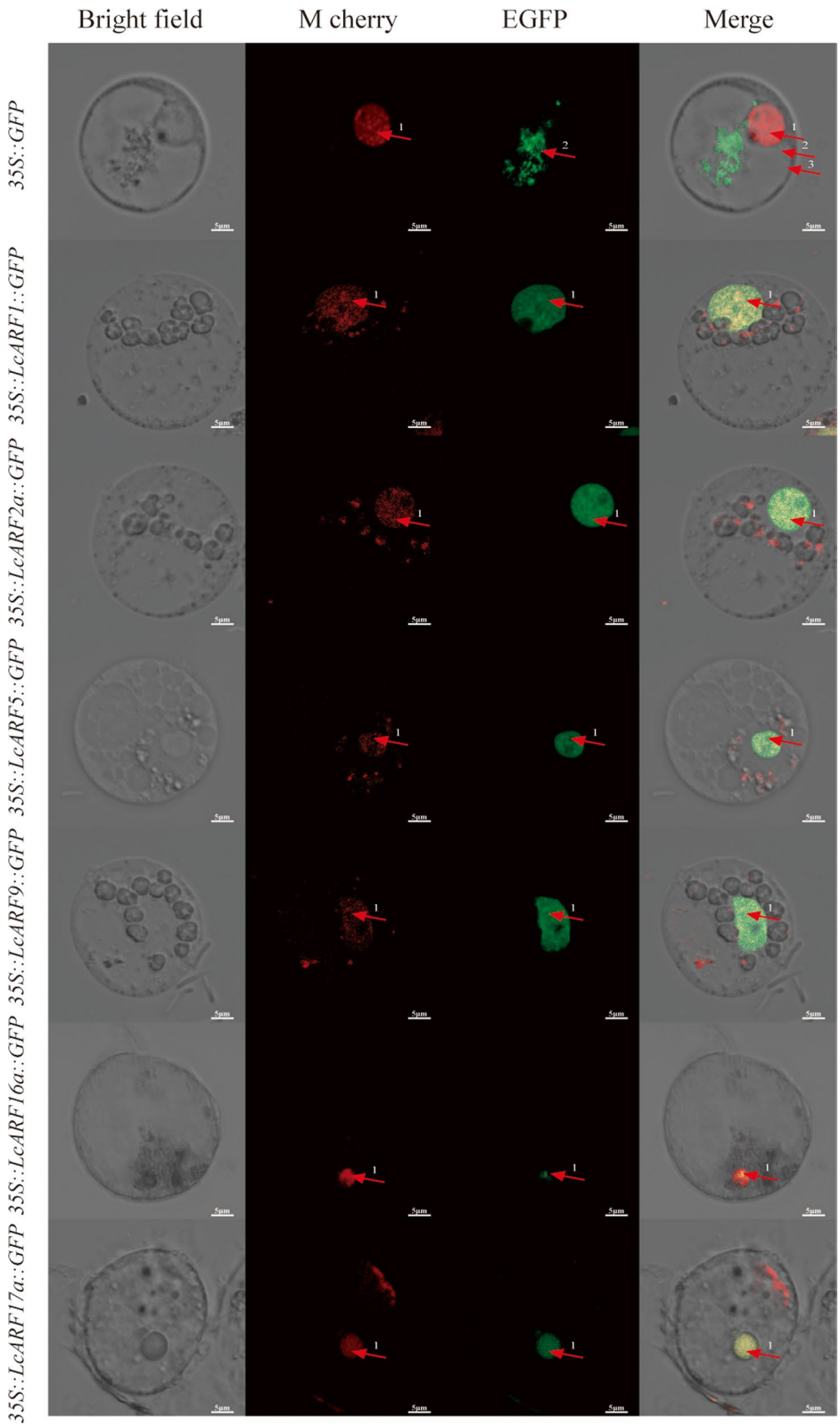

Fig. 7 Subcellular localization of LcARF in *L. chinense* protoplasts. The red fluorescence signal was the nuclear localization signal of H2B, and the green fluorescent signals of the six GFP fusion proteins were particularly strong in the nucleus

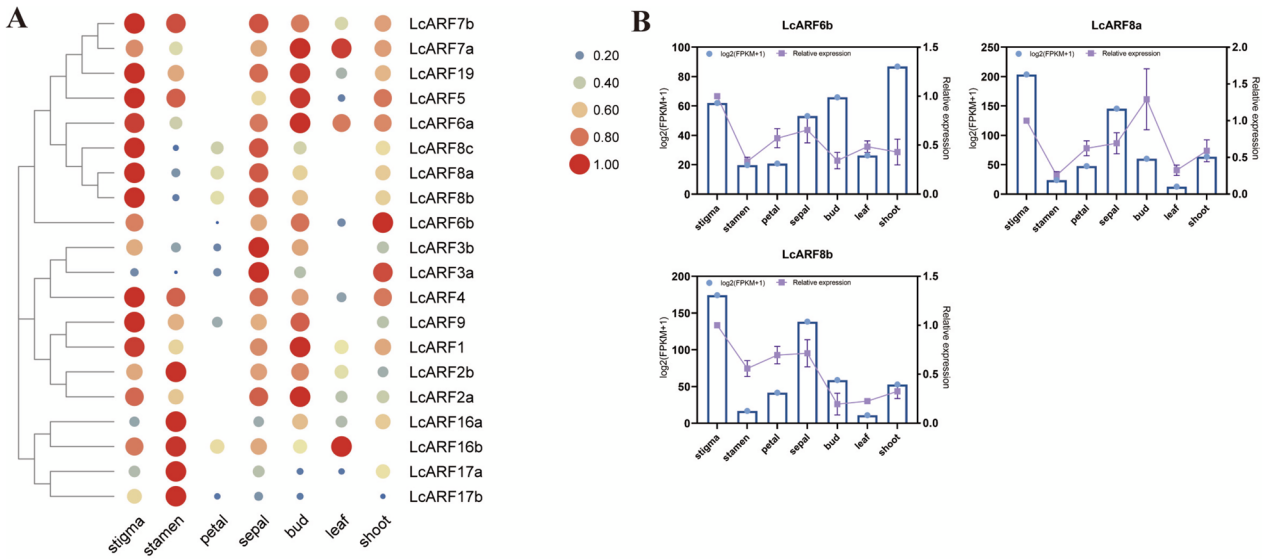

Fig. 8 Expression patterns of *LcARFs* in different tissues, analyzed by qRT-PCR. **A** The expression pattern of *LcARFs* in different tissues. **B** qRT-PCR was used to detect the expression pattern of *LcARFs* in different tissues. The purple line graph represents the results of qRT-PCR experiments, with the scale on the right ordinate of each graph. The blue histogram represents the results of FPKM analysis, with the scale on the left ordinate of each graph

Overexpression of *LcARFs* led to increased cell activity in callus

In our previous experiments focused on somatic embryogenesis, we discovered that the cell viability of callus directly impacted the efficiency of somatic embryogenesis. Based on transcriptome data analysis of *ARF*s in *L. chinense* at various stages of somatic embryo development, we selected *LcARF1/2a*, which had low expression levels in callus, and *LcARF5*, which had high expression levels, for further investigation. The 35S promoter was used to construct the overexpression vector, which was subsequently transferred into the callus of *Liriodendron* through genetic transformation. After conducting qRT-PCR experiments, we chose three overexpression lines for each genotype (Fig. S3). Our findings revealed that the cell area was notably larger in the transgenic lines overexpressing *LcARF2a* than in the wild type (Fig. 10A). Furthermore, additional measurements indicated that the cell length and width of the transgenic lines overexpressing *LcARF1/2a* were notably greater than those of the wild type (Fig. 10B). In addition, the cell aspect ratio of transgenic lines was substantially greater than that of the wild type, indicating that the overexpression of *LcARF1/2a/5* influenced the cell morphology of callus (Fig. 10B). Finally, staining the callus demonstrated that acetic acid magenta stained deeper in the overexpression lines. At the same time, Evans blue staining was lighter compared to the wild type (Fig. 10C). These results suggest that the overexpression of *LcARF1/2a/5* improves the cell viability of *L. chinense* callus.

LcARF1 may be involved in somatic embryogenesis

ARF1 is widely expressed in the thallus of *Marchantia polymorpha*, while *ARF2* is mainly expressed in the marginal area around the thallus of *Marchantia polymorpha* [25]. In order to more intuitively understand the expression pattern of *LcARF* during somatic embryogenesis in *Liriodendron*, we constructed an *LcARF1* promoter-mediated GFP fluorescence signal vector and obtained transgenic lines (Fig. 11). Comparing it to the control group, we found that GFP triggered by the *LcARF1* promoter was primarily expressed in the morphological lower end of the torpedo embryo and morphological lower end of the cotyledon embryo (Fig. 11E, F, K, L). It suggests that *LcARF1* may be involved in regulating hypocotylation and root development. In the globular embryo stage, *pLcARF1*-induced fluorescence signals were distributed throughout the embryo, aligning with *Arabidopsis* report [19] (Fig. 11D, J).

Discussion

Molecular characterization and evolution of the *ARF* gene family in *L. chinense*

Auxin response factors are an essential group of signaling factors in the auxin signaling pathway [23, 26, 27]. They play a crucial role in regulating various physiological and developmental processes in plants by controlling the expression of downstream genes. *ARFs* are primarily involved in regulating plant hormone responses and biological responses to external stimuli [28]. *ARFs* were studied in many plants, including *Arabidopsis* [29],

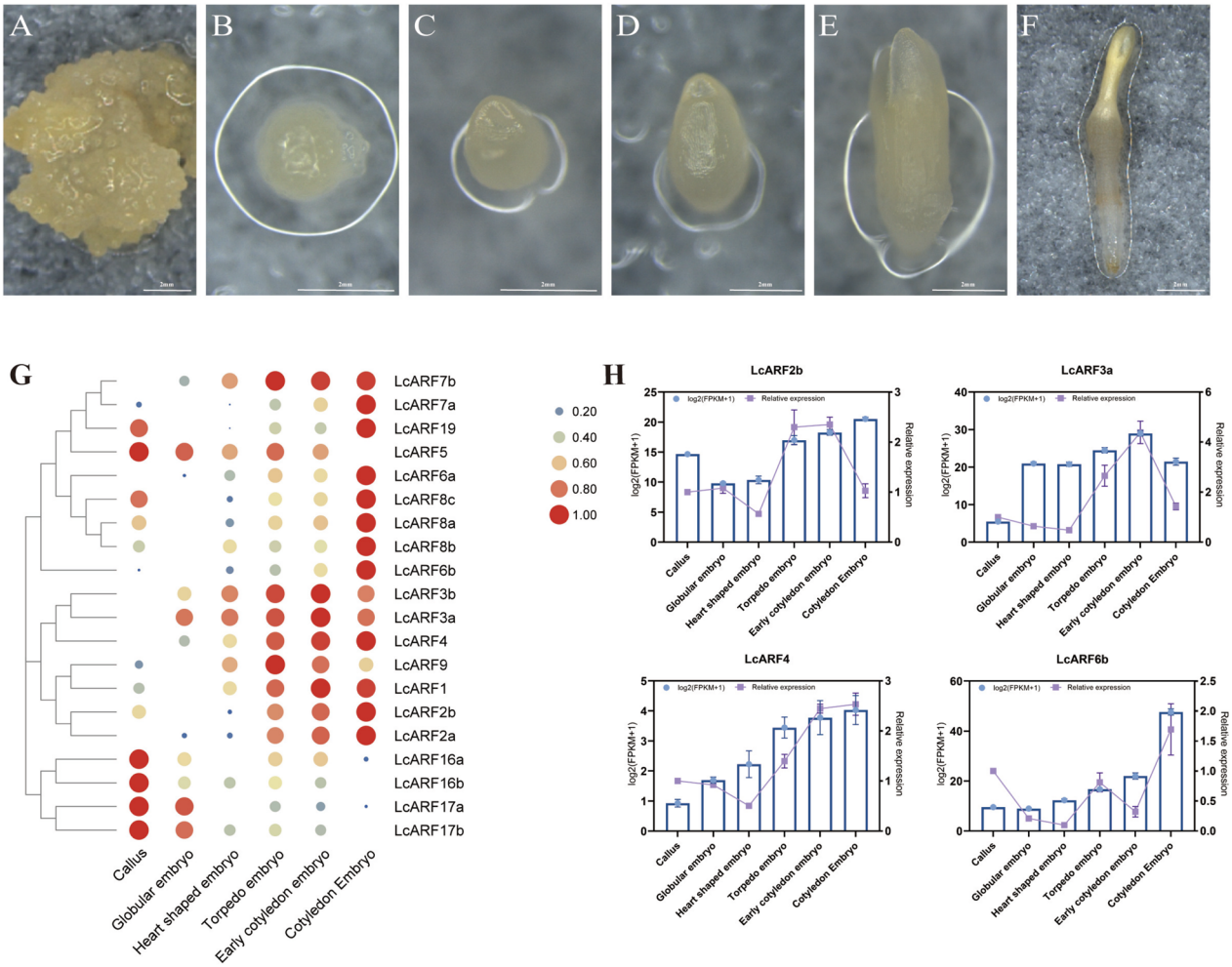

Fig. 9 Expression analysis of *LcARFs* under SE, analyzed by qRT -PCR. **A** Callus. **B** Globular embryo. **C** Heart-shaped embryo. **D** Torpedo embryo. **E** Early cotyledonary embryo. **F** Cotyledon embryo. **G** The expression pattern of *LcARFs* in somatic embryos. **H** qRT -PCR was used to detect the expression pattern of *LcARFs* in different stages of SE. The purple line graph represents the results of qRT-PCR experiments, with the scale on the right ordinate of each graph. The blue histogram represents the results of FPKM analysis, with the scale on the left ordinate of each graph

maize [30], rice [31], tomato [32], *Populus trichocarpa* [33], *Eucalyptus grandis* [34], apple [35], orange [36], physic nut [37], longan [38] and others. In this study, 20 *LcARFs* were identified in *L. chinense*. Given that the genome size of *L. chinense* was 1.75 Gb, nearly 12 times larger than the *Arabidopsis* genome [4], nearly 4 times larger than the rice genome [39], it was surprising that the number of *LcARFs* was less than that of *Arabidopsis* and rice. Apparently, our observations on the *ARF* gene family contradicted with genome complexity between *Arabidopsis*, rice, and *L. chinense*. Class B-ARFs contained 7 members in *L. chinense*, 9 members of rice, and 14 members in *Arabidopsis*. The number of *ARF* genes in *Arabidopsis* distributed in class-B is 2 times that of *L. chinense* and 1.56 times that of rice. Some researchers believed that several independent, small-scale, segmental replication events and chromosomal rearrangements had occurred at the *ARF13* locus, resulting in multiple members of the *AtARF13* family and ultimately the *AtARF* gene family in *Arabidopsis* Expansion [31]. In the evolution of angiosperms, whole-genome or segmental duplication and tandem duplication often occur, leading to the expansion of gene families [38]. In the study of rice and longan, researchers found that *ARFs* evolved through whole genome or segment duplication events, and there was no tandem duplication phenomenon [31, 38]. Our study confirmed that 20 *LcARFs* did not exhibit tandem duplication. It suggested that the number of members of the *LcARF* gene family may be associated with the loss of duplicated genes and the loss of tandem duplication events. Previous studies have shown that most of the *ARFs* of species such as *Arabidopsis* and rice have about six exons [31]. Our study found that the number of exons of *LcARFs* was much higher than that of rice and

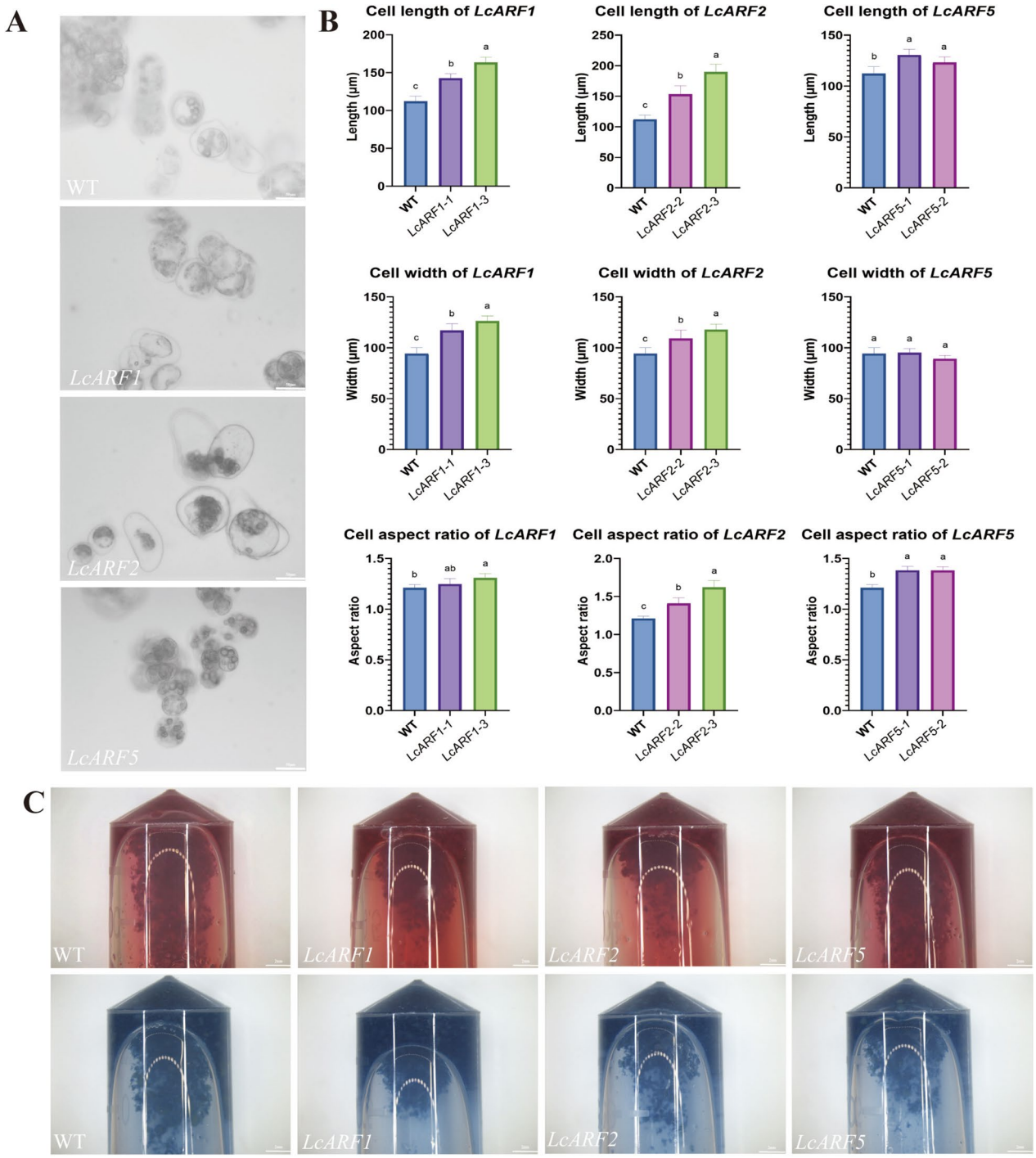

Fig. 10 Phenotypic of calli overexpressing *LcARFs* in *Liriodendron*. **A** Cell morphology of calli from WT and overexpression lines. **B** Cell length, cell width, and cell aspect ratio statistics of calli from WT and overexpression lines. **C** Acetomagenta and Evans blue staining of calli from WT and overexpressed lines. $p < 0.05$

Arabidopsis, and most contained more than ten exons. It indicated that *LcARFs* might have more abundant functions.

Nuclear auxin signaling mediated by *ARF* transcription factors affects plant growth and development by regulating cell division, elongation, and differentiation. The evolutionary origin of the *ARF*-mediated pathway dates back

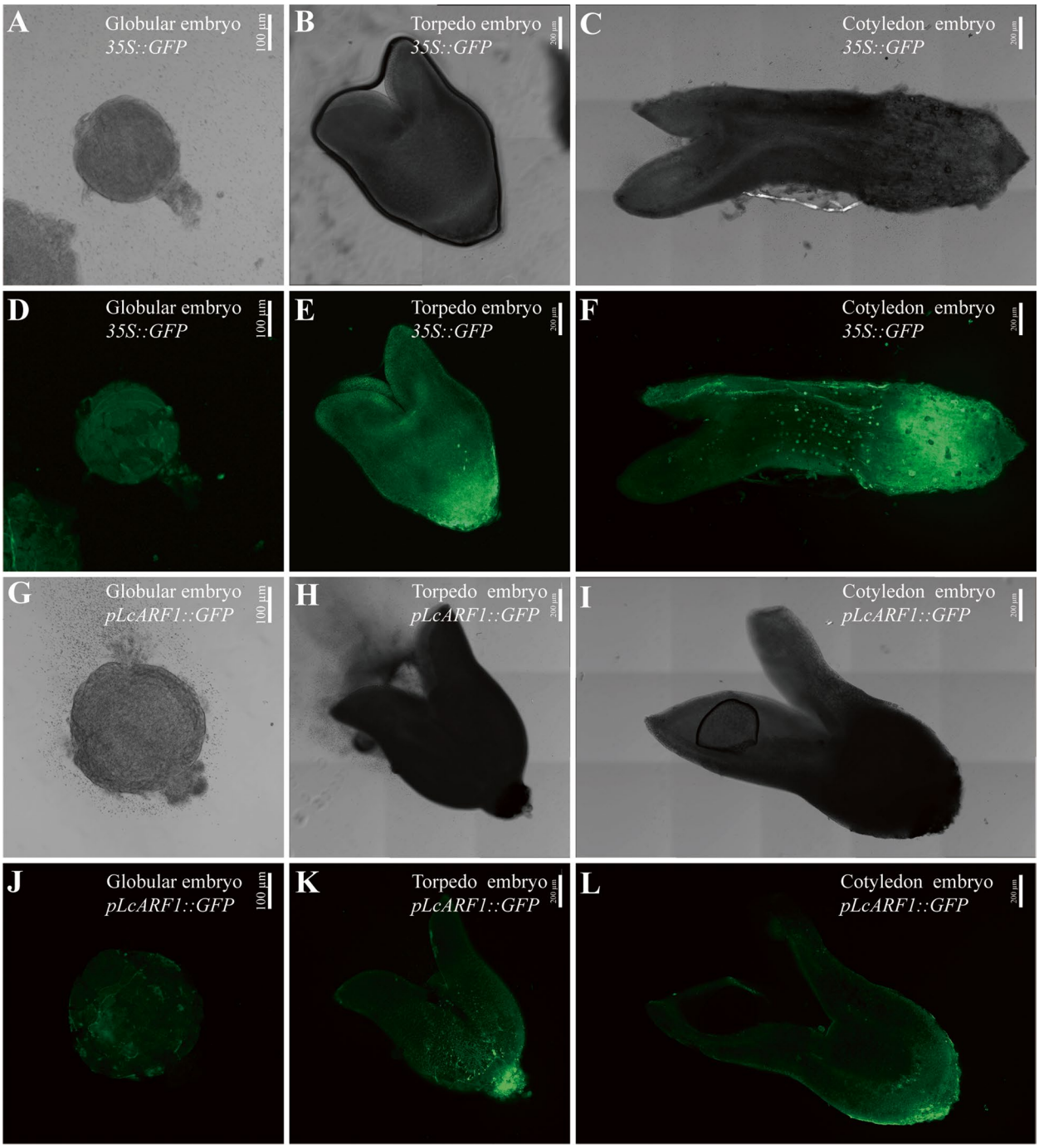

Fig. 11 Expression pattern of *LcARF1* during somatic embryogenesis. **A**, **D**, **G**, **J**: Globular embryo. **B**, **E**, **H**, **K**: Torpedo embryo. **C**, **F**, **I**, **L**: Cotyledon embryos. **A**- **F**: 35S::GFP. **G**- **L**: *pLcARF1::GFP*

to at least the common ancestor of bryophytes and other land plants [40, 41]. Because *L. chinense* is an ancient relic plant, in order to study the evolutionary trajectory of *LcARFs*, we conducted a comprehensive phylogenetic study using LcARF protein sequences from algae to terrestrial plants [42, 43]. Consistent with previous reports, LcARFs were composed of three main classes: class A (LcARF5/6/7/8), class B (LcARF1/2/3/4/9), and class C (LcARF10/16/17). In the three main LcARF*s* classes, we discovered all the bryophyte LcARF sequences, indicating that LcARFs may differentiate into three sets of transcription factors in the common ancestor of bryophytes,

and this primordial event may be prior to the differentiation of liverwort and moss [42]. Surprisingly, in a phylogenetic analysis, we detected that these LcARF sequences formed three independent clusters based on each daughter class. Therefore, we speculate that the LcARF of these bryophytes may be identical to the counterpart of the ancestral *ARF* gene in land plants. The protein domain analysis found that all the 20 *LcARFs* contained Auxin_resp domain, consistent with previous studies, so *LcARFs* and other plants *ARFs* may be evolutionally conserved and functionally similar [44, 45].

LcARFs may play significant roles in the growth and development of *L. chinense*

In recent years, the biological functions of *ARFs* in plant growth and development have been further studied. *ARF* transcription factors mediate the activity of the plant hormone auxin, regulating various aspects of plant development [46]. The researchers found that the *ARF* gene family was involved in the early development of cotton fibers and regulates early senescence in lilies and morning glories [47–49]. Specifically, *ARF* genes under class A were mainly related to shooting regeneration and adventitious root development. For example, in *Arabidopsis*, *AtARF5/7/19* played an important role in leaf vein development, *AtARF5* was involved in shoot regeneration, and *AtARF7/19* was involved in adventitious root development [50–52]. Class B branch genes were mostly related to the development of leaves, stems and roots. Studies on carrots, longan, corn and other species found that *ARF1/2* are involved in the growth and development of vegetative organs [38, 53, 54]. The *ARF* gene of class C is mainly related to gametophyte development. *ARF10/16/17* of grape and *ARF17* of tomato regulate parthenocarpy under the action of miR160 [55–57]. In addition, *AtARF2/4/5* was also involved in *Arabidopsis* gametophyte development [58]. In the reported literature, *ARF4* was a multifunctional gene, which was involved in gametophyte development and shoots regeneration in *Arabidopsis*, regulated the development of wheat roots and stems, and regulated the stomatal switch of tomato to enhance the response to salt damage and waterlogging [50, 59, 60]. In this study, we investigated the differential expression of *LcARFs* in tissues through transcriptome and qRT-PCR experiments. In accordance with findings from other species, classes A and B demonstrated high expression levels in vegetative organs. However, nearly all *ARFs* were found to be expressed in reproductive organs, with class A being predominantly expressed in stigma, class B exhibiting similar expression patterns in both stigma and stamens, and class C showing high expression levels mainly in stamens. This phenomenon may be attributed to the regulation of auxin during the development of male and female gametophytes. Overall, our findings suggest that *LcARF* may play a crucial role in the growth and development of *Liriodendron*, and the study of these key *LcARFs* may aid in the exploration and regulation of related functions at the molecular level. The results of these studies could further enhance the growth and reproduction rate of *Liriodendron*, potentially reducing the cost of wood utilization and promoting the use of afforestation applications.

The role of *LcARFs* in somatic embryogenesis and their regulatory mechanisms

Somatic embryogenesis is induced by transcriptional reprogramming. Somatic cells respond to the induction signal and enter the embryonic development pathway after treatment with plant hormones, mainly auxin, to form somatic embryos [61]. Auxin triggers various molecular regulatory mechanisms during development, including *ARF*, the core component of the auxin signaling pathway [26, 62]. Signaling from *AtARF5* is necessary to form *Arabidopsis* callus shoots [63]. Mutations in *AtARF5* will result in severe patterning defects during embryonic and postembryonic development [64].

Somatic embryogenesis of *L. chinense* is a crucial method to obtain more seedlings. However, *LcARFs* regulating somatic embryogenesis have not been reported in *L. chinense*. Therefore, we explored the expression pattern of *LcARFs* during somatic embryogenesis. Expression trends of *LcARFs* in six stages of embryogenesis were divided into three groups. The first group was mainly class B members, with higher expression levels in torpedo embryos, early cotyledon embryos, and cotyledon embryos. The second group was mainly composed of class C members, which were highly expressed in callus and gradually decreased with the progress of somatic embryogenesis. The remaining *LcARFs*, the last group, were expressed at low levels during early somatic embryogenesis and high levels during the cotyledon embryonic stage. Interestingly, all *LcARFs* expression levels in globular and heart-shaped embryos were very low.As the vitality of callus, which is directly linked to the efficiency of somatic embryogenesis, is of utmost importance, it is critical to examine the effect of *ARFs* expression on callus condition. Based on the findings of transcriptome analysis, we selected *LcARF1/2a/5* for respective overexpression. The results indicated that the callus activity of the overexpression lines was greater than that of the wild type, suggesting that *ARFs* expression may influence or participate in the somatic embryogenesis process of *Liriodendron*. Utilizing the GFP fluorescence reporting system, we determined that *LcARF1* expression was present in the globular embryo, torpedo embryo and cotyledon embryo of *Liriodendron*. Moreover, we observed

that *LcARF1* was mainly concentrated in the lower end of the embryo morphology. These findings suggest that *ARF1* actively participates in the somatic embryogenesis of *Liriodendron* and may potentially regulate root development.

In summary, according to previous research reports, *ARFs* have some essential functions in the process of somatic embryogenesis [65]. These results provide new clues for studying *ARF* genes involved in somatic embryogenesis. Therefore, *LcARFs* are worthy of further revealing their regulatory mechanism through molecular biology experiments. In the future, these important somatic embryogenesis regulatory genes will help improve the efficiency of the asexual reproduction of *L. chinense*, thereby increasing the reproductive ability of *L. chinense* and increasing the yield of trees.

Conclusions

This study comprehensively identified the *Liriodendron chinense ARF* gene family. 20 *LcARFs* gene structures, conserved motifs, phylogeny, cis-acting elements, and protein interaction predictions were analyzed. Then, expression levels of *LcARFs* in different tissues and somatic embryogenesis were analyzed by RNA-seq and qRT-PCR simultaneously. We confirmed that the expression of *LcARFs* was tissue-specific and participated in the somatic embryogenesis process in *Liriodendron*. Furthermore, our study demonstrated that overexpressing *LcARF1/2a/5* enhances the activity of the callus, and we also established the involvement of *LcARF1* in the somatic embryogenesis process. The finding that *ARFs* expression plays a crucial role in the somatic embryogenesis of *Liriodendron* is an important step forward in understanding the breeding process of this species. These results have laid a strong foundation for the optimization of somatic embryogenesis conditions and the improvement of somatic embryo yields, ultimately meeting the needs of horticultural greening and industrial production.

Materials and methods

Datasets and sequence retrieval

The complete genome, transcript/protein sequences, and genome feature file of *L. chinense* were downloaded from https://db.cngb.org/search/project/CNP0000815/ [4]. All ARF proteins of *Arabidopsis* thaliana were obtained from the Plant Transcription Factor Database (http://planttfdb.gao-lab.org/). An ARF Hidden Markov profile (B3 DNA binding (Pfam 02362), Auxin_Resp (Pfam 06507) and AUX_IAA (Pfam 02309)) was retrieved from the Pfam website (http://pfam.xfam.org), and the protein sequences in the *L. chinense* genome were identified using HMMER (v.3.0.1b) and BLASTP. All identified ARFs were further validated by a conserved domain search using the conservative Domains Database (CDD; https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and PFAM (http://pfam.xfam.org/) databases, whose E value cut of < 1E-5, consequently, the redundant and partial sequences were removed manually [66].

Sequence analysis

For the gene structure illustration, we utilized the GFF3 file of the *L. chinense* genome and images were implemented by TBtools software [67]. The motifs analysis of ARF protein was performed by the Multiple Em of Motif Elicitation (MEME Suite) and then demonstrated by TBtools software. For each gene of *ARF*, the several physicochemical properties (i.e., molecular weight (MW), isoelectronic points (PIs), and others) were intended by ExPASY PROTPARAM tools (http://web.expasy.org/protparam/). Subcellular localization was analyzed by PlantmPLoc website (http://www.csbio.sjtu.edu.cn/bioinf/plant- multi/) [68, 69]. The *ARF* promoter sequences (i.e., selected as 2000 bp) were initially imported in Generic File Format (GFF) from the *L. chinense* genome. After that various cis-regulatory elements for each promoter sequence were identified by the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The *L. chinense* genomic database was utilized for the chromosomal locations of *ARF* genes and was mapped based on available information. Protein-protein interaction (PPI) analyses of the LcARF family were performed on the STRING website (http://www.string-db.org) to predict protein interactions. Moreover, we use Cytoscape (3.8.2) software to draw predicted protein network interaction maps.

Phylogenetic analysis

Multiple sequence alignment (MSA) of ARF full-length proteins with FASTA format was done using MUSCLE (v3.8.31) set the following parameters: the maximum number of iterations was 1 with '-maxiters 1' , and find diagonals with '-diags -sv -distance1 kbit20_3'. The trimmed MSA file was generated with trimAl (v1.4) set to 'automated1' mode and then used to construct the ARF phylogenetic tree. The Bayesian phylogenetic tree was constructed using the BEAST software (v2.6.6), by inputting the trimmed file in FASTA format with BEAUti 2 program [70].

Somatic embryogenesis successive stages and various tissue transcriptomes of *L. chinense*

The immature embryos of *L. chinense* were used to induce embryogenic callus. After expanded culture on M13 (3/4MS + 30 g/L sucrose + 2 mg/L 2.4-D + 0.2 mg/L 6-BA) liquid medium, single cells were obtained

by screening with a 400-mesh sieve and placed on 3/4MS (40 g/L sucrose + 0.8 g/L agar) medium to induce somatic embryogenesis. Embryogenic callus and five successive stages of somatic embryogenesis, i.e., globular embryo, heart-shaped embryo, torpedo embryo, early cotyledon embryo and cotyledon embryo, were collected and used for transcriptome sequencing. We used the transcriptomes of various tissues, i.e., shoot, leaf, bud, stigma, stamen, sepal, and petal. Among them, the induction of somatic embryogenesis was repeated three test cycles and callus and embryo samples were collected. Tissue samples were taken from 30-year-old *L. chinense* planted in Nanjing Forestry University. RNA-seq was used to sequence the RNA samples of tissues and somatic embryos, and then our genome sequence was used as a reference to draw clean readings. Trimmomatic (v0.36) was used to remove adaptors, poly- (A) tails, and low-quality reads from the original Illumina sequencing data. These clean reads were then matched to the SMRT long read reference sequence, and the number of matched reads for each reference sequence was calculated using RSEM (v1.3.0) to quantify the transcripts. The FPKM values of all mapped *LcARFs* are shown in Table S4 to calculate the transcript abundance of *LcARF* genes. The transcriptome data of this study have not been published.

RNA extraction and qRT-PCR analysis

We collected samples of shoot, leaf, bud, stigma, stamen, sepal, petal and six embryonic development stages (callus, globular embryo, heart-shaped embryo, torpedo embryo, early cotyledon embryo, cotyledon embryo) of *L. chinense*. Tissue samples were taken from the adult trees of *L. chinense* in Nanjing Forestry University (Nanjing, China). Total RNA extraction was performed using a FastPure Total RNA Isolation Kit from Vazyme (Nanjing, China) (RC401) corporation. In this study, we set three biological replicates for each group of samples, and three technical replicates for each biological replicate [70]. All primers for qRT-PCR were designed by Primer3.0 and are listed in Table S5. All experiments were run on 96-well plates. All data generated from qRT-PCR were calculated by the $2^{-\triangle\triangle CT}$ formula. ACT97 was selected as the internal reference gene, one-way analysis of variance was performed using IBM SPSS Statistics 26, and GraphPad Prism 9 was used to draw the histogram.

Subcellular localization

Using the predicted CDS sequences, we designed primers to clone the CDS sequences of six *LcARF* genes (*LcARF1/2a/5/9/16a/17a*) utilizing *L. chinense* cDNA as a template. Following this, we proceeded to clone the full-length coding sequences of *ARFs*, excluding stop codons, into the green fluorescent protein (GFP) vector (pCAMBIA1302) for subcellular localization. The vector construction and primer information are detailed in the Table S5. The recombinant plasmid and the vector with nuclear localization signal (35S::H2B::Mcherry) were transiently transformed into protoplasts of *Liriodendron* callus by PEG-mediated transformation [71]. After overnight incubation at 23 °C in the dark, a Zeiss LSM 480 fluorescence confocal microscopy for detecting GFP and M cherry signal.

Obtaining genetically modified *Liriodendron*

To obtain transgenic *Liriodendron* callus, we first selected *Liriodendron* callus with a good growth state (fine particles, uniform size, light yellow color) as the experimental material. Then, we used the plant binary expression vectors (*35S::LcARF1::GFP*, *35S::LcARF2a::GFP*, *35S::LcARF5::GFP*) constructed in 4.6 and introduced the vector into callus cells by Agrobacterium-mediated transformation [8]. Through this process, transgenic lines were obtained. We used ACT97 as an internal reference gene, and used qRT-PCR experiments to screen transgenic lines with overexpression of the target gene. The specific primers used in the experiment are listed in Table S5.

Detection of callus activity of transgenic lines

The callus from the transgenic lines was carefully rinsed with phosphate buffer (0.1M, pH = 8.0) and placed onto a slide. The cell morphology was observed under an inverted microscope. Using the ImageJ software, the cells' length and width were measured, and the cell aspect ratio was calculated. At least 100 cells were measured randomly for each line. Acetic acid magenta and Evans blue staining were used to assess callus cell activity further. Acetic acid magenta is capable of staining chromosomes or chromatin into a shade of purplish-red, facilitating the observation of cell mitosis. In an acidic environment, the carboxyl group of magenta acetate can become charged, allowing it to polarly bind to structural components such as chromosomes and nuclei, thereby staining them. Typically, darker staining indicates higher cell viability. On the other hand, Evans blue is capable of binding to proteins to form Evans blue protein complexes, which cannot penetrate normal cell membranes. However, when cells are damaged, Evans blue can penetrate the membrane and bind to proteins, turning them blue. Thus, deeper staining indicates more damaged cells. The calli from the transgenic lines were treated with stains and washed thrice with PBS after 10 min. The stained calli were then observed using a stereomicroscope.

Induction and observation of somatic embryogenesis

Suspension cultured calluses were screened using 150-mesh and 400-mesh sieves. The single cells on the 400 mesh sieve were rinsed with 3/4MS liquid medium and placed on 3/4MS medium supplemented with agar (0.8 g/L) to induce somatic embryogenesis [5]. The globular embryos, torpedo embryos and cotyledon embryos of the experimental group (*pLcARF1::GFP*) and the control group (*35S::GFP*) were selected to observe the distribution of GFP fluorescence signals in embryos.

Abbreviations

qRT-PCR	Quantitative real-time reverse transcription PCR
IAA	Indole Acetic Acid
ABA	Abscisic Acid
MeJA	Methyl Jasmonate

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-04765-7.

Additional file 1: Table S1. *LcARFs* position of Chromosome.

Additional file 2: Table S2. Evolutionary tree of LcARF.

Additional file 3: Table S3. List of domains, MR, CDS, cis-regulatory elements information of the identified LcARF.

Additional file 4: Table S4. FPKM values for different tissue and somatic embryogenesis processes in *LcARF*.

Additional file 5: Table S5. List of primers.

Additional file 6: Figure S1. Genomic distribution of *LcARF* genes on chromosomes.

Additional file 7: Figure S2. Classification of *Liriodendron chinense* ARF proteins.

Additional file 8: Figure S3. qRT-PCR was used to detect the overexpression of target genes in the transgenic callus.

Acknowledgments

We would like to express our sincere gratitude to Dr. Dechen Xu and Dr. Xianlong Dai for their unwavering support and invaluable assistance throughout our research.

Authors' contributions

LX performed the genome-wide analysis of *LcARFs*, qPCR experiments, and subcellular localization, bred genetically modified, and wrote this manuscript. YL, WHW, and ZDH performed the assistance of analysis. SCH and YRL helped the promoter design and qPCR experiments, and JJZ provided the transcriptome data of somatic embryogenesis. JHC and JSS were responsible for the overall concept and experimental designs, data integration, analysis and interpretation, and manuscript preparation. All authors approved the final manuscript.

Funding

This research was supported by the National Key Research and Development Program of China during the 14th Five-year Plan Period (2021YFD2200102), the Youth Foundation of the Natural Science Foundation of Jiangsu Province (Grant No. BK20210614), the Natural Science Foundation of China (32071784) and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

All data analyzed during this study are included in this article and its additional files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 28 August 2023 Accepted: 24 January 2024
Published online: 07 February 2024

References

1. Sheng Y, Hao Z, Peng Y, Liu S, Hu L, Shen Y, Shi J, Chen J. Morphological, phenological, and transcriptional analyses provide insight into the diverse flowering traits of a mutant of the relic woody plant *Liriodendron chinense*. Hortic Res-Engl. 2021; 8(1):174.
2. Li T, Yuan W, Qiu S, Shi J. Selection of reference genes for gene expression analysis in *Liriodendron* hybrids' somatic embryogenesis and germinative tissues. Sci Rep-UK. 2021;11(1):4957.
3. Zhong YD, Sun XY, Liu EY, Li YQ, Gao Z, Yu FX. Expressed sequence tag analysis of functional genes associated with adventitious rooting in *Liriodendron* hybrids. Genet Mol Res. 2016;15(2):7606.
4. Chen J, Hao Z, Guang X, Zhao C, Wang P, Xue L, Zhu Q, Yang L, Sheng Y, Zhou Y, et al. *Liriodendron* genome sheds light on angiosperm phylogeny and species–pair differentiation. Nat Plants. 2019;5(1):18–25.
5. Tingting L, Jinhui C, Jisen S, Jin X: Deep sequencing combined with microarray hybridization to identify novel and conserved microRNAs during somatic embryogenesis of hybrid yellow-poplar (*Liriodendron chinense* (Hemsl.) Sarg. X *L. tulipifera* Linn.). BMC Proc. 2011, 5(Suppl 7):P74.
6. Ali A, Zhang J, Zhou M, Chen T, Shah L, Rehman SU, Hayat S, Shi J, Chen J. Chitosan oligosaccharides stimulate the efficacy of somatic embryogenesis in different genotypes of the *Liriodendron* hybrid. Forests. 2021;12(5):557.
7. Chen T, Yang D, Fan R, Zheng R, Lu Y, Cheng T, Shi J, Chen J. γ-Aminobutyric acid a novel candidate for rapid induction in somatic embryogenesis of *Liriodendron* hybrid. Plant Growth Regul. 2022;96(2):293–302.
8. Wang D, Lu F, Lu Y, Cheng T, Shi J, Chen J, Hao Z. Identification of miR397a and its functional characterization in callus growth and development by regulating its target in *Liriodendron*. Forests. 2021;12(7):912.
9. Wang D, Lu F, Wang P, Zhang J, Lu L, Zhen Y, Lu Y, Cheng T, Shi J, Chen J. The regulatory roles of microRNAs and associated target genes during early somatic embryogenesis in *Liriodendron Sino-Americanum*. Phyton (Buenos Aires). 2021;90(5):1445–63.
10. Zheng Q, Zheng Y, Ji H, Burnie W, Perry SE. Gene regulation by the AGL15 transcription factor reveals hormone interactions in somatic embryogenesis. Plant Physiol. 2016;172(4):2374–87.
11. Boer DR, Freire-Rios A, van den Berg WAM, Saaki T, Manfield IW, Kepinski S, López-Vidrieo I, Franco-Zorrilla JM, de Vries SC, Solano R, et al. Structural basis for DNA binding specificity by the Auxin-dependent ARF transcription factors. Cell. 2014;156(3):577–89.
12. Shen C, Wang S, Bai Y, Wu Y, Zhang S, Chen M, Guilfoyle TJ, Wu P, Qi Y. Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa L.*). J Exp Bot. 2010;61(14):3971–81.
13. Tiwari SB, Hagen G, Guilfoyle T. The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell. 2003;15(2):533–43.
14. Luo J, Zhou J, Zhang J. *Aux/IAA* gene family in plants: molecular structure, regulation, and function. Int J Mol Sci. 2018;19(1):259.

15. Roosjen M, Paque S, Weijers D. *Auxin response factors*: output control in auxin biology. J Exp Bot. 2018;69(2):179–88.
16. Guseman JM, Hellmuth A, Lanctot A, Feldman TP, Moss BL, Klavins E, Calderon Villalobos LIA, Nemhauser JL. Auxin-induced degradation dynamics set the pace for lateral root development. Development. 2015;142(5):905–9.
17. Reeves PH, Ellis CM, Ploense SE, Wu M, Yadav V, Tholl D, Chételat A, Haupt I, Kennerley BJ, Hodgens C, et al. A regulatory network for coordinated flower maturation. PLoS Genet. 2012;8(2):e1002506.
18. De Smet I, Lau S, Ehrismann JS, Axiotis I, Kolb M, Kientz M, Weijers D, Jürgens G. Transcriptional repression of *BODENLOS* by HD-ZIP transcription factor HB5 in *Arabidopsis thaliana*. J Exp Bot. 2013;64(10):3009–19.
19. Rademacher EH, Möller B, Lokerse AS, Llavata-Peris CI, van den Berg W, Weijers D. A cellular expression map of the Arabidopsis *AUXIN RESPONSE FACTOR* gene family. Plant J. 2011;68(4):597–606.
20. Rademacher EH, Lokerse AS, Schlereth A, Llavata-Peris CI, Bayer M, Kientz M, Freire Rios A, Borst JW, Lukowitz W, Jürgens G, et al. Different auxin response machineries control distinct cell fates in the early plant embryo. Dev Cell. 2012;22(1):211–22.
21. Muto H, Nagao I, Demura T, Fukuda H, Kinjo M, Yamamoto KT. Fluorescence cross-correlation analyses of the molecular Interaction between an Aux/IAA Protein, MSG2/IAA19, and protein-protein interaction domains of auxin response factors of *Arabidopsis* expressed in HeLa cells. Plant Cell Physiol. 2006;47(8):1095–101.
22. Hardtke CS, Ckurshumova W, Vidaurre DP, Singh SA, Stamatiou G, Tiwari SB, Hagen G, Guilfoyle TJ, Berleth T. Overlapping and non-redundant functions of the *Arabidopsis* auxin response factors *MONOPTEROS* and *NONPHOTOTROPIC HYPOCOTYL 4*. Development. 2004;131(5):1089–100.
23. Wójcikowska B, Gaj MD. Expression profiling of *AUXIN RESPONSE FACTOR* genes during somatic embryogenesis induction in Arabidopsis. Plant Cell Rep. 2017;36(6):843–58.
24. Indoliya Y, Tiwari P, Chauhan AS, Goel R, Shri M, Bag SK, Chakrabarty D. Decoding regulatory landscape of somatic embryogenesis reveals differential regulatory networks between *japonica* and *indica* rice subspecies. Sci Rep-UK. 2016;6(1):23050.
25. Kato H, Mutte SK, Suzuki H, Crespo I, Das S, Radoeva T, Fontana M, Yoshitake Y, Hainiwa E, van den Berg W, et al. Design principles of a minimal auxin response system. Nat Plants. 2020;6(5):473–82.
26. Weijers D, Wagner D. Transcriptional responses to the auxin hormone. Annu Rev Plant Biol. 2016;67(1):539–74.
27. Ma Q, Grones P, Robert S. Auxin signaling: a big question to be addressed by small molecules. J Exp Bot. 2018;69(2):313–28.
28. Chandler JW. Auxin response factors. Plant Cell Environ. 2016;39(5):1014–28.
29. Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui A, Nguyen D, et al. Functional genomic analysis of the *AUXIN RESPONSE FACTOR* Gene family members in *Arabidopsis thaliana*: unique and overlapping functions of *ARF7* and *ARF19*. Plant Cell. 2005;17(2):444–63.
30. Liu Y, Jiang H, Chen W, Qian Y, Ma Q, Cheng B, Zhu S. Genome-wide analysis of the auxin response factor (ARF) gene family in maize (*Zea mays*). Plant Growth Regul. 2011;63(3):225–34.
31. Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Han B, Tao Y. Genome-wide analysis of the *auxin response factors* (ARF) gene family in rice (*Oryza sativa*). Gene. 2007;394(1–2):13–24.
32. Kumar R, Tyagi AK, Sharma AK. *Genome-wide analysis of* auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Mol Genet Genomics. 2011;285(3):245–60.
33. Kalluri UC, Difazio SP, Brunner AM, Tuskan GA. Genome-wide analysis of *Aux/IAA* and *ARF* gene families in *Populus trichocarpa*. BMC Plant Biol. 2007;7:59.
34. Yu H, Soler M, Mila I, San CH, Savelli B, Dunand C, Paiva JA, Myburg AA, Bouzayen M, Grima-Pettenati J, et al. Genome-wide characterization and expression profiling of the *AUXIN RESPONSE FACTOR* (*ARF*) gene family in *Eucalyptus grandis*. PLoS One. 2014;9(9):e108906.
35. Luo X, Sun M, Xu R, Shu H, Wang J. ZHANG S: *Genomewide identification and expression analysis of the* ARF gene family in apple. J Genet. 2014;93(3):785–97.
36. Li S, OuYang W, Hou X, Xie L, Hu C, Zhang J. Genome-wide identification, isolation and expression analysis of auxin response factor (*ARF*) gene family in sweet orange (*Citrus sinensis*). Front Plant Sci. 2015;6:119.
37. Tang Y, Bao X, Liu K, Wang J, Zhang J, Feng Y, Wang Y, Lin L, Feng J, Li C. *Genome-wide identification and expression profiling of the* auxin response factor (ARF) gene family in physic nut. PLoS One. 2018;13(8):e201024.
38. Shen X, Chen R, Chen X, Munir N, Zhang S, Xu X, Lin Z, Zhang J, Li X, Lin Y, et al. Molecular evolution and expression analysis of ADP-ribosylation factors (*ARFs*) from longan embryogenic callus. Gene. 2021;777:145461.
39. Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X et al: A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). science. 2002, 296(5565):79-92.
40. Kohchi T. The Roles of the sole Activator-type Auxin response factor in pattern formation of *Marchantia polymorpha*. Plant Cell Physiol. 2017;58(10):1642–51.
41. Sun J, Li GS. Leaf dorsoventrality candidate gene *CpARF4* has conserved expression pattern but divergent tasiR-ARF regulation in the water fern *Ceratopteris pteridoides*. Am J Bot. 2020;107(11):1470–80.
42. Song S, Hao L, Zhao P, Xu Y, Zhong N, Zhang H, Liu N. Genome-wide identification, expression profiling and evolutionary analysis of auxin response factor gene family in potato (*Solanum tuberosum* group Phureja). Sci Rep-UK. 2019;9(1):1755.
43. Martin-Arevalillo R, Thévenon E, Jégu F, Vinos-Poyo T, Vernoux T, Parcy F, Dumas R. *Evolution of the* auxin response factors from charophyte ancestors. PLoS Genet. 2019;15(9):e1008400.
44. Chen C, Zeng Z, Liu Z, Xia R. Small RNAs, emerging regulators critical for the development of horticultural traits. Hortic Res-Engl. 2018;5(1):14–63.
45. Hu W, Zuo J, Hou X, Yan Y, Wei Y, Liu J, Li M, Xu B, Jin Z. *The* auxin response factor gene family in banana: genome-wide identification and expression analyses during development, ripening, and abiotic stress. Front Plant Sci. 2015;6:742.
46. Powers SK, Holehouse AS, Korasick DA, Schreiber KH, Clark NM, Jing H, Emenecker R, Han S, Tycksen E, Hwang I, et al. Nucleo-cytoplasmic partitioning of ARF proteins controls auxin Responses in *Arabidopsis thaliana*. Mol Cell. 2019;76(1):177–90.
47. Zhang X, Cao J, Huang C, Zheng Z, Liu X, Shangguan X, Wang L, Zhang Y, Chen Z. *Characterization of cotton* ARF factors and the role of GhARF2b in fiber development. BMC Genomics. 2021;22(1):202.
48. Lombardi L, Arrom L, Mariotti L, Battelli R, Picciarelli P, Kille P, Stead T, Munné-Bosch S, Rogers HJ. Auxin involvement in tepal senescence and abscission in *Lilium*: a tale of two lilies. J Exp Bot. 2015;66(3):945–56.
49. Wang H, Chang X, Lin J, Chang Y, Chen J, Reid MS, Jiang C: Transcriptome profiling reveals regulatory mechanisms underlying corolla senescence in petunia. Hortic Res-Engl. 2018, 5(1).
50. Zhang MM, Zhang HK, Zhai JF, Zhang XS, Sang YL, Cheng ZJ. *ARF4* regulates shoot regeneration through coordination with ARF5 and IAA12. Plant Cell Rep. 2021;40(2):315–25.
51. Schuetz M, Fidanza M, Mattsson J. *Identification* of auxin response factor-encoding genes expressed in distinct phases of leaf vein development and with overlapping functions in leaf formation. Plants. 2019;8(7):242.
52. Lee HW, Cho C, Pandey SK, Park Y, Kim M, Kim J. *LBD16* and *LBD18* acting downstream of *ARF7* and *ARF19* are involved in adventitious root formation in Arabidopsis. BMC Plant Biol. 2019;19(1):46.
53. Asakura Y, Ishigaki E, Sugiyama R, Kurosaki F. Cloning and expression of cDNAs encoding ADP-ribosylation factor in carrot seedling. Plant Sci (Limerick). 2007;172(2):189–95.
54. Jincheng-Yuan Jinhui-Song. Hailian-Ma, Xiaoqing-Song, Huiping-Wei, Yinghui-Liu: *Ectopic expression a maize* ADP-ribosylation factor gene in Arabidopsis, increase plant size and growth rate. J Plant Biochem Biot. 2013;24(2):161–6.
55. Damodharan S, Zhao D, Arazi T. A common miRNA160-based mechanism regulates ovary patterning, floral organ abscission and lamina outgrowth in tomato. Plant J. 2016;86(6):458–71.
56. de Jong M, Wolters-Arts M, Schimmel BCJ, Stultiens CLM, de Groot PFM, Powers SJ, Tikunov YM, Bovy AG, Mariani C, Vriezen WH, et al. *Solanum lycopersicum* AUXIN RESPONSE FACTOR 9 regulates cell division activity during early tomato fruit development. J Exp Bot. 2015;66(11):3405–16.
57. Zhang W, Abdelrahman M, Jiu S, Guan L, Han J, Zheng T, Jia H, Song C, Fang J, Wang C. VvmiR160s/*VvARFs* interaction and their spatio-temporal expression/cleavage products during GA-induced grape parthenocarpy. BMC Plant Biol. 2019;19(1):111.
58. Liu Z, Miao L, Huo R, Song X, Johnson C, Kong L, Sundaresan V, Yu X. *ARF2–ARF4* and *ARF5* are essential for female and male gametophyte development in arabidopsis. Plant Cell Physiol. 2018;59(1):179–89.

59. Wang J, Wang R, Mao X, Li L, Chang X, Zhang X, Jing R. *TaARF4* genes are linked to root growth and plant height in wheat. Ann Bot-London. 2019;124(6):903–15.
60. Bouzroud S, Gasparini K, Hu G, Barbosa MAM, Rosa BL, Fahr M, Bendaou N, Bouzayen M, Zsögön A, Smouni A, et al. Down regulation and Loss of *Auxin Response Factor 4* function using CRISPR/Cas9 alters plant growth, stomatal function and improves tomato tolerance to salinity and osmotic stress. Genes-Basel. 2020;11(3):272.
61. Wójcikowska B, Wójcik AM, Gaj MD. Epigenetic regulation of auxin-induced somatic embryogenesis in plants. Int J Mol Sci. 2020;21(7):2307.
62. Quintana-Escobar AO, Nic-Can GI, Galaz Avalos RM, Loyola-Vargas VM, Gongora-Castillo E. Transcriptome analysis of the induction of somatic embryogenesis in *Coffea canephora* and the participation of ARF and Aux/IAA genes. PeerJ. 2019;7:e7752.
63. Ckurshumova W, Smirnova T, Marcos D, Zayed Y, Berleth T. Irrepressible *MONOPTEROS/ARF5* promotes de novo shoot formation. N Phytol. 2014;204(3):556–66.
64. Vidaurre DP, Ploense S, Krogan NT, Berleth T. *AMP1* and *MP* antagonistically regulate embryo and meristem development in *Arabidopsis*. Development. 2007;134(14):2561–7.
65. Xiao Y, Li J, Zhang Y, Zhang X, Liu H, Qin Z, Chen B. Transcriptome analysis identifies genes involved in the somatic embryogenesis of Eucalyptus. BMC Genomics. 2020;21(1):1–803.
66. Wu W, Zhu S, Zhu L, Wang D, Liu Y, Liu S, Zhang J, Hao Z, Lu Y, Cheng T, et al. Characterization of the *Liriodendron Chinense MYB* gene family and its role in abiotic stress response. Front Plant Sci. 2021;12:641280.
67. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202.
68. Chou K, Shen H, Newbigin E. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. PLoS One. 2010;5(6):e11335.
69. Chou K, Shen H. Large-scale plant protein subcellular location prediction. J Cell Biochem. 2007;100(3):665–78.
70. Wu W, Zhu S, Xu L, Zhu L, Wang D, Liu Y, Liu S, Hao Z, Lu Y, Yang L, et al. Genome-wide identification of the *Liriodendron chinense WRKY* gene family and its diverse roles in response to multiple abiotic stress. BMC Plant Biol. 2022;22(1):25.
71. Huo A, Chen Z, Wang P, Yang L, Wang G, Wang D, Liao S, Cheng T, Chen J, Shi J. Establishment of transient gene expression systems in protoplasts from *Liriodendron* hybrid mesophyll cells. PLoS One. 2017;12(3):e172475.

Publisher's Note