

Article

Not peer-reviewed version

Genome-Wide Characterization and Identification of Arf Gene Family Reveals the Regulation of RrARF5 in AsA Metabolism in *Rosa roxburghii* Tratt. Fruits

[Tu Feng](#)[†], [Zhengliang Sun](#)[†], [Mingchun Liu](#), [Hong Zhao](#), Yizhong Zhang, [Pedro Garcia-Caparros](#), Bin Yang, [Yingdie Yang](#)^{*}

Posted Date: 28 July 2025

doi: 10.20944/preprints202507.2203.v1

Keywords: *R. roxburghii*; ARF; genome-wide identification; fruit development; vitamin C



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Genome-Wide Characterization and Identification of ARF Gene Family Reveals the Regulation of RrARF5 in AsA Metabolism in *Rosa roxburghii* Tratt. Fruits

Tu Feng ^{1,†}, Zhengliang Sun ^{2,†}, Mingchun Liu ², Hong Zhao ¹, Yizhong Zhang ¹, Pedro Garcia-Caparrós ³, Bin Yang ⁴ and Yingdie Yang ^{1,*}

¹ School of Ecological Engineering, Guizhou University of Engineering Science, Bijie 551700, China

² Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, 610065, China

³ Higher Engineering School, University of Almeria, 04120, Almeria, Spain

⁴ Deyang Guoke Dual Carbon Research Institute, Deyang, 618221, China

* Correspondence: yingdiey2@163.com

† These authors contributed equally to this work.

Abstract

Rosa roxburghii Tratt., a fruit crop known for its high vitamin C content and other nutritional compounds, has not yet been studied for its Auxin Responsive Factor (ARF) gene family, which plays a crucial role in plant growth and fruit development. In the present study, we identified 14 ARF genes (designated as RrARFs) in *R. roxburghii*, which are distributed across seven chromosomes and grouped into four subfamilies. Cis-acting elements analyses revealed that these genes might be involved in various biological processes, including plant development, light responses, cell cycle regulation, phytohormone responses, and responses to abiotic and biotic stresses. Gene expression analysis demonstrated differential expression of RrARF genes across different tissues and stages of fruit development, with four members showing higher expression during the fruit ripening stages. Furthermore, coexpression analysis identified that *RrARF5* was highly coexpressed with *RrMDHAR1*, a key enzyme involved in vitamin C biosynthesis. Moreover, transactivation assays and transient overexpression experiments confirmed that *RrARF5* activates the transcription of *RrMDHAR1*. The outcomes of this study enhance our understanding of the ARF gene family's role in vitamin C accumulation in *R. roxburghii* and its broader impact on fruit quality regulation.

Keywords: *R. roxburghii*; ARF; genome-wide identification; fruit development; vitamin C

1. Introduction

Auxin is a significant plant hormone that plays a pivotal role in various plant developmental processes, including embryogenesis, vascular bundle formation, flower and fruit development, and root growth (Du et al., 2012). The transcriptional regulation of several gene families, such as Gretchen Hagen 3 (GH3), Aux/IAA, and Auxin Response Factor (ARF) families, has been shown to influence auxin synthesis and transport, thereby affecting overall plant growth and development (Tiwari et al., 2001; Park et al., 2007; El-Sharkawy et al., 2014). Of these families, ARF plays a significant role in regulating plant growth and development by binding to the TGTCTC DNA motif, mediating hormonal responses, and consequently affecting plant developmental processes (Guilfoyle et al., 1998).

A typical ARF consists of three distinct structural domains. The first is the highly conserved N-terminal B3-like DNA-binding domain (DBD), which is present in all ARFs and is essential for

binding to target DNA sequences. The second domain is the C-terminal dimerization domain (CTD), which facilitates the formation of heterodimers with proteins from the Aux/IAA family. The third domain is the variable intermediate region (MR), which can act either as an activation domain (AD) or a repression domain (RD) (Guilfoyle and Hagen, 2007; Zhou et al., 2017; Ye et al., 2020). The regulatory activity of ARF depends on the proper presence and interaction of all three domains (Wang et al., 2007).

The identification of ARF gene families has been accomplished in numerous plant species, including *Arabidopsis thaliana* L. (Yoko et al., 2005), *Fragaria vesca* L. (Wang et al., 2018), *Vitis vinifera* L. (Wan et al., 2014), *Citrus sinensis* L. (Li et al., 2015), *Capsicum annuum* L. (Zhang et al., 2017), *Solanum lycopersicum* L. (Kumar et al., 2011) and *Oryza sativa* L. (Wang et al., 2007). ARF family genes play critical roles in various developmental and physiological processes in plants, such as embryogenesis (Dolf et al., 2014), lateral root growth (Yoko et al., 2007), leaf expansion (Wilmoth et al., 2010), and leaf senescence (Lim et al., 2010). In addition, ARF gene family members have also been shown to function as crucial regulators in fruit development, ripening, and quality metabolism (Juan et al., 2007; Vriezen et al., 2008; Yuan et al., 2018; Yuan et al., 2019; Yi et al., 2022; Gambhir et al., 2024). MdARF5 was reported to activate the expression of ethylene-related genes, thereby initiating apple fruit ripening (Yue et al., 2020). Similarly, CpARF2 is involved in the regulation of fruit ripening in papaya by promoting the transcriptional activity of *CpEIL1* (Zhang et al., 2020). In tomato, SlARF10 and SlARF6A have been demonstrated to facilitate chlorophyll accumulation by activating *SlGLK1* expression (Yuan et al., 2018; 2019). However, the role and mode of action of ARF in Vc biosynthesis and metabolism remain largely unknown.

The wild fruit crop *R. roxburghii*, belonging to the genus *Rosa* within the family Rosaceae, is widely distributed in the mountainous regions of southwestern and south-central China. Its cultivation has expanded significantly, with an estimated planting area exceeding 30,000 hectares and an annual yield of approximately 2,000–3,000 tons (Lu et al., 2016). The fruit is edible and characterized by a unique aroma, crunchy texture, and a slightly acidic and astringent flavor (Yang et al., 2014). Notably, *R. roxburghii* is recognized for its exceptional nutritional profile, particularly its high content of ascorbic acid (Vitamin C) (Liu et al., 2016). Ascorbic acid is a vital metabolite for most organisms, serving as an antioxidant and cofactor in various biological processes, including stress resistance, phytohormone biosynthesis, cell division, and cell expansion (An et al., 2007). During plant growth, AsA accumulation is primarily regulated by its biosynthesis, metabolism and translocation. So far, four major biosynthetic pathways have been identified in plants: the L-galactose pathway, the inositol pathway, the D-galacturonic acid pathway and the L-gulose pathway. AsA metabolism is further divided into degradation and recycling pathways. Within the L-galactose pathway, L-galactose dehydrogenase (GalDH) functions as a key precursor enzyme, while L-galactono-1,4-lactone dehydrogenase (GalLDH) catalyzes the final step, converting L-galactono-1,4-lactone into AsA. Two additional enzymes, GDP-L-galactose phosphorylase (GGP) and GDP-D-mannose 3', 5'-differential isomerase (GME), are critical regulatory components of this biosynthetic pathway. AsA also plays an essential physiological role in scavenging reactive oxygen species (ROS), serving as a specific electron donor in the reduction of hydrogen peroxide (H_2O_2) to H_2O and O_2 . The recycling of oxidized AsA is mediated by monodehydroascorbate reductase (MDHAR) and dehydroascorbate reductase (DHAR), which are crucial for maintaining cellular AsA levels. In previous studies of kiwifruit, the enzymatic activities of GalDH, GalLDH, MDHAR, and DHAR showed a similar trend of AsA accumulation, with significantly positive correlations observed between enzyme activity and AsA content, which suggesting a coordinated regulation (Liao et al., 2021). Since humans and many other animals lack the endogenous capacity to synthesize AsA, it must be obtained from dietary sources like fresh fruits and vegetables (Smirnoff 2018). The pattern of AsA accumulation varies among fruit species: in kiwifruit and apple, peak AsA levels are observed during the young fruit stage; in jujube, during fruit expansion stage; and in strawberry and tomato, during fruit ripening stage (Liao et al., 2023). Previous studies have shown that AsA content increases progressively in *R. roxburghii* throughout fruit development and ripening process (Zong et al., 2023).

Due to these characteristics, *R. roxburghii* is renowned for its exceptionally high AsA concentration, earning it the title "King of Vitamin C" and becoming a commercially significant fruit crop. Therefore, elucidating the regulatory mechanisms underlying AsA accumulation during *R. roxburghii* fruit development holds substantial significance for both crop improvement and nutritional enhancement.

Recently, genome-wide data for *R. roxburghii* have been assembled (Zong et al., 2023). However, despite the ARF gene family is widely studied in other species, its role in *R. roxburghii* remains unexplored. Given the ARF family's known roles in fruit development and ripening, it is likely that these genes play a similar role in *R. roxburghii*. Investigating the presence and function of ARF genes in this species could offer critical insights into the mechanism behind its high ascorbic acid accumulation during the ripening stage, potentially leading to strategies for enhancing the fruit quality of *R. roxburghii*.

2. Materials and Methods

2.1. Identification of ARF Family Genes in *R. roxburghii* Genome

The *R. roxburghii* genome sequence was obtained from the CNSA (CNGB Nucleotide Sequence Archive) with the accession number CNP0004212 (<https://db.cngb.org/cnsa/>) (Zong et al., 2023). To effectively identify the ARF family members in the *R. roxburghii* genome, candidate protein sequences containing the Auxin Resp (PF06507) and B3 DNA-binding domain (PF02362) structural motifs were downloaded from the Pfam protein family database (<http://Pfam.xfam.org/>). Candidate ARF genes were identified from the *Rosa roxburghii* genome ($e\text{-value} \leq e^{-10}$) using the "Simple HMM search" in the TBtools software (Chen et al. 2020). Using the NCBI-CDD (<https://www.ncbi.nlm.nih.gov/cdd>) and SMART (<http://smarteml.de/>) databases to confirm each putative RrARFs. Using the SMART database (<http://SMART.embl-heidelberg.de/>), Markov model (HMM) files were utilized to confirm each putative RrARFs. The number of amino acid residues (AA), molecular weight (MW), and theoretical isoelectric point (pI) of the RrARFs proteins were calculated using ExpAsy software (<http://cn.expasy.org/tools>).

2.2. Chromosomal Distribution of RrARFs

The locations of RrARFs on chromosomes were extracted from the genome annotation files. The genes densities across the entire chromosomes were determined and visualized using TBtools software (Chen et al., 2020). For further analysis of conserved motifs and domains, only RrARFs that were anchored to chromosomes were displayed.

2.3. Analysis of Cis-Acting Elements in RrARFs Promoters

The 2000 bp sequences upstream of the initiation codon (ATG) of each *R. roxburghii* ARF gene (RrARFs) were extracted from the *R. roxburghii* genome database using TBtools software (Chen et al., 2020). PlantCARE database was used to further analyze the cis-acting regulatory elements presented in the promoter regions of ARF family genes (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

2.4. Phylogenetic Analysis of the *R. roxburghii* ARF Gene Family

The ARF protein sequences from *R. roxburghii* (14 sequences), *Arabidopsis thaliana* (23 sequences), *Fragaria vesca* (12 sequences), *Malus domestica* (29 sequences) and *Vitis vinifera* (19 sequences) were subjected to multiple sequence alignment using the default settings of Clustal W in MEGA 7 software. A phylogenetic tree was then constructed based on the results of the sequence alignment using the neighbor-joining method, in which the bootstrap value was set to 1000 (Kumar et al., 2016). The protein sequences of ARFs from *Arabidopsis thaliana*, *Fragaria vesca*, *Malus domestica* and *Vitis vinifera* species were obtained from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). The final phylogenetic tree was edited and visualized using the iTOL online tool (<https://itol.embl.de/>).

2.5. Gene Structure and Conserved Motifs Characterization

Conserved motifs in RrARFs were identified using the MEME server (<https://meme-suite.org/>) with the following parameters: a maximum number of 10 motifs, a minimum motif width of 6, and a maximum motif width of 50. We performed structural domain analysis using the NCBI Conserved Domain Database (CDD) to ascertain the types and location of structural domains within all RrARF sequences. In addition, we used TBtools software to visualize the exon/intron structures of the RrARFs genes, as well as the conserved motifs and structural domains of the corresponding RrARF proteins (Chen et al., 2020).

2.6. Gene Expression Analysis of ARF Family Genes

RNA-Seq data from different tissues and organs of *R. roxburghii*, including root, stem, leaf, flower, and fruits harvested in different developmental and ripening stages were downloaded from the CNSA (<https://db.cngb.org/cnsa/>) under accession number CNP0004212 (Zong et al., 2023). Data filtering and quality control were performed using fastp software, and the resulting clean data were used for subsequent analyses (Chen et al., 2018). The *R. roxburghii* genome was used as a reference for read alignment (<https://db.cngb.org/cnsa/>), and String Tie was used to quantify the aligned reads. Differential expression TPM (Transcripts Per Kilobase of exon model per Million mapped reads) across different tissues was performed using TBtools (Chen et al., 2020). The normalized data were then used to plot expression heat maps with TBtools.

2.7. Weighted Correlation Network Analysis (WGCNA)

The top 50% of genes, ranked according to their expression levels, were selected for WGCNA analysis. After an initial threshold screening, a soft-threshold power of $\beta=7$ was selected to strengthen the original proportionality matrix to obtain the derivation of an adjusted neighbor-joining matrix. To better assess the correlation of expression patterns among genes, the neighbor-joining matrices were further transformed into topological overlap matrices (TOMs). Subsequently, gene-shearing analysis was performed using a topological heterogeneity dissimilarity matrix, defined as $\text{dissTOM} = 1 - \text{TOM}$. Clustering and module partitioning were then performed, with a minimum module size set at thirty genes ($\text{minModuleSize}=30$). Module merging was based on a similarity threshold of 0.3 ($\text{cutHeight}=0.3$), and the network type was defined as "signed" ($\text{type} = \text{"signed"}$ or $\text{networkType} = \text{"signed"}$). Visualization of the network was performed using Cytoscape software (Otasek et al., 2019).

2.8. Dual-Luciferase (Dual-LUC) Reporter Assay

To investigate the transcriptional regulation of *RrMDHAR1* by RrARF5, the coding sequence of RrARF5 was cloned into the pGreenII 62SK vector to function as the effector construct. The promoter sequence of *RrMDHAR1* was cloned into the pGL3 vector as the reporter construct. Leaf protoplasts from *N. benthamiana* were then prepared following the methodology reported by Pei et al. (2024). The dual luciferase reporter assay kit was used to quantify the transcriptional activity, according to the manufacturer's instructions.

2.9. Accession Numbers

The accession numbers of VvARFs genes in *Vitis vinifera* were reported in the study of Wan et al. (2014). The accession numbers of AtARFs genes in *Arabidopsis thaliana* and MdARFs genes in *Malus domestica* were reported in the study of Luo et al. (2014). The accession numbers of FvARFs genes in *Fragaria vesca* were reported in the study of Wang et al. (2018).

3. Results

3.1. Identification of ARF Family Genes in *R. roxburghii* Genome

Following an HMMER search and subsequent analysis using the NCBI Conserved Domain Database (CDD) search tool, 14 ARF genes were identified in *R. roxburghii* genome (Table 1). Each of the corresponding RrARF proteins contained the characteristic structural domains of ARF proteins

(Auxin Response Factor, Auxin Resp). These genes were then sequentially named *RrARF1* to *RrARF14* based on their chromosomal locations (Figure 1, Table 1).

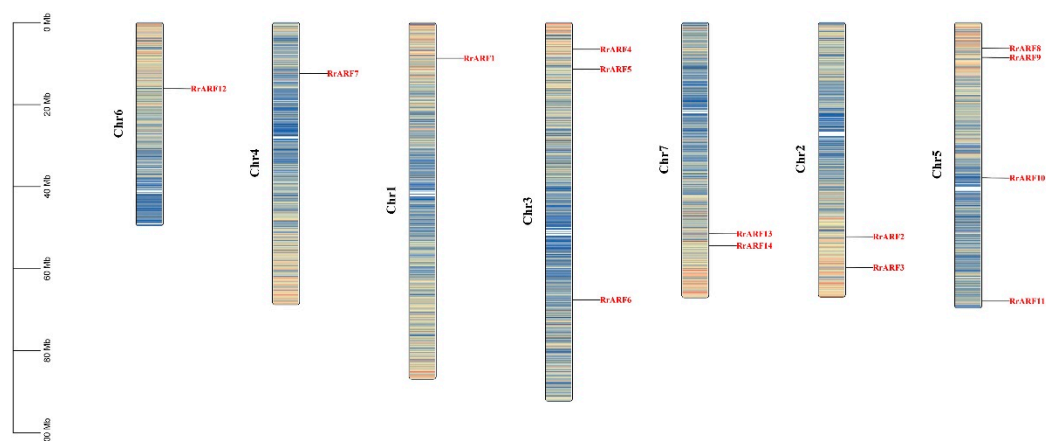


Figure 1. Distribution of ARF family genes on *Rosa roxburghii* chromosomes. The gene density was visualized with the color intensity.

Table 1. Characterization of *RrARF* genes in *Rosa roxburghii*.

Gene	Gene	CDS	Chromosome	Protein		
Name	ID	Size	Location	Length(aa)	MW(KDa)	pI
RrARF1	Rr100998	3549	Chr1	1182	132.8	6.15
RrARF2	Rr203848	2031	Chr2	676	75.5	5.83
RrARF3	Rr204729	2799	Chr2	932	102.6	5.25
RrARF4	Rr300761	2232	Chr3	743	81.5	6.31
RrARF5	Rr301243	2598	Chr3	865	95.3	6.09
RrARF6	Rr304962	2145	Chr3	714	78.7	8.56
RrARF7	Rr400968	3384	Chr4	1127	124.7	6.24
RrARF8	Rr500747	2019	Chr5	672	74.6	6.23
RrARF9	Rr500973	2607	Chr5	868	96.5	6.07
RrARF10	Rr503508	2016	Chr5	671	74.4	6.10
RrARF11	Rr505360	1713	Chr5	570	62.6	6.13

RrARF12	Rr601745	2112	Chr6	703	77.3	6.57
RrARF13	Rr703521	2061	Chr7	686	76.0	6.56
RrARF14	Rr703808	2025	Chr7	674	74.4	6.54

The amino acid (aa) lengths of the RrARF proteins vary from 570 aa (RrARF11) to 1182 aa (RrARF1), resulting in an average protein length of 791.64 aa. Accordingly, the molecular weights of the RrARF proteins ranged from 62.6 kDa (RrARF11) to 132.8 kDa (RrARF1). The isoelectric points for the 14 RrARF proteins were mainly below 7, except for RrARF6, which exhibited a higher pI of 8.56, suggesting that RrARF proteins may function in diverse cellular microenvironments.

To investigate the evolutionary relationships among RrARF homologous genes, 97 ARF proteins were obtained from five different plant species: *R. roxburghii*, *Arabidopsis thaliana*, *Malus domestica* (apple), *Vitis vinifera* (grape) and *Fragaria vesca* (wild strawberry). A phylogenetic tree was constructed using the neighbour-joining method, which is presented in Figure 2. Based on the characteristics of the protein structural domains of the ARF family in *Arabidopsis thaliana*, our phylogenetic tree classified these ARF homologues into four distinct groups. The distribution of the 14 RrARFs members across these subgroups was as follows: subgroup I (1 member), subgroup II (4 members), subgroup III (5 members) and subgroup IV (4 members).

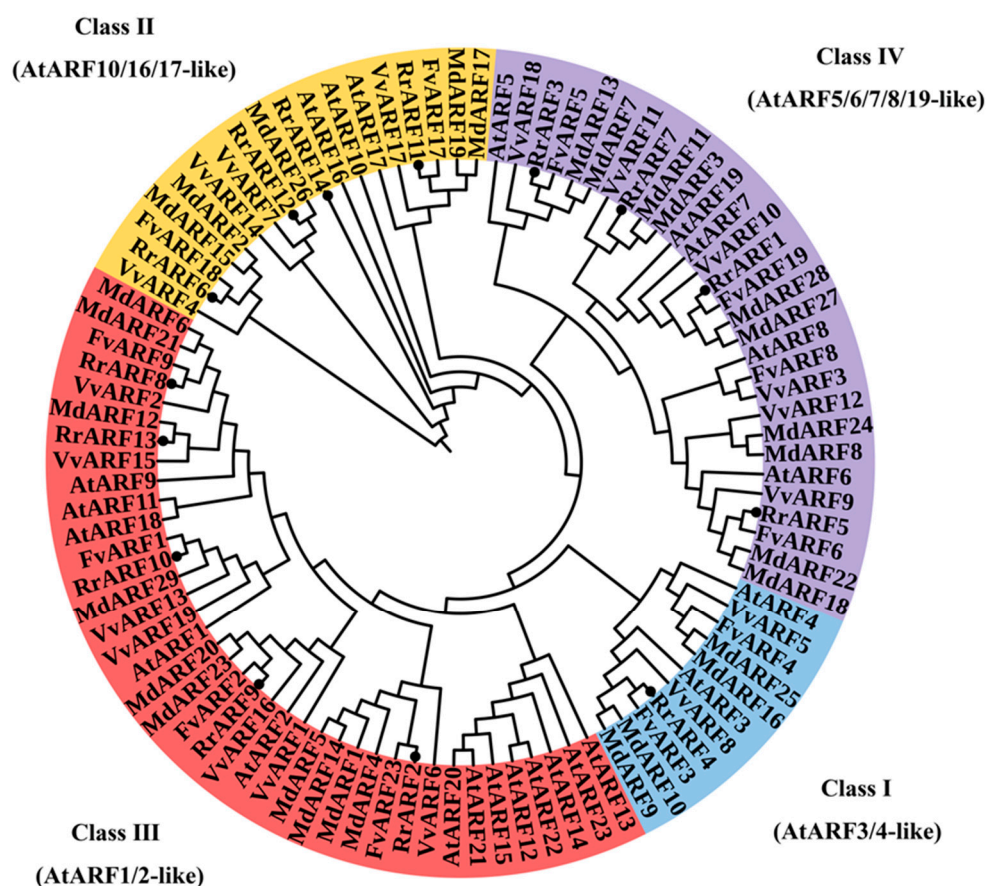


Figure 2. Phylogenetic analysis of ARF family genes in different plant species. The ARF family genes are divided into four groups, represented by different colors. At, *Arabidopsis thaliana*; Fv, *Fragaria vesca*; Md, *Malus domestica*; Rr, *Rosa roxburghii*; and Vv, *Vitis vinifera*.

3.2. Analysis of Conserved Structural Domains and Promoter Sequences of ARF Family Members in *R. roxburghii*

To gain further insight into the structural characteristics of the RrARF members and to ascertain its potential function, an in-depth examination was conducted on its exon-intron configuration, conserved structural domain, and motifs composition (Figure 3).

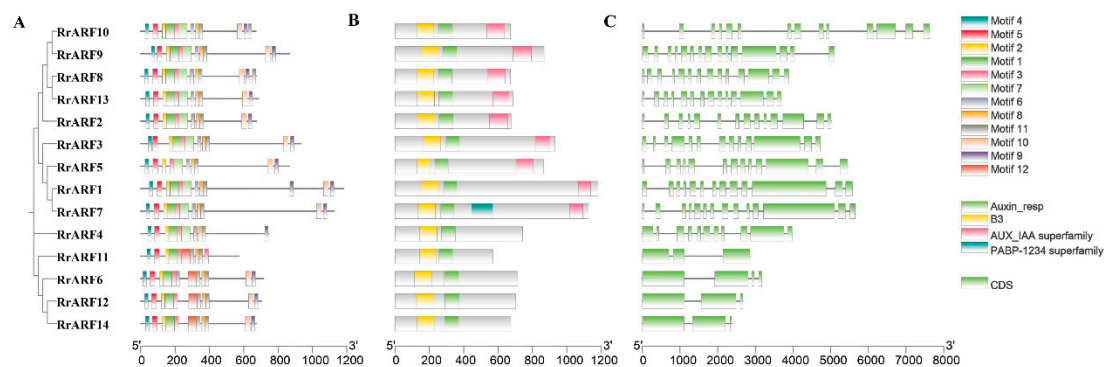


Figure 3. Conserved motifs, typical ARF structures, and gene structure analysis of *Rosa roxburghii* ARF family genes. (A) Distribution of conserved motifs in RrARF proteins. (B) Analysis of RrARFs protein domains. (C) Exon/intron structure of the *RrARF* members. Different colored boxes represent different themes. The scale bars are shown at the bottom.

To fully confirm the accuracy of the conserved structural domains in RrARF, multiple sequence comparison was performed using ClustalW. The typical ARF structure is characterized by the presence of three conserved structural domains: B3 (DBD), Auxin-responsive, and AUX/IAA superfamily (CTD) (Xing et al., 2011). The results showed that all members of the RrARF family possessed both the B3 structural domain (DBD) and the Auxin-responsive structural domain. However, the AUX/IAA domain (CTD) was lacking in *RrARF4*, *RrARF6*, *RrARF11*, *RrARF12*, and *RrARF14* (Figure 3B; Figure S1).

In addition, the RrARF protein sequence was further analyzed by the online software MEME and 12 motifs ranging from 21 to 50 amino acids in length were identified (Figure 3A; Table S1). Among these motifs, motifs 1, 2, 3, 5 belong to the B3 structural domain (DBD), motifs 6, 8, 11 belong to the auxin response structural domain, and motifs 9, 10 belong to the AUX/IAA structural domain (CTD).

Gene structure analysis revealed that the majority of the coding sequences were interrupted by introns, with members of the same phylogenetic group often sharing similar gene structures (Figure 3C). Interestingly, although RrARF9 and RrARF10 from the group of class III are similar in their protein structures, their exon-intron distributions showed differences. In addition, although RrARF1 and RrARF5 from the Class IV group have the same number of exon-introns, their CDS sizes are different. These differences may indicate that RrARFs have diverse splicing and rich transcripts.

Given the crucial role of *cis*-acting elements in regulating gene expression, we retrieved the 2-kb promoter sequences of all identified RrARFs genes and submitted them to the PlantCare database for *cis*-acting elements prediction and analysis. This analysis revealed the presence of hormone-responsive elements (e.g., MeJA, abscisic acid, salicylic acid, gibberellin, etc.), wound-responsive elements, and stress-related elements (e.g., low temperature, drought, etc.) as well as light-responsive elements (Table S2). Notably, the TGA-element and AuxRR-core motifs, which are indicative of a typical auxin response, were only identified in the promoters of *RrARF1*, *RrARF7*, *RrARF9*, and *RrARF13*, suggesting a putative role for these genes in auxin-mediated growth processes.

3.3. Expression Profiles of RrARFs in Different Tissues and Developmental Stages

To investigate the potential functions of the RrARF gene family in *R. roxburghii*, we analyzed the spatiotemporal expression profiles using published RNA-seq data from various tissues and different developmental stages of *R. roxburghii*, including root, stem, leaf, and flower, as well as fruit tissues sampled at five different developmental and ripening stages. The expression profiles of the 14 RrARF family genes in different tissues and during development fruit developmental stages can be divided into four subgroups (Figure 4). The 4 genes of RrARF that make up subgroup I (*RrARF5*, *RrARF11*, *RrARF12*, and *RrARF14*) were highly expressed in the latter stages of fruit growth and ripening, indicating a potential role in the processes of late fruit maturity and ripening. Two RrARF genes (*RrARF1* and *RrARF6*) that are mostly expressed in leaf and flower tissues make up subgroup II. This suggests that the two genes may have a role in leaf and flower development. Four genes of RrARF (*RrARF7*, *RrARF8*, *RrARF9*, and *RrARF10*) found in subgroup III displayed strong expression in roots, stems, and young fruits, suggesting that these members may be crucial to the development of both vegetative and early fruit. Subgroup IV contains four RrARF genes (*RrARF2*, *RrARF3*, *RrARF4*, and *RrARF13*) with specific expression patterns in the root and stem, indicating their vital function in root and stem growth and development.

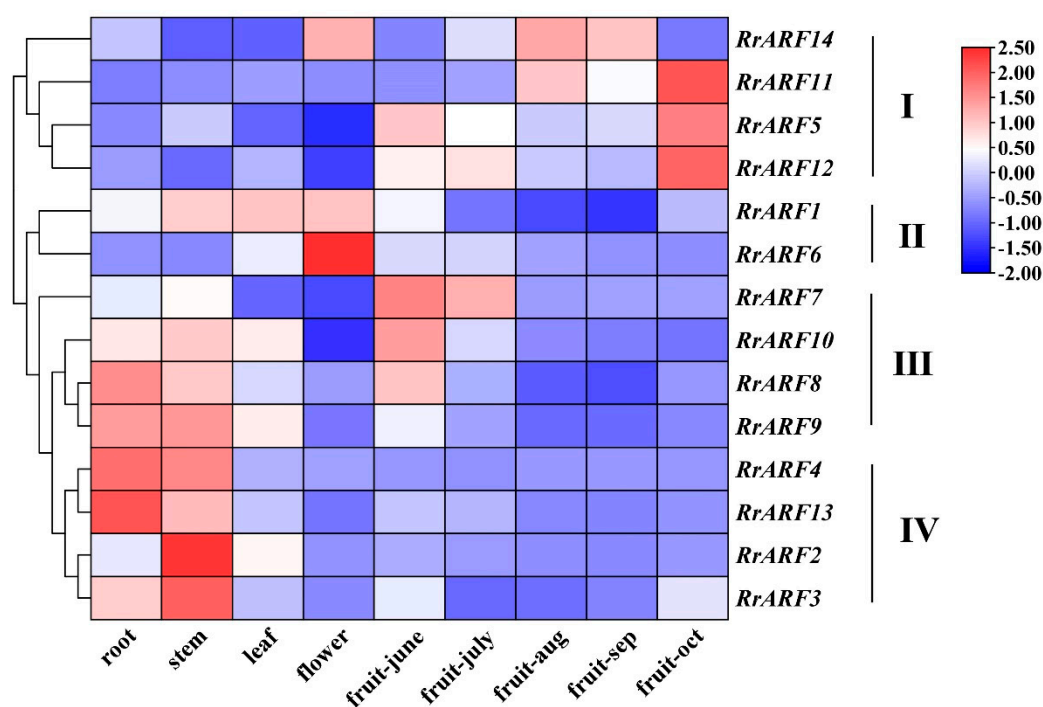


Figure 4. Expression patterns of RrARFs genes in different tissues. The heatmap showed the expression patterns of RrARFs in various tissues and fruit developmental stages, as derived from transcriptomic data.

3.4. Co-Expressed Gene Networks of RrARFs

To investigate the expression patterns of RrARF genes associated with vitamin C synthesis and metabolism, a co-expression network was constructed using weighted gene co-expression network analysis (WGCNA). The expression profiles of 14,310 genes were grouped into 14 modules, with the remaining unclustered genes placed in grey modules (Figure 5A; Figure S5). The number of genes within each module was provided in Table S3. The turquoise module was the most gene-rich, comprising 3,890 genes. In contrast, the module with the lowest number of genes was the cyan module, which contained 49 genes (Table S3).

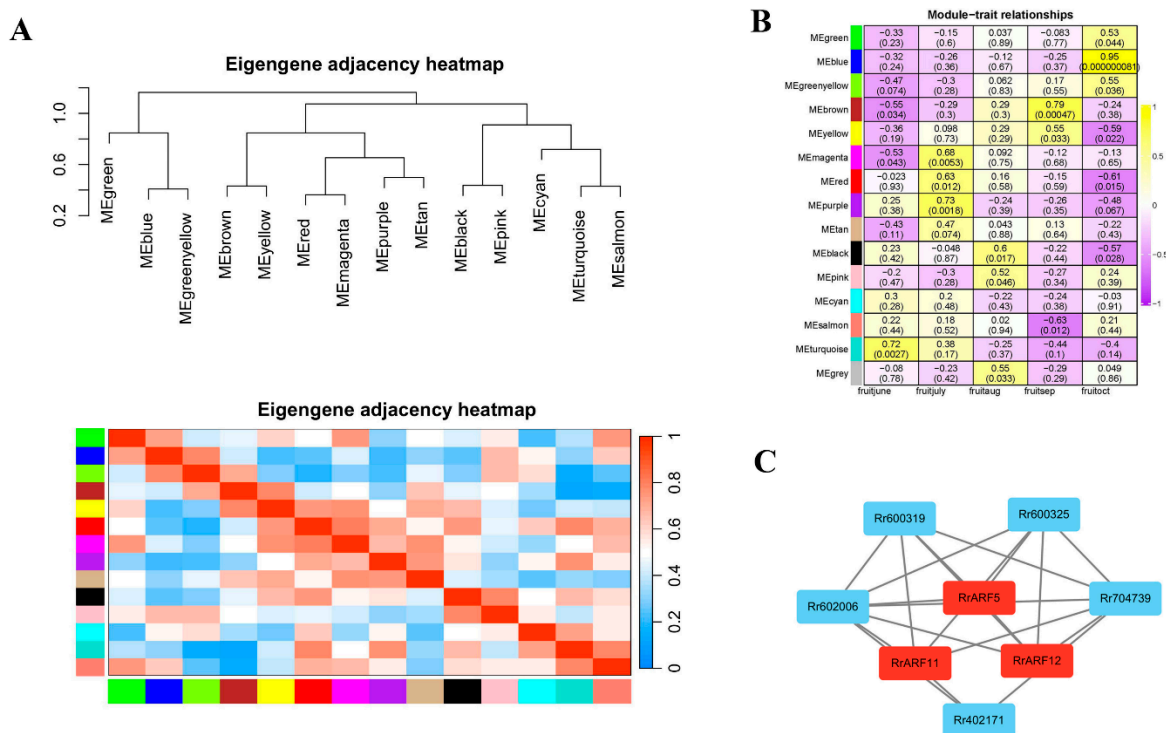


Figure 5. WGCNA analysis of *RrARFs* genes and genes related to vitamin C synthesis and metabolism in *R. roxburghii* (A) Gene clustering dendrogram in WGCNA. The upper part shows the gene clustering dendrogram and the lower part shows the assigned modules with the same colour for the same modules. (B) The relationships between the identified modules and samples at different stages of fruit development were visualised using heatmaps. The numbers displayed within the boxes represent the correlation coefficients and corresponding p-values between the identified modules and the relevant samples. (C) The co-expression network of *RrARFs* genes and genes related to vitamin C synthesis and metabolism in *R. roxburghii* within the blue module.

The correlation between the 14 modules and different periods of fruit development was analyzed (Figure 5B). Among these, the MEblue module showed the highest correlation with fruit-oct stage ($r = 0.95$, $p < 0.00000081$), which was a crucial phase for vitamin C synthesis in *R. roxburghii*. Consequently, the *RrARFs* genes within the blue module were subsequently identified as the core genes in regulation of Vc biosynthesis. A gene coexpression network was then constructed using these core genes and their coexpression genes related to vitamin C synthesis and metabolism (Figure 5C). As shown in Figure 5C, three ripening-related ARF genes (*RrARF5*, *RrARF11*, and *RrARF12*) showed a high correlation with five genes involved in VC biosynthesis, suggesting a putative regulatory relationship between these ARF and the five VC biosynthesis genes. The weight values between two genes in the gene interaction network are shown in Table S5.

3.5. *RrARF5* Involved in VC Biosynthesis by Activating *RrMDHAR1* Transcription

The analysis of the public RNA-seq data enabled the examination of the expression of genes associated with vitamin C synthesis and metabolism within the context of a co-expression network. The results demonstrated that the gene *Rr602006*, named as *RrMDHAR1*, which encodes monodehydroascorbate reductase (MDHAR) (Figure S2), exhibited high expression levels in *R. roxburghii* during fruit development (Table S4). Additionally, the expression patterns of *RrARF5* and *RrMDHAR1* exhibited a high positive correlation during fruit development and ripening (Figure 5C; Figure S4), suggesting a positive regulatory relationship between *RrARF5* and *RrMDHAR1*. Promoter analysis revealed the presence of two TGA elements and an AuxRR core ARF transcription factor regulatory element in the promoter region of *RrMDHAR1* (Figure S3). Therefore, it was hypothesized that *RrMDHAR1* may be a target of *RrARF5*. To further investigate their relationship, we performed

a dual-luciferase reporter assay in *Nicotiana benthamiana* leaf protoplasts by co-transfecting the effector construct 35S:RrARF5 with the reporter construct *pro:RrMDHAR1* (Figure 6A). The results showed that RrARF5 significantly activated the expression of *RrMDHAR1* compared to the control (Figure 6B), thereby supporting the hypothesis that RrARF5 critically involved in Vc biosynthesis by regulating *RrMDHAR1* in *R. roxburghii* fruits.

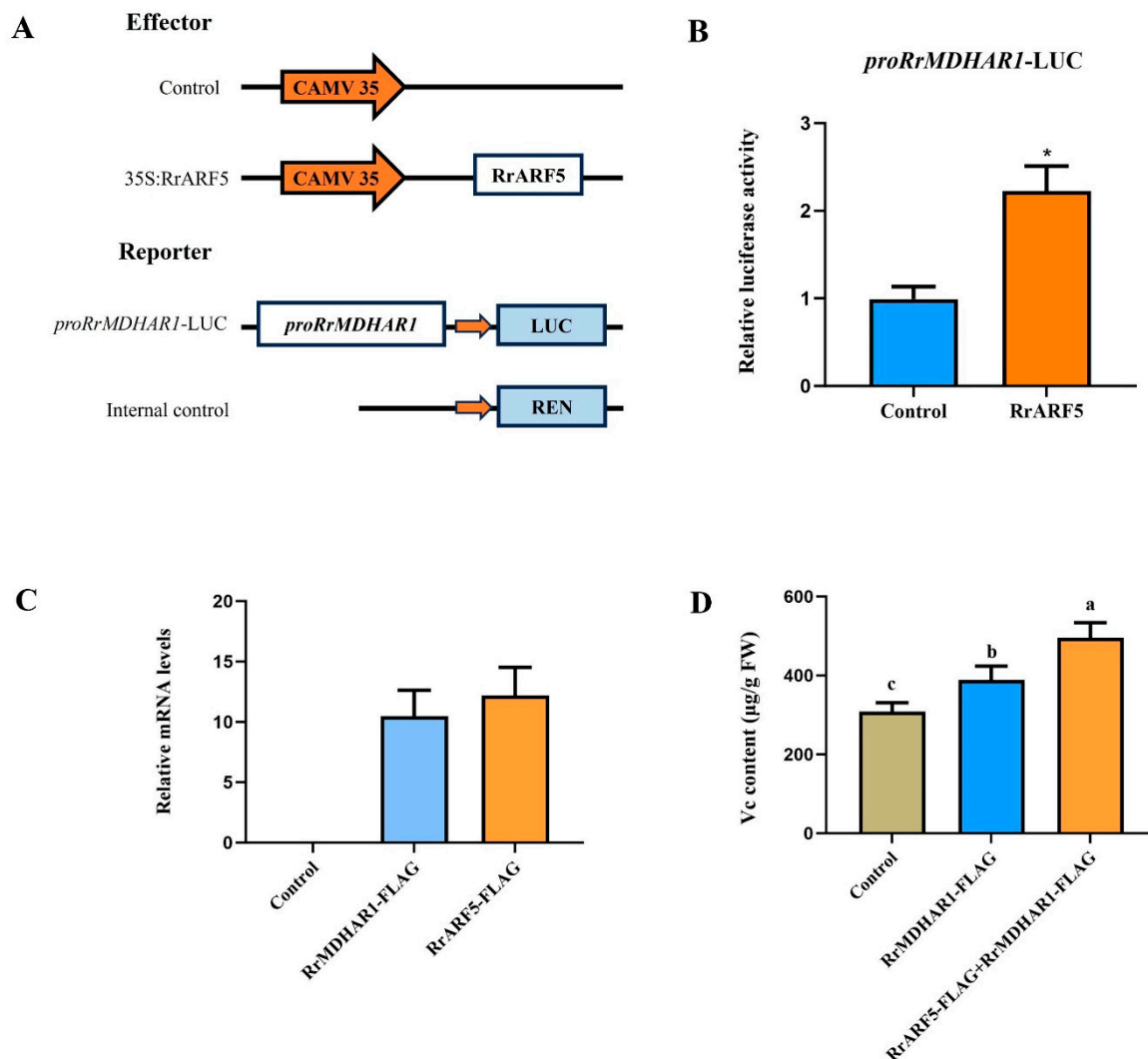


Figure 6. RrARF5 regulates Vc accumulation by activating the expression of RrMDHAR1 gene. (A) Schematic diagram of effector and reporter constructs used in the transient expression assays. (B) The activation effect of RrARF5 on the transcription of *RrMDHAR1*. (C) Transcript accumulation of the chimeric RrMDHAR1-FLAG and RrARF5-FLAG in the transient expression in fruits of tomato. Transcript levels were assessed by qRT-PCR in pericarp of injected tomato fruits. Sl-Actin gene as used as internal control. (D) The content of the vitamin C in fruits of control, injected with 35S:RrMDHAR1-FLAG construct, and injected with both 35S:RrMDHAR1-FLAG and 35S:RrARF5-FLAG tomato. Control, injected with empty vector. FW, fresh weight. Values are means \pm SD of three replicates.

To further validate the regulation of RrARF5 in Vc accumulation by regulating *RrMDHAR1*, we constructed 35S:RrMDHAR1-FLAG and 35S:RrARF5-FLAG vectors and transiently expressed these vectors in tomato fruits (Figure 6C). As shown in Figure 6D, ectopic expression of 35S:RrMDHAR1-FLAG resulted in a significant increase in Vc content compared to the control (injected with empty vector). Moreover, ectopic expression of both 35S:RrMDHAR1-FLAG and 35S:RrARF5-FLAG in tomato fruits led to significantly higher content of Vc compared to the fruits only ectopic expression

of *35S:RrMDHAR1-FLAG* (Figure 6D). These data further indicated that RrARF5 modulates Vc accumulation in fruits by regulation of *RrMDHAR1*.

4. Discussion

ARF family is a central component of the auxin signaling pathway, regulating a multitude of processes integral to plant development. The ARF transcription factor family has been extensively studied in various plant species, such as *Arabidopsis thaliana* (Guilfoyle and Hagen, 2007), tomato (*Solanum lycopersicum*) (Zouine et al., 2014), papaya (*Carica papaya* L.) (Liu et al., 2015), longan (*Dimocarpus longan* L.) (Peng et al., 2020), apple (*Malus domestica*) (Luo et al., 2014), and rice (*Oryza sativa*) (Wang et al., 2007). Nevertheless, comprehensive investigations of the entire ARF gene family in *R. roxburghii* have not yet been conducted. In the present study, we identified 14 RrARF transcription factors in *R. roxburghii*, which were unevenly distributed across seven chromosomes (Figure 1). Phylogenetic analysis of the 14 RrARF proteins showed that they can be classified into four major groups, with each group showing a close evolutionary relationship to the FvARF genes (Figure 2). This observation indicates that ARF genes in these taxa may have originated from a common ancestor, or alternatively, that they may share highly conserved functions (Paterson et al., 2004). Moreover, we identified twelve conserved motifs in RrARF family in *R. roxburghii* (Figure 3). Although the number of members differs in each phylogenetic group, the motifs within a group are markedly conserved. In addition to the conserved motifs, further research is required to determine whether additional motifs may be associated with novel functions.

The majority of ARFs are comprised of three structural domains with conserved features (Zouine et al., 2014). However, five RrARFs (RrARF4, RrARF6, RrARF11, RrARF12, and RrARF14) lack the Aux/IAA-binding domain (Supplemental Figure S1), which is consistent with the structure of the AtARF3 from the Arabidopsis subfamily that lacks the CTD structural domain. Moreover, the ARF members that lack the Aux/IAA-binding domain can also be found in tomato, rice and orchid species (Kumar et al., 2011; Wu et al., 2011; Bai et al. 2023). Since the Aux/IAA-binding domain is critical for the interaction between ARF and Aux/IAA, the lack of this domain in the five RrARFs may suggest these proteins exert their function in an auxin-dependent manner. It is noteworthy that RrARF7 contains a PABP-1234 superfamily structural domain in addition to the three typical ARF domains (Figure 3). However, the function of this conserved domain and also the function of RrARF7 require further studies. Additionally, the presence of cis-acting elements associated with plant development, abiotic and biotic stresses, light responses, and phytohormone responses in the promoter regions of the RrARF family genes and their distinct expression patterns in different tissues and fruit developmental stages support their diverse functions in plant growth and fruit development. However, the functional significances of these RrARF genes require further investigation.

High vitamin C content is one of the most important quality traits of fruits of *R. roxburghii*, understanding the regulatory mechanism of vitamin C biosynthesis and metabolism is vital for the improvement of fruit quality of *R. roxburghii*. In the present study, thorough coexpression analysis, transactivation assays, and transient overexpression assays, we revealed that RrARF5 acts as positive regulator of Vc biosynthesis by activation of the expression of *RrMDHAR1*, a structure gene which was reported to be correlated with ascorbic acid content (Stevens et al., 2008). These findings contribute to a better understanding of the function of ARF gene family in fruit quality regulation, while also providing a theoretical foundation for further research into the molecular mechanisms underlying ascorbic acid accumulation in *R. roxburghii*.

It is worth pointing out that although we identified the ARF gene family in *R. roxburghii* and revealed the role of one ARF member in Vc accumulation, the function of other ARF genes in different plant growth and developmental process remain largely unknown. In the future, we will focus on the development of the plant genetic transformation technology of *R. roxburghii* and uncovering the biological function of ARF family members.

Acknowledgments: The authors gratefully acknowledge the financial support by the Regional First-class Discipline of Ecology in Guizhou Province (XKTJ[2020]22), the Technology Projects of Guizhou Province, China (No.ZK[2021]090), the Outstanding Young Scientific and Technological Talents of Sichuan Province (2022JDJQ0057).

References

- An Huaming, Fan Weiguo, Chen Ligeng, Asghar Saleem & Liu Qinglin. 2007. Molecular characterisation and expression of L-galactono-1,4-lactone dehydrogenase and L-ascorbic acid accumulation during fruit development in *Rosa roxburghii*. *The Journal of Horticultural Science and Biotechnology*. 82, 627-635.
- Bai Yiwei, Ma Yanjun, Chang Yanting, Zhang Wenbo, Deng Yayun, Zhang Na, Zhang Xue, Fan Keke, Hu Xiaomeng, Wang Shuhua, Jiang Zehui, Hu Tao. 2023. Identification and transcriptome data analysis of ARF family genes in five Orchidaceae species. *Plant molecular biology*. 112, 85-98.
- Chen Chengjie, Chen Hao, Zhang Yi, Thomas Hannah R, Frank Margaret H, He Yehua, Xia Rui. 2020. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant*. 13, 1194-1202.
- Chen Shifu, Zhou Yanqing, Chen Yaru, Gu Jia. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34, i884-i890.
- Dolf Weijers, Eva Benkova, Katja EJäger, Alexandra Schlereth, Thorsten Hamann, Marika Kientz, Jill C Wilmoth, Jason W Reed, and Gerd Jürgens. 2014. Developmental specificity of auxin response by pairs of arf and AUX/IAA transcriptional regulators. *The EMBO Journal*. 24, 1874-1885.
- Du Hao, Wu Nai, Fu Jing, Wang Shiping, Li Xianghua, Xiao Jinghua, Xiong Lizhong. 2012. A GH3 family member, OsGH3-2, modulates auxin and abscisic acid levels and differentially affects drought and cold tolerance in rice. *Journal of Experimental Botany*. 63, 6467-6480.
- El-Sharkawy Islam, Sherif Sherif M, Jones Brian, Mila Isabelle, Kumar Prakash P, Bouzayen Mondher, Jayasankar Subramanian. 2014. TIR1-like auxin-receptors are involved in the regulation of plum fruit development. *Journal of Experimental Botany*. 65, 5205-5215.
- Gambhir Priya, Raghuvanshi Utkarsh, Kumar Rahul, Sharma Arun Kumar. 2024. Transcriptional regulation of tomato fruit ripening. *Physiology and Molecular Biology of Plants*. 30, 289-303.
- Guilfoyle Tom J, Hagen Gretchen. 2007. Auxin response factors. *Current Opinion in Plant Biology*. 10, 453-460.
- Guilfoyle Tom J, Ulmasov T, Hagen Gretchen. 1998. The ARF family of transcription factors and their role in plant hormone-responsive transcription. *Cellular and Molecular Life Sciences*. 54, 619-627.
- Juan Carlos Serrani, Mariano Fos, Alejandro Atarés, José L. García-Martínez. 2007. Effect of gibberellin and auxin on parthenocarpic fruit growth induction in the cv Micro-Tom of tomato. *Journal of Plant Growth Regulation*. 26, 211-221.
- Kumar Rahul, Tyagi Akhilesh K, Sharma Arun K. 2011. Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. *Molecular Genetics and Genomics*. 285, 245-260.
- Kumar Sudhir, Stecher Glen, Tamura Koichiro. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*. 33, 1870-1874.
- Li Sibe, OuYang Weizhi, Hou Xiaojin, Xie LiangLiang, Hu Chungeng, Zhang Jinzhi. 2015. Genome-wide identification, isolation and expression analysis of auxin response factor (ARF) gene family in sweet orange (*Citrus sinensis*). *Frontiers in Plant Science*. 6, 119.
- Liao Guanglian, Chen Lu, He Yanqun, Li Xishi, Lv Zhengxin, Yi Shuyao, Zhong Min, Huang Chunhui, Jia Dongfeng, Qu Xueyan, Xu Xiaobiao. 2021. Three metabolic pathways are responsible for the accumulation and maintenance of high AsA content in kiwifruit (*Actinidia eriantha*). *BMC Genomics*. 22, 13.
- Liao Guanglian, Xu Qiang, Allan Andrew C, Xu Xiaobiao. 2023. L-Ascorbic acid metabolism and regulation in fruit crops. *Plant physiology*. 192, 1684-1695.
- Lim Pyung Ok, Lee In Chul, Kim Junyoung, Kim Hyo Jung, Ryu Jong Sang, Woo Hye Ryun, Nam Hong Gil. 2010. Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity. *Journal of Experimental Botany*. 61, 1419-1430.

- Liu Kaidong, Yuan Changchun, Li Haili, Lin Wanhuang, Yang Yanjun, Shen Chenjia, Zheng Xiaolin. Genome-wide identification and characterization of auxin response factor (ARF) family genes related to flower and fruit development in papaya (*Carica papaya* L.). 2015. *BMC Genomics*. 16, 901.
- Liu Menghua, Zhang Qi, Zhang Yuanhe, Lu Xianyuan, Fu Weiming & He Jingyu. 2016. Chemical analysis of dietary constituents in *Rosa roxburghii* and *Rosa sterilis* fruits. *Molecules*. 21, 1204.
- Liu Xigang, Dinh Thanh Theresa, Li Dongming, Shi Bihai, Li Yongpeng, Cao Xiuwei, Guo Lin, Pan Yanyun, Jiao Yuling, Chen Xuemei. 2014. AUXIN RESPONSE FACTOR 3 integrates the functions of AGAMOUS and APETALA2 in floral meristem determinacy. *The Plant journal: for cell and molecular biology*. 80, 629-641.
- Lu Min, An Huaming, Li Liangliang. (2016). Genome survey sequencing for the characterization of the genetic background of *Rosa roxburghii* Tratt and leaf ascorbate metabolism genes. *PLoS One*, 11(2), e0147530.
- Luo Xiaocui, Sun Meihong, Xu Ruirui, Shu Huairui, Wang Jiawei, Zhang Shizhong. 2014. Genomewide identification and expression analysis of the ARF gene family in apple. *Journal of Genetics*. 93, 785-797.
- Otasek David, Morris John H, Bouças Jorge, Pico Alexander R, Demchak Barry. 2019. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biology*. 20, 185.
- Park Jung-Eun, Park Ju-Young, Kim Youn-Sung, Staswick Paul E, Jeon Jin, Yun Ju, Kim Sun-Young, Kim Jungmook, Lee Yong-Hwan, Park Chung-Mo. 2007. Gh3-mediated auxin homeostasis links growth regulation with stress adaptation response in Arabidopsis. *The Journal of Biological Chemistry*. 282, 10036-10046.
- Paterson Andrew H, Bowers John E, Chapman Brad A. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proceedings of the National Academy of Sciences*. 101, 9903-9908.
- Pei Yangang, Xue Qihan, Shu Peng, Xu Weijie, Du Xiaofei, Wu Mengbo, Liu Kaidong, Pirrello Julien, Bouzayen Mondher, Hong Yiguo, Liu Mingchun. 2024. Bifunctional transcription factors SIERF.H5 and H7 activate cell wall and repress gibberellin biosynthesis genes in tomato via a conserved motif. *Developmental Cell*. 59, 1345-1359.
- Peng Yuan, Fang Ting, Zhang Yiyong, Zhang Mengyuan, Zeng Lihui. 2020. Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in Longan (*Dimocarpus longan* L.). *Plants*. 9, 221.
- Shen Chenjia, Yue Runqing, Sun Tao, Zhang Lei, Xu Luqin, Tie Shuanggui, Wang Huizhong, Yang Yanjun. 2015. Genome-wide identification and expression analysis of auxin response factor gene family in *Medicago truncatula*. *Frontiers in Plant Science*. 6, 1-13.
- Smirnoff Nicholas. 2018. Ascorbic acid metabolism and functions: a comparison of plants and mammals. *Free Radical Biology and Medicine*. 122, 116-129.
- Stevens Rebecca, Page Dan, Gouble Barbara, Garchery Cecile, Zamir Dani, Causse Mathilde. 2008. Tomato fruit ascorbic acid content is linked with monodehydroascorbate reductase activity and tolerance to chilling stress. *Plant, Cell & Environment*. 31, 1086-1096.
- Tiwari Shiv B, Wang Xiaojun, Gretchen Hagen, Guilfoyle Tom J. 2001. AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *The Plant Cell*. 13, 2809-2822.
- Vriezen Wim H, Feron Richard, Maretto Fabio, Keijman Jasper, Mariani Celestina. 2008. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. *The New Phytologist*. 17, 60-76.
- Wan Sibao, Li Weili, Zhu Yueying, Liu Zhanmin, Huang Weidong, Zhan Jicheng. 2014. Genome-wide identification, characterization and expression analysis of the auxin response factor gene family in *Vitis vinifera*. *Plant Cell Reports*. 33, 1365-1375.
- Wang Dekai, Pei Kemei, Fu Yaping, Sun Zongxiu, Li Sujuan, Liu Heqin, Tang Kan, Han Bin, Tao Yuezhi. 2007. Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). *Gene*. 394, 13-24.
- Wang Shaoxi, Shi Fengyan, Dong Xiangxiang, Li Yuxiang, Zhang Zhihong, Li He. 2018. Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in strawberry (*Fragaria vesca*). *Journal of Integrative Agriculture*. 18, 1587-1603.

- Wilmoth Jill C, Wang Shucai, Tiwari Shiv B, Joshi Atul D, Hagen Gretchen, Guilfoyle Thomas J, Alonso Jose M, Ecker Joseph R, Reed Jason W. 2010. NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. *The Plant Journal*. 43, 118-130.
- Wu Jian, Wang Feiyan, Cheng Lin, Kong Fuling, Peng Zhen, Liu Songyu, Yu Xiaolin, Lu Gang. 2011. Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum*. *Plant cell reports*. 30, 2059-2073.
- Xing Hongyan, Ramesh N Pudake, Guo Ganggang, Xing Guofang, Hu Zhaorong, Zhang Yirong, Sun Qixin, Ni Zhongfu. 2011. Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. *BMC Genomics*. 12, 178.
- Yang Hao, Hu Jiwei, Huang Xianfei, Zhou Chao, Li Lingyun, Fan Mingyi. 2014. Risk Assessment of Heavy Metals Pollution for *Rosa sterilis* and Soil from Planting Bases Located in Karst Areas of Guizhou Province. *Applied Mechanics and Materials*. 700, 475-481.
- Ye Youju, Wang Jianwen, Wang Wei, Xu LiAn. 2020. ARF family identification in *Tamarix chinensis* reveals the salt responsive expression of TcARF6 targeted by miR167. *PeerJ*. 8, e8829.
- Yi ShanNa, Mao Jianxin, Zhang Xinyu, Li Xiaoming, Zhang Zhihong, LiHe. 2022. FveARF2 negatively regulates fruit ripening and quality in strawberry. *Frontiers in Plant Science*. 13, 1023739-1023739.
- Yoko Okushima, Hidehiro Fukaki, Makoto Onoda, Athanasios Theologis, Masao Tasaka. 2007. ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in *Arabidopsis*. *The Plant Cell*. 19, 118-130.
- Yoko Okushima, Overvoorde Paul J, Arima Kazunari, Alonso Jose M, Chan April, Chang Charlie, Ecker Joseph R, Hughes Beth, Lui Amy, Nguyen Diana, Onodera Courtney, Quach Hong, Smith Alison, Yu Guixia, Theologis Athanasios. 2005. Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19. *The Plant Cell*. 17, 444-463.
- Yuan Yujin, Mei Lihua, Wu Mengbo, Wei Wei, Shan Wei, Gong Zehao, Zhang Qian, Yang Fengqing, Yan Fang, Zhang Qiang, Luo Yingqing, Xu Xin, Zhang Wenfa, Miao Mingjun, Lu Wangjin, Li Zhengguo, Deng Wei. 2018. SlARF10, an auxin response factor, is involved in chlorophyll and sugar accumulation during tomato fruit development. *Journal of Experimental Botany*. 69, 5507-5518.
- Yuan Yujin, Xu Xin, Gong Zehao, Tang Yuwei, Wu Mengbo, Yan Fang, Zhang Xiaolan, Zhang Qian, Yang Fengqing, Hu Xiaowei, Yang Qichen, Luo Yingqing, Mei Lihua, Zhang Wenfa, Jiang Caizhong, Lu Wangjin, Li Zhengguo, Deng Wei. 2019. Auxin response factor 6A regulates photosynthesis, sugar accumulation, and fruit development in tomato. *Horticulture Research*. 6, 1-16.
- Yue Pengtao, Lu Qian, Liu Zhi, Lyu Tianxing, Li Xinyue, Bu Haidong, Liu Weiting, Xu Yaxiu, Yuan Hui, Wang Aide. 2020. Auxin-activated MdARF5 induces the expression of ethylene biosynthetic genes to initiate apple fruit ripening. *The New Phytologist*. 226, 1781-1795.
- Zhang Huanxin, Cao Ning, Dong Chunjuan, Shang Qingmao. 2017. Genome-wide identification and expression of ARF gene family during adventitious root development in hot pepper (*Capsicum annuum*). *Horticultural Plant Journal*. 3, 151-164.
- Zhang Tao, Li Weijin, Xie Ruxiu, Xu Ling, Zhou Yan, Li Haili, Yuan Changchun, Zheng Xiaolin, Xiao Langtao, Liu Kaidong. 2020. CpARF2 and CpEIL1 interact to mediate auxin-ethylene interaction and regulate fruit ripening in papaya. *The Plant Journal*. 103, 1318-1337.
- Zhou Zheng, Schenke Dirk, Miao Ying, Cai Daguang. 2017. Investigation of the crosstalk between the flg22 and the UV-B-induced flavonol pathway in *Arabidopsis thaliana* seedlings. *Plant, Cell & Environment*. 40, 453-458.
- Zong Dan, Liu Huan, Gan Peihua, Ma Shaojie, Liang Hongping, Yu Jinde, Li Peilin, Jiang Tao, Sahu Sunil Kumar, Yang Qingqing, Zhang Deguo, Li Laigeng, Qiu Xu, Shao Wenwen, Yang Jinlong, Li Yonghe, Guang Xuanmin, He Chengzhong. 2023. Chromosomal-scale genomes of two *Rosa* species provide insights into genome evolution and ascorbate accumulation. *The Plant Journal*. 117, 1264-1280.
- Zouine Mohamed, Fu Yongyao, Chateigner-Boutin Anne-Laure, Mila Isabelle, Frasse Pierre, Wang Hua, Audran Corinne, Roustan Jean-Paul, Bouzayen Mondher. 2014. Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. *Plos One*. 9, e84203.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.