

The ARF gene family in the medicinal plant *Macleaya cordata*: genome-wide survey and gene expression profiling

Peng Yang

Hunan University of Medicine

Yuxin Qi

Hunan University of Medicine

Bin Huang

Hunan University of Medicine

xueshuang huang (✉ xueshuanghuang@126.com)

Hunan University of Medicine

Research Article

Keywords: ARFs, Genome-wide, *Macleaya cordata*, Stress treat

Posted Date: August 12th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1934675/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Auxin response factors (ARFs) are a class of transcription factor that direct plant gene responses to auxin, and which is critical for plant growth and development. *Macleaya cordata*, a medicinal plant species from the Papaveraceae family, has not yet been subjected to a genome-wide investigation of the *ARF* gene family. We identified ARFs in *Macleaya cordata* using bioinformatic approaches and studied their structure, chromosome distribution, conserved protein domains, phylogenetic trees, and expression profiles. As a result, 17 ARFs in *Macleaya cordata* were discovered, all of which had the Auxin-resp domain and B3 DNA binding domain. In *Macleaya cordata*, fragment replication events were discovered in four gene pairs, but no tandem repeats of the *ARF* genes were observed. *ARF5*, *ARF3*, *ARF8*, and *ARF16* were all substantially expressed in the roots, which suggested they might be involved in plant stress responses. Furthermore, the expression of *ARF8* and *ARF16* altered significantly when exposed to salt and cold stress. This result adds further evidence that these two *ARF* genes may have a function in plant stress tolerance. Our findings lay the groundwork for understanding the molecular basis and regulatory mechanisms of ARFs in *Macleaya cordata*.

Introduction

The auxin response factor (ARF) family is a plant-specific family of transcription factors that can specifically bind to the auxin response element (AuxRE; TGTCTC) in the promoter region of auxin responsive genes to activate or inhibit gene expression [1–3]. Most ARF proteins have three conserved domains: the B3 DNA binding domain, with DNA binding activity, is responsible for direct binding to AuxREs in the upstream promoter region of a gene as regulated by auxin; the Auxin-resp domain, which can activate or inhibit transcript activity; and the Aux-IAA domain, which is a region that mediates the homodimerization and heterodimerization of Aux/IAAs and ARFs to regulate the expression of downstream auxin-responsive genes (ARGs) via different combinations of ARFs and AuxREs [1, 2, 4].

The *ARF* gene families of many plants have been systematically analyzed, for example, 23 *ARF* genes in *Arabidopsis thaliana*, 31 *ARF* genes in *Zea mays*, and 25 *ARF* genes in *Oryza sativa* have been identified [5–9]. The function of some *ARF* genes has also been identified [3, 5, 10]. *AtARF5* (MONOPTEROS) was the first *ARF* gene identified by map-based cloning in plants, and it plays an important role in the process of embryo establishment and the normal formation of vascular tissue [10]. Furthermore, *AtARF7* and *AtARF19*, the most widely studied *ARF* genes in Arabidopsis, display some redundancy in function and play an important role in lateral root development [11–13]. Attia et al. [14] used transgenic rice lines to demonstrate that reduction in the expression of the *OsARF1* gene, resulted in low seedling vigor, stunting, short leaves, leaf curling, and sterility, indicating that *OsARF1* is important for the development of vegetative and reproductive organs effect in rice.

Macleaya cordata (with the Chinese name “Bo-luo-hui”) is a perennial herb that belongs to the Papaveraceae family and is typically prescribed as a traditional antibacterial medicine [15, 16]. The whole-genome sequencing of *Macleaya cordata* was completed in 2017, which laid a good foundation for molecular biology research of the *Macleaya cordata* plant [17]. In contrast to other plants, there are currently no reports on the analysis of the *ARF* gene family in *Macleaya cordata*, and thus no bioinformatics or expression analyses has been performed for this species. Therefore, a systematic analysis of the *ARF* gene family must first be conducted based on the complete genome sequence of *Macleaya cordata*. This study uses bioinformatic methods to identify *ARF* genes across the genome in *Macleaya cordata* and analyzes their genic structure, chromosome distribution, conserved protein domains, molecular evolution, and expression profiles in order to lay the groundwork for further research into this highly important family of plant-specific transcription factors.

Materials & Methods

Data collection and identification

The *Macleaya cordata* genome was obtained from the National Center for Biotechnology Information (NCBI) [17]. We identified all potential protein sequences in the *Macleaya cordata* genome using BLAST and HMMER (E-value $1e^{-5}$) [18, 19] and utilized the Pfam database to validate the existence of the B3, Auxin resp, and AUX IAA domains in previously reported ARFs [20].

ExPASy software was used to calculate the isoelectric point (pI) and molecular weight (MW) of the ARFs in *Macleaya cordata* [21]. WoLF PSORT software was used to predict the subcellular localization of ARFs in *Macleaya cordata*.

Analysis of the motif, domain, and chromosome distribution

MEME was used to find ARF protein motifs in *Macleaya cordata* *ARF* gene family members [22]. The domain of the ARF proteins in *Macleaya cordata* was confirmed using the CD-Search Tool [23]. TBtools was used to visualize the analysis findings of the motif and domain analysis of the *Macleaya cordata* *ARF* gene family members [24].

Multiple sequence alignment, phylogenetic tree construction, and selection pressure analysis

PlantTFDB was used to obtain all deposited *ARF* gene sequence of these two model plant species (*Arabidopsis thaliana* and *Oryza sativa*) (Supplementary Materials Data 1) [25]. All the ARFs in the three species analyzed in this study were aligned and compared using Clustal Omega software [26]. Based on the alignment findings, we utilized MEGA X (10.0.5) software to build the neighbor-joining (NJ) and maximum-likelihood (ML) trees (Poisson correction and bootstrap = 1,000 repetitions) [27]. EvolView software was used to curate the created NJ tree [28, 29]. The Ka/Ks value of the homologous ARFs in *Macleaya cordata* was calculated using the MCScanX software and KaKs Calculator 2.0 software [30–33].

Analysis of the promoter *cis*-acting elements

PlantCare was used to investigate the numerous *cis*-acting elements in *Macleaya cordata* ARF promoter sequences [34] and TBtools was used to visualise the generated data [24].

Analysis of gene expression

Three groups of two-year-old full-flowering *Macleaya cordata* with favorable and consistent growing conditions were selected for analysis. The control group received no treatment. Treatment 1 was sprayed with 200 mol L⁻¹ NaCl for 3 days, whereas treatment 2 was cultivated at 4°C for 3 days. Based on Trizol method, total RNA was extracted from the roots, stems, leaves, and flowers of control and treated *Macleaya cordata* plants using a plant RNA extraction kit (V1.5) (DNase I) (Chengdu Biofit Biotechnologies Co., Ltd). The reverse transcription kit was HiScript II Q Select RT SuperMix for qPCR (+ gDNA wiper) provided by Vazyme (<http://www.vazyme.com/>). The reaction solution was prepared on an ice box, and the reverse transcription reaction was performed immediately after adding the sample to mix by gently patting. A more specific qPCR process was performed as described previously [35]. The relevant data of the primers for amplification are shown in Supplementary Materials Data 2. Three independent biological replicates were used for this experiment. The expression data in *Macleaya cordata* represents the standardized base number and average value for distinct tissues and developmental phases and is utilized by R software to build a heat map representation. The Spearman correlation distance was used to group genes with similar expression characteristics.

Results

Identification and analysis of ARFs in *Macleaya cordata*

In our investigation, we utilized local BLAST searches to find sequences with high similarity to *Arabidopsis thaliana* ARFs ($E < 0.001$), then it was merged with the Hmmer analysis to find sequences including the Auxin-resp domain (PF06507) and the B3 DNA binding domain (PF02362) as candidates for actual ARF genes in *Macleaya cordata*. Multiple alternate splicing sequences originating from the same location in the *Macleaya cordata* genome were found in the candidate sequences, and the longest transcribed sequence was selected for additional screening to determine all putative ARF candidates. To further characterize specific sequences, we employed the SMART domain database. The incomplete sequences in the discovered coding area were deleted, resulting in the identification of 17 ARF genes in the *Macleaya cordata* genome (Table 1). We discovered substantial changes in the physical and chemical characteristics of the 17 ARF protein sequences by examining their physical and chemical properties: the amino acid (AA) length of the ARF proteins range from 382 (ARF14) to 1,164 AA (ARF11), while the length of the coding region was from 1,146 (ARF14) to 3,492 base pairs (bp) (ARF11), the molecular weight ranged from 42,681.72 (ARF14) to 129,457.11 Da (ARF11), and the isoelectric point ranged from 5.42 (ARF6) to 8.39 (ARF14).

Table 1

Basic information of the auxin response factors (ARFs) in *Macleaya cordata*. At, *Arabidopsis thaliana*; pl, isoelectric point; ORF, open reading frame.

Name	Locus	Homologous Genes in At	Scaffold	Start	End	Strand	pl	MW (Da)	Protein (aa)	ORF (bp)	Subcellular Localization
ARF1	rna397	AT4G23980.1	MVGT01002338.1	442,300	446,922	+	6.94	84,711.97	761	2283	nucl
ARF2	rna643	AT1G30330.1	MVGT01001648.1	436,289	444,624	-	6.15	110,372.07	1009	3027	nucl
ARF3	rna1763	AT1G30330.1	MVGT01001357.1	131,747	140,026	-	6.01	107,621.25	964	2892	nucl
ARF4	rna3297	AT4G30080.1	MVGT01002031.1	303,492	306,455	+	6.66	76,582.34	698	2094	nucl
ARF5	rna5176	AT5G60450.1	MVGT01001732.1	662,290	669,001	+	5.89	95,333.13	865	2595	nucl
ARF6	rna6659	AT1G19850.1	MVGT01001340.1	224,273	229,838	+	5.42	105,853.15	962	2886	nucl
ARF7	rna6914	AT4G23980.2	MVGT01001025.1	257,367	261,233	+	6.03	78,990.78	710	2130	nucl
ARF8	rna8226	AT4G30080.1	MVGT01000437.1	3,382,482	3,385,897	+	6.49	79,758.84	729	2187	nucl
ARF9	rna8271	AT5G60450.1	MVGT01000437.1	3,737,980	3,743,427	+	5.66	89,133.92	803	2409	nucl
ARF10	rna8683	AT5G62000.3	MVGT01000438.1	1,298,346	1,303,370	+	5.94	93,528.96	842	2526	nucl
ARF11	rna11116	AT1G19220.1	MVGT01000213.1	661,537	668,552	-	5.94	129,457.11	1164	3492	nucl
ARF12	rna14079	AT1G19220.1	MVGT01001693.1	59,014	67,293	+	6.22	128,851.22	1146	3438	nucl
ARF13	rna15418	AT1G59750.1	MVGT01004210.1	144,756	152,688	+	5.94	85,076.81	764	2292	nucl
ARF14	rna15781	AT1G77850.1	MVGT01002935.1	194,920	196,065	+	8.39	42,681.72	382	1146	cyto
ARF15	rna17460	AT5G37020.1	MVGT01002401.1	77691	87,128	-	5.88	97,347.21	875	2625	nucl
ARF16	rna21466	AT5G62000.3	MVGT01002051.1	1,412,338	1,417,403	-	6.7	96,563.23	871	2613	nucl
ARF17	rna21764	AT5G62000.3	MVGT01000924.1	961,249	966,266	+	5.8	75,377.31	672	2016	nucl

According to these findings, all ARF proteins in *Macleaya cordata* were found to contain the Auxin-resp domain (PF06507) and the B3 domain (PF02362) (Fig. 1). Except for AtARF12, McARF14, AtARF13, AtARF18, McARF4, and McARF8, the C-terminus of all other ARF protein sequences contained the Aux-IAA domain (PF02309). This suggests that the encoding ARF proteins of *Macleaya cordata* may have function in a similar manner to the ARF proteins of *Arabidopsis thaliana*, and may enhance apical dominance, lateral root formation, floral organ development, and the response to environmental stress.

To investigate the conservation and diversity of ARF proteins further, 10 conserved motifs ($E < 0.1$) were found and examined using the MEME database (Fig. 2 and Supplementary Figure S1). Except for ARF14, which lacks motif5, all the analyzed ARF proteins have list motifs in numerical order motif2, motif6, motif10, motif1, motif3, motif5, motif4, and motif8. We discovered that motif9 and motif7 together make up the Aux-IAA domain. The motif sequence schematic presented in Supplemental Figure S1 showed a high similarity of motifs contained within the three domains of the ARF proteins, demonstrating that these three domains are highly conserved in the ARF family.

Evolutionary analysis of the ARFs in *Macleaya cordata*

To better understand the diversity and evolutionary connections of the ARF proteins in *Macleaya cordata*, we employed a collection of ARF sequences to construct the phylogenetic tree, including the 22 AtARFs, 27 OsARFs, and the 17 McARFs used in this work. The tree was constructed using two methods: NJ (Fig. 3) and ML (Supplementary Figure S2).

The ARFs of the three plant species focused on in our study are clearly split into five branches. *Arabidopsis thaliana* ARF protein and *Macleaya cordata* ARF protein are the most closely linked, while rice ARF proteins were revealed to be more distantly related. This ARF protein grouping is consistent with the general evolutionary connection of these plants.

Homology and selective pressure analysis of the ARF genes in *Macleaya cordata*

The two basic mechanisms for gene family expansion are fragment replication and tandem duplication. We used BLASTP and MCScanX to identify collinear genes with which to analyze the amplification pattern of the ARF gene family in *Macleaya cordata*. Four gene pairs, ARF1 and ARF7, ARF2 and ARF3, ARF4 and ARF8, and ARF10 and ARF16, were shown to have fragment duplication events. Tandem repeats of individual ARF genes were not discovered in any of the ARF genes in *Macleaya cordata*. This shows that fragment replication is the primary method of ARF gene family expansion in *Macleaya cordata*.

Then, we calculated the Ka/Ks (the ratio of the number of nonsynonymous substitutions per nonsynonymous site (Ka) to the number of synonymous substitutions per synonymous site (Ks)) of the segmentally duplicated ARF genes in *Macleaya cordata* using the KaKs Calculator 2.0 to further investigate the potential evolutionary mechanism of the ARF genes in *Macleaya cordata*. We hypothesized that if Ka/Ks = 1, this gene was subjected to pure selection. The results demonstrate that all analysed ARF gene pairs returned values significantly less than one (Table 2), implying that they were subjected to highly purifying selection during evolution.

Table 2
Ka/Ks of the ARF genes in *Macleaya cordata*. The results were calculated and generated by MCScanX and KaKs Calculator 2.0. Ks: synonymous substitution rate; Ka: nonsynonymous substitution rate.

Seq_1	Seq_2	Ka	Ks	Ka_Ks	EffectiveLen	Average	
						S-sites	N-sites
ARF1	ARF7	0.18	0.91	0.20	2106	483.50	1622.50
ARF2	ARF3	0.15	0.78	0.20	2871	661.25	2209.75
ARF4	ARF8	0.10	0.83	0.12	2088	496.08	1591.92
ARF10	ARF16	0.14	0.77	0.19	2505	576.50	1928.50

Cis-element analysis of the ARF gene promoters in *Macleaya cordata*

PlantCare was used to investigate the promoter regions (CDS upstream 2,000 bp) of the ARF genes identified in *Macleaya cordata*. As shown in Fig. 4, even gene pairs that are evolutionarily related to each other can have variances in cis-elements. Most ARFs, in particular, had cis-elements linked to jasmonate acid induction, anaerobic response, and light responsiveness. This may indicate that they are transcriptionally responsive to such environmental conditions. Cis-elements associated to auxin induction were identified in the promoter regions of ARF1, ARF2, ARF4, ARF5, ARF7, ARF10, ARF12, ARF13, ARF14, ARF15, and ARF16.

Expression of ARFs in *Macleaya cordata*

We used RT-qPCR to examine the tissue-specific expression pattern of each ARF gene to further understand the features and functional impacts of ARF genes in *Macleaya cordata*. The heatmap depicts the findings of ARF expression in four *Macleaya cordata* tissues (i.e., flowers, leaves, stems, and roots) (Fig. 5).

The 17 *ARFs* were grouped into five subgroups based on their expression domains. Tissue specificities differed between groupings. Only *ARF11* was found to be strongly expressed in flowers, suggesting that it may play a role in floral development. *ARF5*, *ARF8*, *ARF3*, and *ARF16* were found to be strongly expressed in roots, suggesting that they may mediate roles in responding to the surrounding environment. The high expression of *ARF3* and *ARF6* in leaves suggests that these two *ARFs* may play roles in leaf growth and development. *ARF6*, *ARF12*, and *ARF17* were found to be highly expressed in stems. *ARF1*, *ARF9*, *ARF10*, and *ARF15* were found to be strongly expressed in the leaves and stems. The expression of *ARF2*, *ARF4*, *ARF7*, *ARF13*, and *ARF14* was revealed by RT-qPCR to be highly concentrated in stem and root tissues. The low expression of *ARF1*, *ARF6*, *ARF9*, *ARF10*, *ARF12*, *ARF15*, and *ARF17* in roots may suggest that they have an inhibitory effect on root growth and development, but further experimental verification is needed.

The gene expression analysis revealed that several *ARFs* (*ARF3*, *ARF5*, *ARF8*, and *ARF16*) had a high level of expression in the roots. High levels of gene expression in the root system of a plant may indicate a role for the gene in responding to environmental stress, thus we used RT-qPCR to further examine the expression of *ARF3*, *ARF5*, *ARF8*, and *ARF16* in *Macleaya cordata* plants exposed to salt and cold stress. As shown in Fig. 6, under the conditions of salt and cold stress, the expression of each assessed ARF changed to differing degrees. The expression of *ARF5* in the roots of salt and cold treated plants decreased significantly. The expression of *ARF8* in the stems of salt and cold stressed plants increased significantly. The expression of *ARF16* in salt-treated roots increased significantly, which may indicate that it plays a very important role in stress resistance. However, their specific functions need further experimental analysis.

Discussion

ARFs are a family of plant-specific transcription factor that controls the expression of ARGs. In this investigation, we identified 17 ARF loci across the 16 scaffolds that represent the *Macleaya cordata* genome (Fig. 4). The number of *ARF* genes encoded in *Arabidopsis thaliana* and *Oryza sativa* was greater than the number of *ARF* genes detected in this study, a finding that likely stems from the lack of completeness of the *Macleaya cordata* genome. We discovered that most ARF proteins in *Macleaya cordata* had the Aux-IAA domain at the C-terminus, exactly as has been previously reported for the ARF proteins of *Arabidopsis thaliana* [36], based on protein structure analysis (Figs. 1 and 2). This suggests that the *ARF* gene family is remains reasonably highly conserved across plant species.

Gene expression analyses (Fig. 5) revealed that *Macleaya cordata* ARFs are expressed in a range of tissues and frequently display tissue-specific expression. This finding is similar with earlier *ARF* gene expression studies in *Arabidopsis thaliana* and other plants, indicating that the ARF proteins may mediate a role in the normal physiological process of *Macleaya cordata* [5, 9, 37].

The ARF gene family controls plant growth and auxin metabolism and so plays a critical role in plant development. The entire genome sequence of *Macleaya cordata* was used as a background for studying the gene structure, chromosome distribution, conserved protein domain, molecular evolution, and expression profile of the *Macleaya cordata* ARF gene family to better understand the structure and function of the ARF in *Macleaya cordata*. While we feel the results reported here are reliable, there are several caveats to be aware of, such as the fact that the genome of *Macleaya cordata* is not yet complete, and further study is needed to fully understand this thorny topic.

Conclusions

In *Macleaya cordata*, we found 17 ARFs and investigated them thoroughly. The collinearity analysis showed that fragment replication events were discovered in four *ARF* gene pairs, but no tandem repeats of the *ARF* genes were observed. *ARF5*, *ARF3*, *ARF8*, and *ARF16* were all substantially expressed in the roots, which suggested they might be involved in plant stress responses. Furthermore, the expression of *ARF8* and *ARF16* altered significantly when exposed to salt and cold stress. All in all, our findings serve as a basis for future study into the regulation of this gene family's function, as well as a reference for revealing the gene family's regulatory mechanism in terms of its significance in plant growth and development.

Declarations

Statements and Declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contribution

Xueshuang Huang and Peng Yang designed the study, Peng Yang and Yuxin Qi performed the treatments, Peng Yang carried out the real time experiments, Peng Yang, Hailang He and Xin Huang carried out the promoter studies, Peng Yang, Yuxin Qi and Hailang He analysed the data and Peng Yang and Xueshuang Huang wrote the paper.

Compliance with Ethical Standards

This article does not contain any studies with human participants or animals performed by any of the authors.

Fundings

This research was supported by Scientific Research Incubation Fond of Hunan University of Medicine (No. 20KJPY07) and Hunan Provincial Natural Science Foundation of China (2021JJ40386).

References

1. Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. *Plant Cell* 15(2):533–543
2. Guilfoyle TJ, Hagen G (2007) Auxin response factors. In 10:453–460
3. Ellis CM, Nagpal P, Young JC, Hagen G, Guilfoyle TJ, Reed JW (2005) AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* 132(20):4563–4574
4. Abel S, Oeller PW, Theologis A (1994) Early auxin-induced genes encode short-lived nuclear proteins. *Proc Natl Acad Sci USA* 91:326–330
5. Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui A, Nguyen D et al (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: Unique and overlapping functions of ARF7 and ARF19. *Plant Cell* 17(2):444–463
6. Kalluri UC, Difazio SP, Brunner AM, Tuskan GA (2007) Genome-wide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*. *BMC Plant Biol* 7:59–59
7. Kumar R, Tyagi AK, Sharma AK (2011) 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* 285(3):245–260
8. Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Han B, Tao Y (2007) Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). *Gene* 394:13–24
9. Xing H, Pudake RN, Guo G, Xing G, Hu Z, Zhang Y, Sun Q, Ni Z (2011) : **Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize.** *BMC Genomics* 12
10. Hardtke CS, Berleth T (1998) The *Arabidopsis* gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO J* 17(5):1405–1411
11. Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui A, Nguyen D et al (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: Unique and overlapping functions of ARF7 and ARF19. *Plant Cell* 17:444–463
12. Liscum E, Briggs WR (1995) Mutations in the NPH1 locus of *Arabidopsis* disrupt the perception of phototropic stimuli. *Plant Cell* 7:473–485
13. Stowe-Evans EL, Harper RM, Motchoulski AV, Liscum E (1998) NPH4, a conditional modulator of auxin-dependent differential growth responses in *Arabidopsis*. *Plant Physiol* 118:1265–1275
14. Attia KA, Abdelkhalik AF, Ammar MH, Wei C, Yang J, Lightfoot DA, El-Sayed WM, El-Shemy HA (2009) : **Antisense phenotypes reveal a functional expression of OsARF1, an auxin response factor, in transgenic rice.** *Current Issues in Molecular Biology* 11(SUPPL. 1)
15. Kosina P, Gregorova J, Gruz J, Vacek J, Kolar M, Vogel M, Roos W, Naumann K, Simanek V, Ulrichova J (2010) Phytochemical and antimicrobial characterization of *Macleaya cordata* herb. *Fitoterapia* 81:1006–1012
16. Sai CM, Li DH, Xue CM, Wang KB, Hu P, Pei YH, Bai J, Jing YK, Li ZL, Hua HM (2015) Two Pairs of Enantiomeric Alkaloid Dimers from *Macleaya cordata*. *Org Lett* 17:4102–4105
17. Liu X, Liu Y, Huang P, Ma Y, Qing Z, Tang Q, Cao H, Cheng P, Zheng Y, Yuan Z et al (2017) The Genome of Medicinal Plant *Macleaya cordata* Provides New Insights into Benzylisoquinoline Alkaloids Metabolism. *Mol Plant* 10:975–989
18. Wheeler TJ, Eddy SR (2013) Hmmer: DNA homology search with profile HMMs. *Bioinformatics* 29:2487–2489
19. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10:421
20. El-Gebali S, Mistry J, Bateman A, Eddy S, Luciani A, Potter S, Qureshi M, Richardson L, Salazar G, Smart A et al (2018) : **The Pfam protein families database in 2019.** *Nucleic acids research* 47
21. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, Bairoch A (2005) : **Protein Identification and Analysis Tools on the ExPASy Server.** In: *The Proteomics Protocols Handbook*. Edited by Walker JM. Totowa, NJ: Humana Press; : 571–607
22. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) : **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res* 37(Web Server issue):W202-208
23. Lu S, Wang J, Chitsaz F, Derbyshire M, Geer R, Gonzales N, Gwadz M, Hurwitz D, Marchler G, Song J et al (2019) : **CDD/SPARCLE: the conserved domain database in 2020.** *Nucleic acids research* 48
24. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* 13(8):1194–1202
25. Guo A, Chen X, Gao G, Zhang H, Zhu Q-H, Liu X, Zhong Y-F, Gu X, He K, Luo J (2008) PlantTFDB: A comprehensive plant transcription factor database. *Nucleic Acids Res* 36:D966–969
26. Fábó M, mi PY, Joon L, Nicola B, Tamer G, Nandana M, Prasad B, N TAR, C PS, Research FRDJNA (2019) : **The EMBL-EBI search and sequence analysis tools APIs in 2019(W1):W1**
27. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549
28. he Z, Zhang H, Gao S, Lercher M, Chen W-H, Hu S (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res* 44:gw370

29. Zhang H, Gao S, Lercher M, Hu S, Chen W-H (2012) EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Res* 40:W569–572

30. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H et al (2012) MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* 40:1–14

31. Wang D, Wan H-L, Zhang S, Yu J (2009) γ-MYN: A new algorithm for estimating Ka and Ks with consideration of variable substitution rates. *Biol Direct* 4:20

32. Wang D, Zhang S, He F, Zhu J, Hu S, Yu J (2009) How do variable substitution rates influence Ka and Ks calculations? *Genom Proteom Bioinform* 7(3):116–127

33. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J (2010) KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinform* 8(1):77–80

34. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30:325–327

35. Wang K, Cheng Y, Yi L, He H, Zhan S, Yang P (2021) Genome-wide identification of the Tubby-Like Protein (TLPs) family in medicinal model plant *Salvia miltiorrhiza*. *PeerJ* 9:e11403

36. Ulmasov T, Hagen G, Guilfoyle TJ (1999) Dimerization and DNA binding of auxin response factors. *Plant J* 19(3):309–319

37. Wang D, Pei K, Fu Y, Sun Z, Li S, Liu H, Tang K, Han B, Tao Y (2007) Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). *Gene* 394(1–2):13–24

Figures

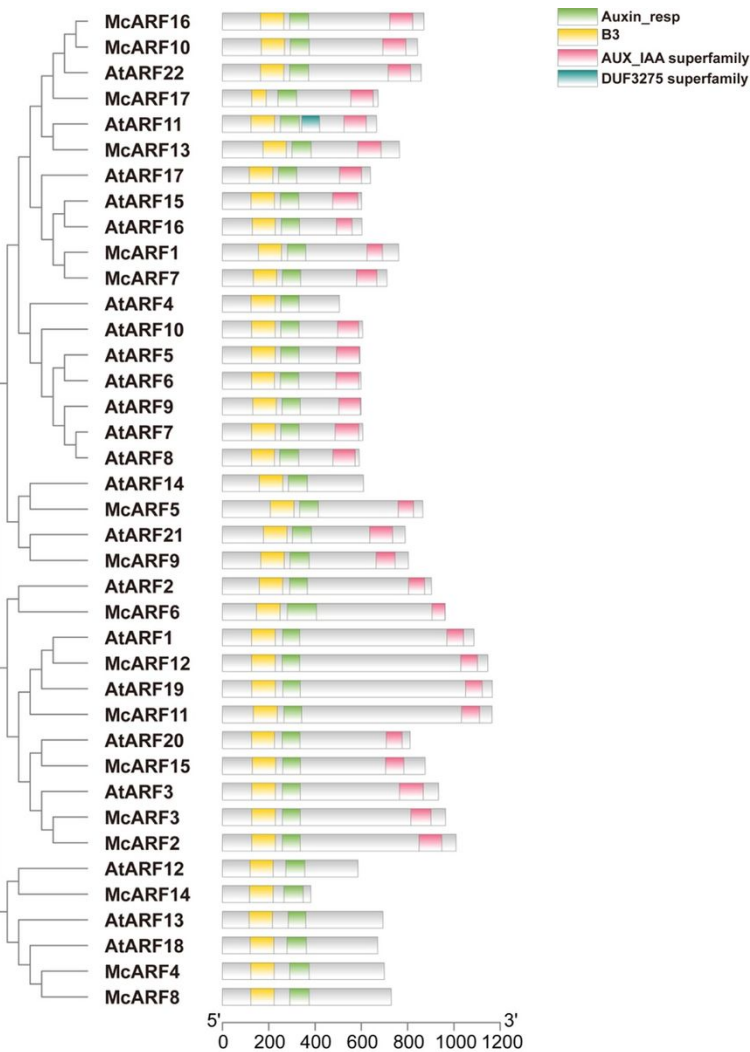


Figure 1

The domain structure of ARFs in *Macleaya cordata* and *Arabidopsis thaliana*. Different domains, including Auxin-resp, B3, AUX-IAA superfamily, and DUF3275 superfamily are represented by colorful blocks.

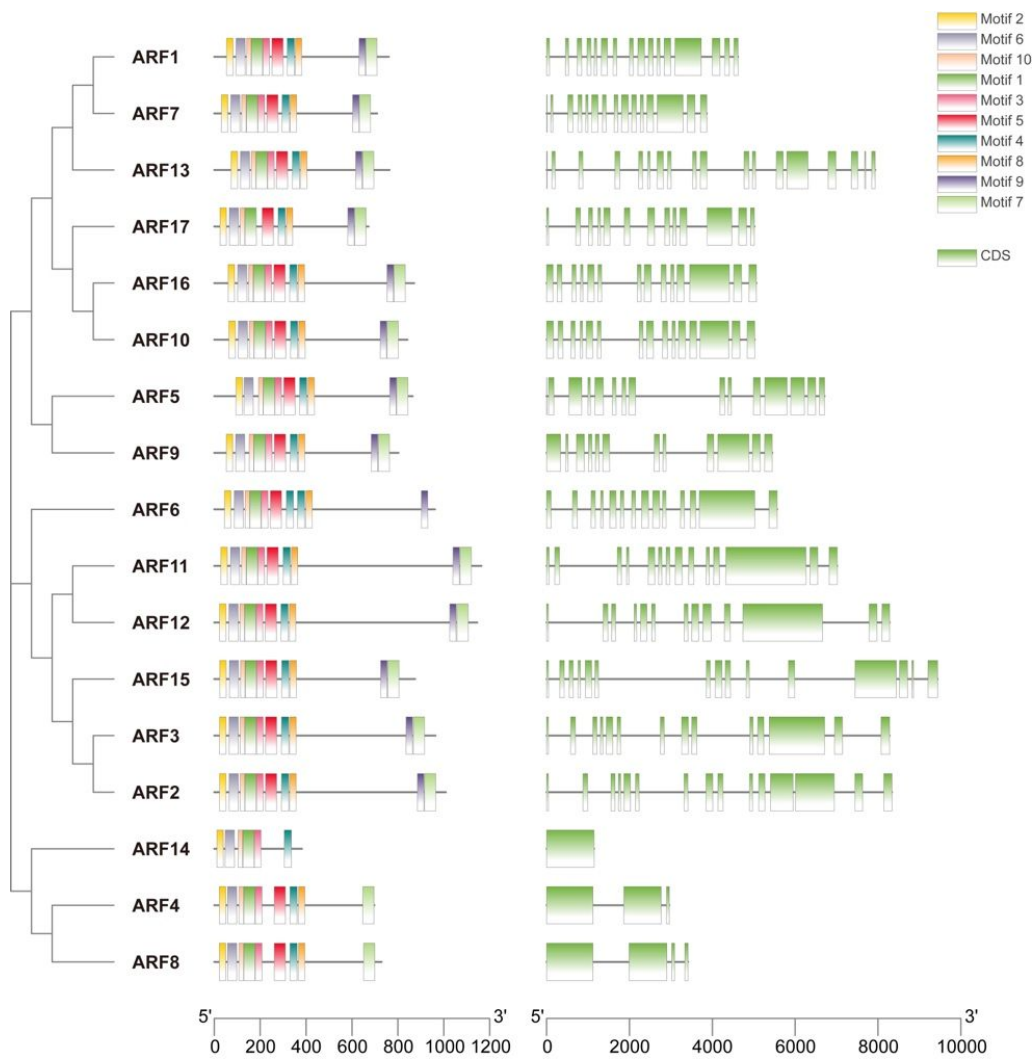


Figure 2

The motif and gene structure of the ARFs in *Macleaya cordata*. Conserved motifs are showed in different colors. For specific sequence information of motif, please refer to the Supplementary Figure S1.

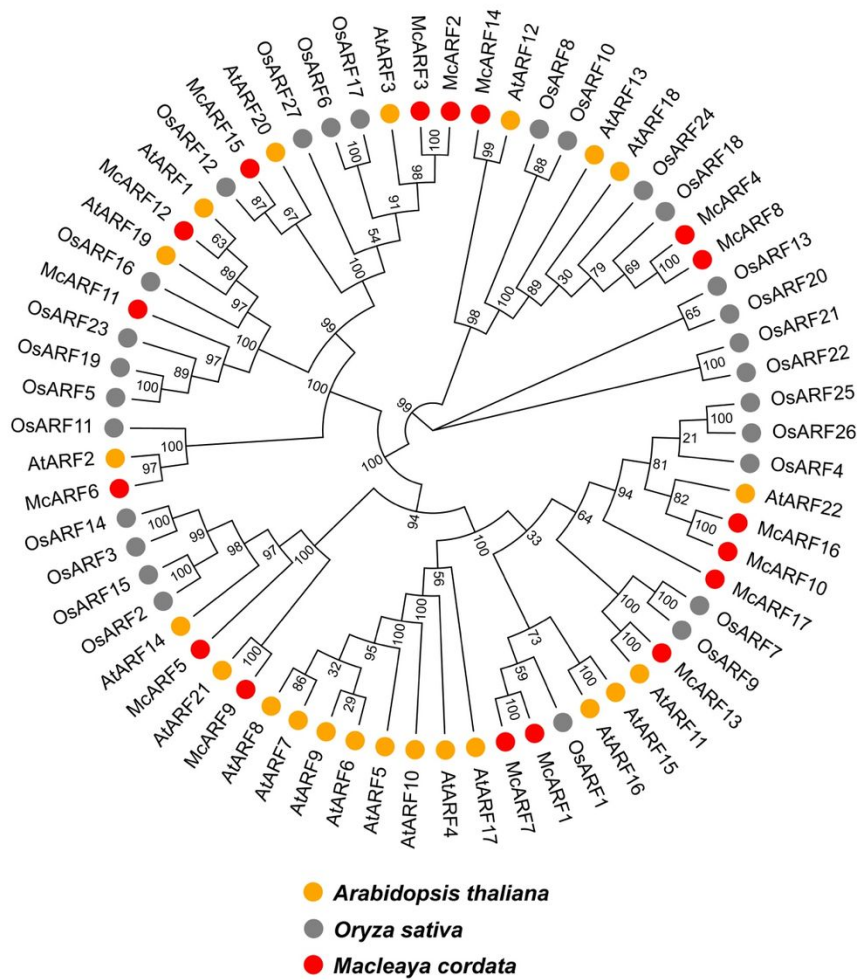


Figure 3

The NJ tree of the ARF proteins in *Arabidopsis thaliana*, *Oryza sativa*, and *Macleaya cordata*. Each was separated into five classes as indicated by a distinct hue. Different labels are used to identify individuals of various species and the bootstrap values are displayed on the branches.

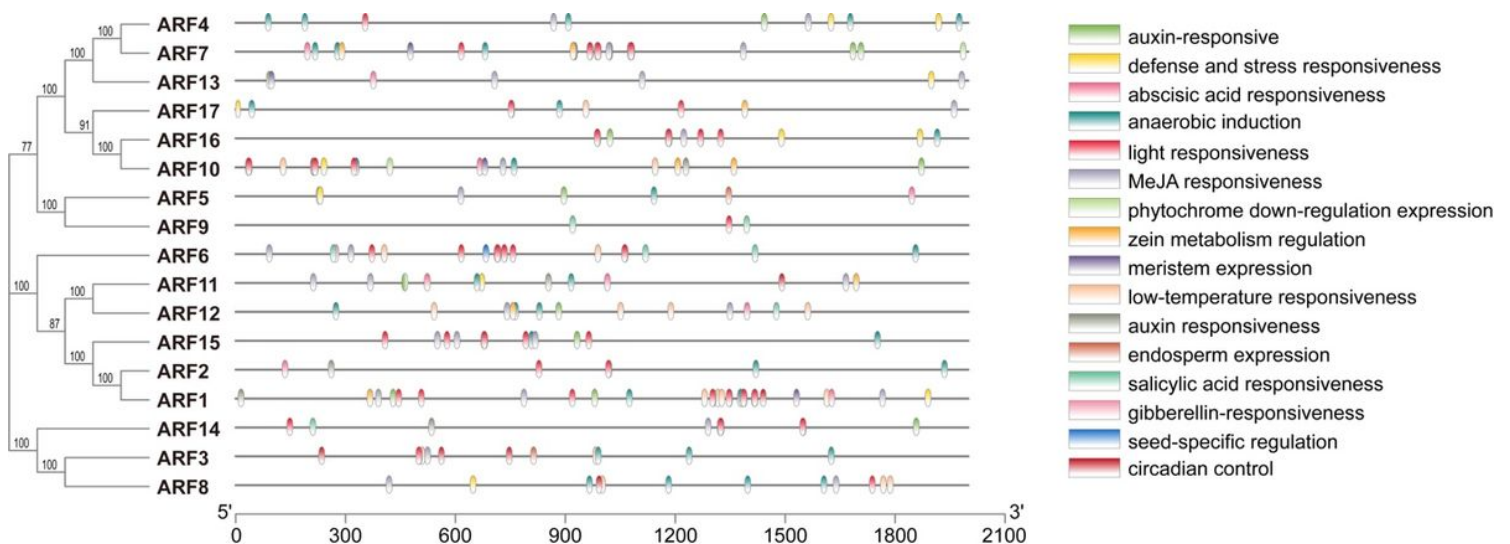


Figure 4

Cis-elements of the ARF gene promoter regions of *Macleaya cordata*. These *cis*-elements are represented by different colors, and were indicated by their positions on the promoters of ARF genes. MeJA, methyl jasmonate.

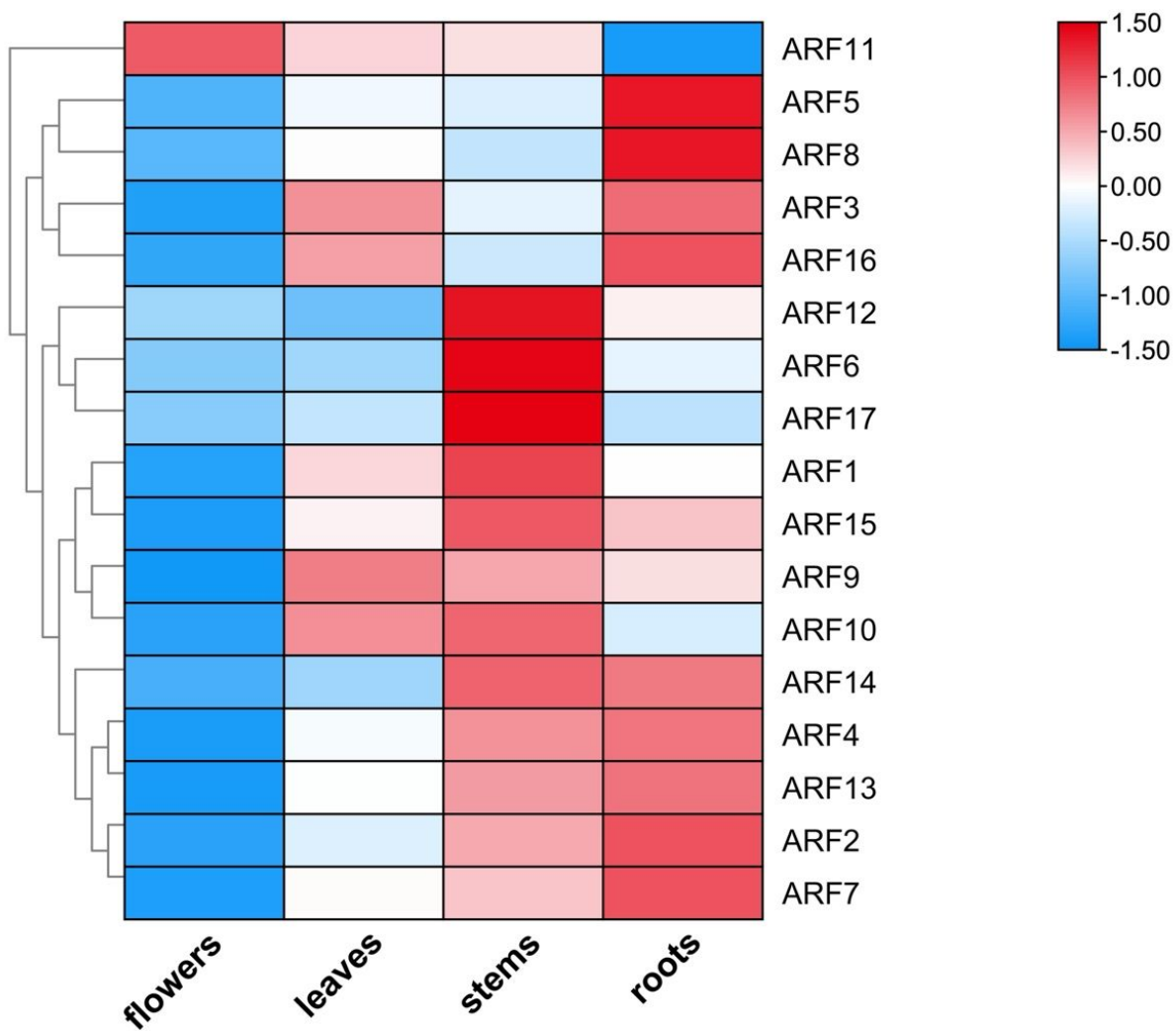


Figure 5

ARF gene expression in different *Macleaya cordata* tissues. Three independent biological replicates were used for this experiment. Data were converted to log2FC and visualized by heat map.

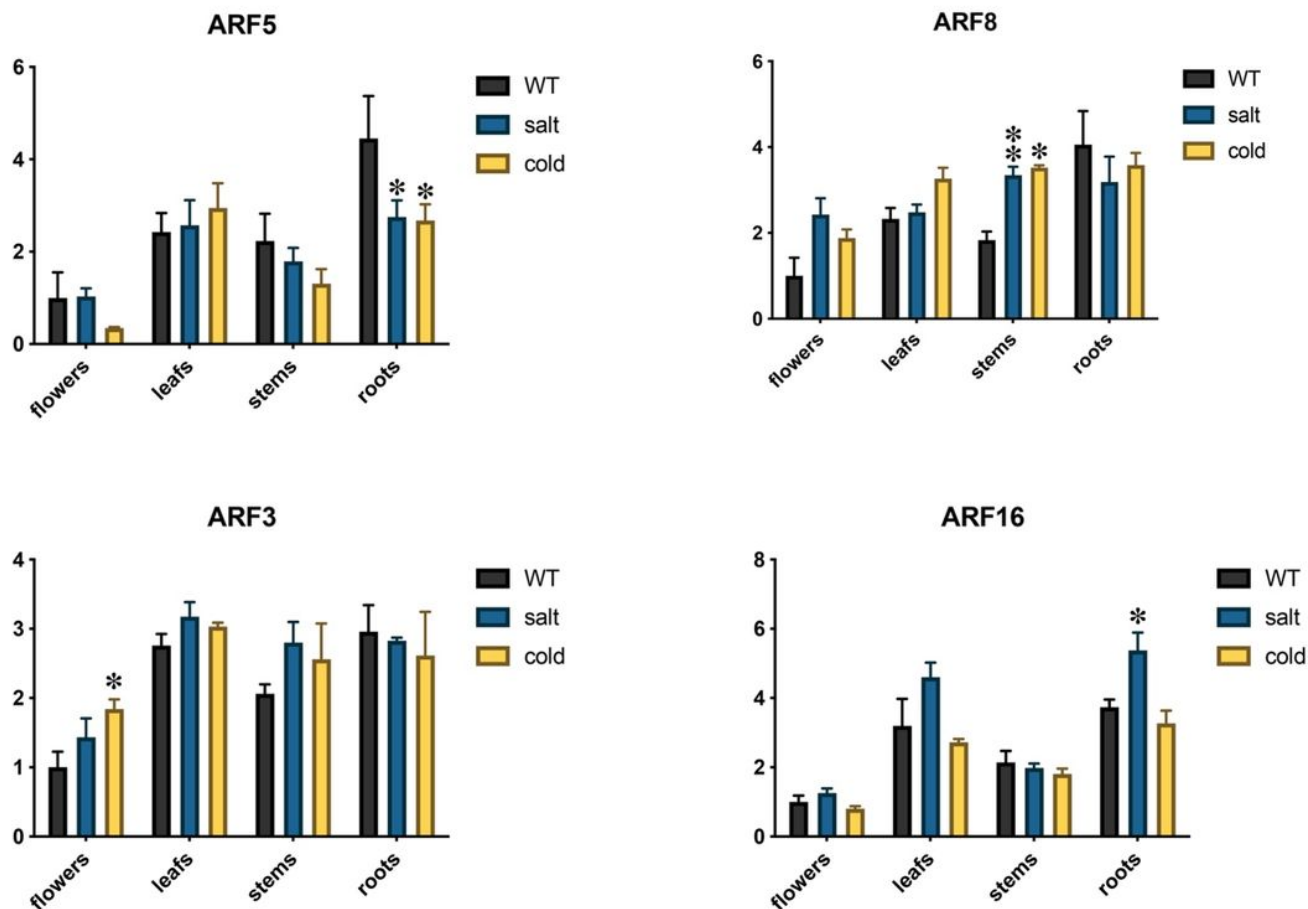


Figure 6

Relative expression levels of *Macleaya cordata* ARF genes under salt stress and cold stress. WT: Wild type, no treatment; salt: salt treat; cold: cold treat. Three independent biological replicates were used for this experiment. *, $p < 0.05$; **, $p < 0.01$ (Student's t-test).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SFigure1.jpg](#)
- [SFigure2.jpg](#)
- [SupplementaryMaterialsData1.xlsx](#)
- [SupplementaryMaterialsData2.xlsx](#)
- [SupplementaryMaterialsData3.xlsx](#)