horticulturae

MDPI

Article

Genome-Wide Identification and Bioinformatics Analysis of Auxin Response Factor Genes in Highbush Blueberry

Yu Zong [1,2,†], Lili Gu [1,†], Zhuli Shen [1], Haiting Kang [1], Yongqiang Li [1,2], Fanglei Liao [1,2], Lishan Xu [1,2,*] and Weidong Guo [1,2,*]

[1] College of Chemistry and Life Sciences, Zhejiang Normal University, Jinhua 321004, China; yzong@zjnu.cn (Y.Z.); liligu7@163.com (L.G.); szl_77@163.com (Z.S.); haitingkang@163.com (H.K.); lyq@zjnu.cn (Y.L.); fangleiliao@zjnu.cn (F.L.)

[2] Zhejiang Provincial Key Laboratory of Plant Biotechnology, Jinhua 321004, China

* Correspondence: xls@zjnu.cn (L.X.); gwd@zjnu.cn (W.G.)

† These authors contributed equally to this work.

Abstract: Auxin response factors (ARFs) are a transcription factor family that regulates the expression of auxin phase-responsive genes. Here, we performed a genome-wide investigation of the tetraploid blueberry (*Vaccinium corymbosum* cv. 'Draper') genome sequence. Physical and chemical properties, phylogenetic evolution, gene structure, conservative motifs, chromosome location, and cis-acting elements of blueberry *ARF* genes were comprehensively evaluated. A total of 70 blueberry *ARF* genes (*VcARF*) were found in its genome, which could be divided into six subfamilies. *VcARF* genes were unevenly distributed on 40 chromosomes and were observed to encode protein sequences ranging in length from 162 to 1117 amino acids. Their exon numbers range from 2 to 22. *VcARF* promoter regions contain multiple functional domains associated with light signaling, aerobic metabolism, plant hormones, stress, and cell cycle regulation. More family members of *VcARF* genes were discovered in blueberry than in previously studied plants, likely because of the occurrence of whole-genome duplication and/or tandem duplication. *VcARF* expression patterns were analyzed at different stages of fruit development, and *VcARF3*, *VcARF4*, *VcARF14*, *VcARF37*, and *VcARF52* were observed to play important roles. *VcARF3* and *VcARF4* appeared to function as repressors, while *VcARF14* acted as an essential factor in fruit firmness differences between firm and soft flesh cultivars.

Keywords: genome wide; blueberry; auxin response factor; fruit firmness

Citation: Zong, Y.; Gu, L.; Shen, Z.; Kang, H.; Li, Y.; Liao, F.; Xu, L.; Guo, W. Genome-Wide Identification and Bioinformatics Analysis of Auxin Response Factor Genes in Highbush Blueberry. *Horticulturae* **2021**, *7*, 403. https://doi.org/10.3390/horticulturae7100403

Academic Editor: Dong Zhang

Received: 3 September 2021
Accepted: 7 October 2021
Published: 15 October 2021

1. Introduction

The ubiquitous involvement of auxin in almost all aspects of plant development and in plant responses to the environment, abiotic stress, and growth tropisms underlines its importance [1]. The initiation and regulation of these processes are mostly accomplished by the expression and regulation of auxin-related genes [2], which have received considerable research attention. These genes include the Skp1–Cullin–F-Box protein complex, containing the transport inhibitor response 1 protein (SCF^{TIR1}) auxin receptor and its related auxin signaling F-box protein receptor family members, and two families of partially redundant proteins: the auxin response factors (ARFs) and their cognate auxin/indole-3-acetic acid (Aux/IAA) repressors [3–5]. ARFs activate or inhibit the expression of auxin response genes by binding to auxin-responsive elements (AuxREs) in the promoter region and recruit Aux/IAA to perform their functions.

As transcription factors, ARFs possess a modular structure and consist of three major domains: an amino-terminal DNA-binding domain (DBD), a middle domain, and a carboxy-terminal PB1 (Phox and Bem1) domain contained within a region previously known as domain III/IV [2,6,7]. The B3-type DBD is flanked by dimerization domains (DDs), at least in ARF1 and ARF5, and also contains a Tudor-like ancillary domain. The DBD recognizes and binds to the TGTCTC element (AuxRE) in the promoter region of

auxin-responsive genes [8]. The DD mediates the dimerization of ARF1 and ARF5, which is essential for ARF5 cooperative binding to target DNA [9]. The function of the Tudor-like ancillary domain is unknown, but it might be involved in an interaction with the DD. The middle domain of ARFs either activates or represses transcription [10,11]; repressor domains are enriched in proline, serine, and threonine, and activator domains are enriched in glutamine. This basis has been used to identify and classify ARFs from many plant species through genome-wide analyses.

Most information concerning ARF function derives from both forward genetic analysis and reverse genetic phenotypic analysis in Arabidopsis, which showed that ARF functional redundancy is universal [12,13]. The Arabidopsis genome contains 22 full-length ARF genes and one incomplete ARF23 [14]. This is thought to be a pseudogene because it contains a stop codon in the DBD, leading to a partially encoded protein [15]. AtARF1 mutations do not themselves confer phenotypes but can enhance the phenotypic traits of *AtARF2* mutations [16]. *AtARF2* mutations result in late flowering leaf senescence, and an increased seed number. *ARF3*/*ARF4* plays an important role in the development of reproductive and vegetative tissues, and *AtARF3* also has a function in floral meristem determination, floral organ pattern formation, and pistil development [17,18]. *AtARF6* and *AtARF8* encode a pair of functionally redundant transcription factors that regulate the pistil and stamen development of immature flowers. The stamen filaments of *arf6* and *arf8* double mutants are shorter than those of the wildtype and show delayed anther dehiscence, leading to female sterility [19]. Although almost no phenotypic change was detected between *AtARF7* or *AtARF19* single-mutant and wildtype plants, *AtARF7* and *AtARF19* double mutants showed inhibited adventitious root formation, a reduced number of lateral roots, and no leaf cell enlargement, indicating that the two genes are complementary in function [20].

ARFs were also reported to be involved in the fruit ripening process, with most fruit development studies being carried out in tomato. Of the 21 *ARF* tomato genes [21], *SlARF2* encodes the main regulatory factor that controls fruit ripening through ethylene signals and biosynthesis [22], and *SlARF4* functions in the accumulation of fruit chlorophyll and affects early fruit development by regulating glucose metabolism [23]. *SlARF9* regulates cell division at an early stage of fruit development [24], while *SlARF10* controls the formation of chlorophyll, starch synthesis, and sugar accumulation in fruits [25]. However, the response of most *ARF* genes to environmental and hormonal signals has been poorly studied, and there are still unidentified key factors in the *ARF* expression regulatory network. The temporal and spatial expression of *ARF* genes is also unclear.

Following advances in sequencing technology, plant genome sequencing is now more affordable. Thus, several *ARF* plant gene families have been identified at the whole-genome level, including, but not limited, to rice (*Oryza sativa*) [14], maize (*Zea mays*) [26], tomato (*Solanum lycopersicum*) [21], apple (*Malus domestica*) [27], sweet orange (*Citrus sinensis*) [28], grape (*Vitis vinifera*) [29], longan (*Dimocarpus longan*) [30], and strawberry (*Fragaria vesca*) [31]. However, the genome-wide identification of *ARF* genes has not been performed in blueberry (*Vaccinium* spp.). In comparison with fruit trees, the genome resources of blueberry are limited. The first whole-genome sequence of the blueberry cultivar 'W8520' was released by Bian et al. in 2014 [32], but its assembly quality and usability were limited. Although Gupta et al. [33] released another assembly based on the genome sequencing of the blueberry cultivar 'O'Neal' in 2014, 27.35% of genome sequences contained gaps and mistakes. More recently, Colle and colleagues published a high-quality haplotype-phased blueberry genome [34] that was assembled from the highbush blueberry cultivar 'Draper'. This genome has enabled several molecular and bioinformatics studies of blueberry to be conducted and allows the opportunity to identify transcription factor genes throughout the genome.

In the current study, genome-wide *VcARF*s were identified from the highbush blueberry 'Draper' genome. By analyzing the number of family members, gene structure, and amino acid sequences, we clarified the structural and evolutionary characteristics of the blueberry *ARF* gene family. The expression patterns of different *VcARF* genes were

also estimated from transcriptomic data, and pivotal *VcARF*s that strongly correlate with blueberry fruit development were screened. Moreover, the transcript abundance of key *VcARF* genes was validated in firm flesh cultivar 'Star' and soft flesh cultivar 'O'Neal' at a range of developmental stages. Our findings provide a reference for the genome-wide identification of transcription factor genes and a study on the regulation of *VcARF* genes in blueberry fruit firmness.

2. Materials and Methods

2.1. Plant Materials

The fruits of two southern highbush blueberry cultivars 'Star' (firm flesh) and 'O'Neal' (soft flesh) were collected at four development stages following the sampling strategy of Zifkin et al. [35]. For each cultivar, 50 healthy fruits for every stage were sampled from different fruit setting positions of three plants. A total of 30 fruits of similar sizes at each fruit stage were screened [35]. The fruits were immediately frozen in liquid nitrogen after collection and stored at −80 °C until required.

2.2. RNA Isolation and Reverse Transcription

Total RNA was extracted using a modified CTAB method as described by Chang et al. [36]. The quality and concentration of blueberry RNA were determined using agarose gel electrophoresis and spectrophotometric analyses, respectively. cDNA was synthesized from 1 μg RNA using the PrimeScript™ RT reagent Kit with gDNA Eraser (Takara Biotechnology Co., Ltd., Dalian, China) following the manufacturer's instructions. Synthesized 1st-strand cDNAs were diluted 10-fold and stored at −80 °C for further use.

2.3. Identification of the ARF Gene Family in Blueberry

A local BLAST database for genome and amino acid sequences of the highbush blueberry 'Draper' was constructed using BLAST-2.9.0+ [37]. *A. thaliana* ARF peptide sequences were downloaded from the plant transcription factor database at Peking University (Beijing, China) (http://planttfdb.gao-lab.org/, accessed on 1 April 2020). Potential blueberry ARF transcription factors were retrieved from deduced amino acid sequences using BLASTp in the BLAST toolkit [37]. Query hits with an e-value $\leq$ 1e − 20 were kept for further analyses. Profile hidden Markov models of the ARF conserved domain (PF06507) were downloaded from the Pfam website (http://pfam.xfam.org/, accessed on 1 April 2020) [38]. HMMER3.3 [39] was used to search for homologous ARF proteins in blueberry based on the ARF conserved domain. Other parameters of HMMER were set as default. Candidate ARF genes obtained from BLASTp and HMMER results were merged, and redundant sequences were removed. Non-redundant results were submitted to the NCBI conserved domain database (https://www.ncbi.nlm.nih.gov/cdd, accessed on 3 April 2020) for domain analysis, and the sequences with ARF annotation results were retained for sequential analyses. Prediction and analysis of physicochemical properties of all ARF protein sequences in blueberry were made using the Expasy website (https://web.expasy.org/protparam/, accessed on 4 April 2020).

2.4. Analyses of Amino Acid Sequences and VcARF Gene Structure

MAFFT v7.453 [40] and IQ-TREE1.6.12 [41] were used to align protein sequences of *Arabidopsis* and blueberry and to construct a phylogeny tree based on the maximum likelihood method, respectively. Ultrafast bootstrapping was carried out 1000 times to test the accuracy of the phylogeny tree. Other parameters of IQ-TREE were set as default. Identified *VcARF* sequences were obtained from the 'Draper' genome using a customized Perl script and aligned with MAFFT v7.453. A phylogenetic tree was built with IQ-TREE1.6.12. Gene structural analysis was performed using the gene structure display server online tool (http://gsds.gao-lab.org/index.php, accessed on 6 April 2020). The structure of coding regions, introns, and non-coding regions of all selected *ARF* genes in the blueberry genome were plotted, and the phylogenetic relationship of *VcARF* genes was drawn.

2.5. Chromosome Location of VcARFs and Conservative Motif Analysis of VcARF Proteins

Genomic location information of *VcARF* genes was extracted from annotation general feature format files using Shell script. *VcARF* genes were plotted onto the chromosomes with the chromPlot R package [42]. MEME (https://meme-suite.org/meme/doc/meme.html, accessed on 5 April 2020) was used to analyze the conserved motifs of all VcARF protein sequences. The number of motifs was set as 15, and other parameters were kept as default.

2.6. Identification of Cis Elements on VcARF Promoters

A Perl script was used to retrieve 2 kbp sequences upstream of the transcription start site of *VcARF* genes. Cis elements in the promoter regions were predicted by PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 6 April 2020). The categories and number of those cis elements were calculated and drawn using Microsoft Excel 2013 (Microsoft, Redmond, WA, USA).

2.7. Expression Pattern of VcARF Genes during Fruit Development

The gene expression profile of different plant organs was obtained from a study by Colle et al. (2019) [34]. A heat map was generated by ggplot2 [43] based on the fragments per kilobase of exon per million mapped fragments (FPKM) values. The expression level of *VcARF* transcripts in fruits of 'Star' and 'O'Neal' cultivars at four development stages was determined by quantitative PCR on an ABI StepOne Plus™ Real-Time fluorescence qPCR system (Applied Biosystems Co., Ltd., Beijing, China). Specific primers (Table 1) were designed using Primer Premier 5 software (PREMIER Biosoft International, Palo Alto, CA, USA). The experiments were conducted on three biological replicates, and the results were normalized by using the *VcGAPDH* gene. Gene expression data were analyzed with a relative quantification method ($2^{-\Delta\Delta Ct}$) [44]. Statistical analyses were carried out using SPSS software v.18.0 (IBM, Armonk, NY, USA).

Table 1. *VcARF* genes and their primers used for quantification PCR in the current study.

Gene Name	Forward Primer Sequence (5′-3′)	Reverse Primer Sequence (5′-3′)
VcARF3	GTGCTGGACCCCTTGTTACT	GGCAGCTGTTGATCCAATCC
VcARF4	GCTGGACCCCTTGTTACTCTT	ATTGAACGAAGGCAGCTGTTG
VcARF14	TATGGCGGGACCGTAACAAC	ACTGAGTCACCGGTAAAGAGC
VcARF37	CCGTAACAACCACCCCGATT	GGGGAAGTAGACAACGTGGG
VcARF52	GCACCAGATCACCCGATTCC	TACCAAGGGCAATCCCCTGC
VcGAPDH	TGAGAAAGAATACAAGCCAGAT	CAGGCAACACCTTACCAA

3. Results

3.1. Identification of the Blueberry VcARF Gene Family

We predicted a total of 70 *VcARF* genes in the blueberry genome and named the genes *VcARF1–70*, according to their physical location on the chromosome. *VcARF* gene sequence lengths were highly varied, ranging from 1831 bp (*VcARF47*) to 17,346 bp (*VcARF25*). VcARF protein sequences ranged from 162 (VcARF18) to 1117 (VcARF50) amino acids, with molecular weights of 18.7–124.7 kDa and isoelectric points of 4.76–9.34 (Table 2).

Table 2. *VcARF* gene family in the genome of *Vaccinium corymbosum c.v.* 'Draper'.

Gene Name	Genome Position		Gene Length	Protein Length	Molecular Weight	Isoelectric Point
	Start	End				
VcARF1	18,439,481	18,443,830	4349	692	76,350.16	6.38
VcARF2	17,675,083	17,684,250	9167	573	62,768.41	5.91
VcARF3	35,513,223	35,522,595	9372	710	78,858.21	5.78
VcARF4	5,089,913	5,099,455	9542	706	78,334.61	6.11
VcARF5	17,374,973	17,379,982	5009	692	76,311.12	6.32
VcARF6	26,259,124	26,264,169	5045	707	78,336.45	6.23
VcARF7	23,759,116	23,764,049	4933	692	76,321.11	6.3
VcARF8	16,724,921	16,733,537	8616	709	78,059.18	7.05
VcARF9	22,380,725	22,385,479	4754	692	76,294.09	6.41
VcARF10	27,366,853	27,375,393	8540	889	98,168.12	6.20
VcARF11	30,200,121	30,204,632	4511	683	75,418.63	7.52
VcARF12	30,535,118	30,544,916	9798	782	86,669.81	6.21
VcARF13	31,908,720	31,918,385	9665	1073	118,643.05	6.08
VcARF14	31,971,246	31,975,574	4328	325	36,417.71	5.79
VcARF15	5,638,023	5,646,179	8156	896	99,747.22	5.46
VcARF16	2,964,750	2,975,312	10,562	818	91,211.21	5.88
VcARF17	39,198,442	39,206,540	8098	1092	21,781.74	6.26
VcARF18	7,176,872	7,179,011	2139	162	18,729.58	5.70
VcARF19	13,331,772	13,340,852	9080	574	62,818.47	5.91
VcARF20	30,285,241	30,295,980	10,739	880	98,106.11	6.94
VcARF21	25,800,314	25,808,896	8582	889	98,196.17	6.26
VcARF22	28,595,844	28,600,407	4563	683	75,510.72	7.52
VcARF23	28,958,634	28,968,521	9887	782	86,602.73	6.16
VcARF24	31,429,406	31,436,228	6822	827	92,363.56	6.21
VcARF25	27,786,604	27,803,950	17,346	260	28,558.81	5.67
VcARF26	30,557,305	30,565,768	8463	884	97,559.4	6.09
VcARF27	33,308,214	33,312,630	4416	683	75,381.6	7.85
VcARF28	33,590,293	33,600,143	9850	771	85,367.26	6.09
VcARF29	36,291,973	36,299,899	7926	747	84,171.56	7.97
VcARF30	36,316,721	36,323,616	6895	510	58,026.11	9.34
VcARF31	36,347,477	36,352,792	5315	572	63,904.6	5.55
VcARF32	17,724,006	17,728,602	4596	674	74,482.07	5.98
VcARF33	35,733,011	35,740,217	7206	623	69,462.96	7.82
VcARF34	34,950,200	34,957,557	7357	896	99,694.07	5.46
VcARF35	35,564,675	35,571,979	7304	896	99,804.31	5.46
VcARF36	4,968,620	4,978,140	9520	870	96,855.96	5.81
VcARF37	7,369,136	7,383,171	14,035	993	112,009.2	6.18
VcARF38	12,687,653	12,696,034	8381	884	97,599.42	6.12
VcARF39	9,683,280	9,692,882	9602	782	86,682.81	6.21
VcARF40	9,951,498	9,956,041	4543	683	75,383.64	7.85
VcARF41	23,235,922	23,244,210	8288	574	62,848.5	5.91
VcARF42	29,522,528	29,525,843	3315	267	30,862.32	5.21
VcARF43	35,261,834	35,272,570	10,736	880	98,106.11	6.94
VcARF44	17,822,463	17,827,109	4646	674	74,375.88	6.02
VcARF45	27,973,739	27,982,227	8488	828	92,414.65	6.21
VcARF46	1,273,882	1,279,100	5218	565	63,200.73	5.55
VcARF47	1,307,489	1,309,320	1831	402	45,446.23	8.97
VcARF48	17,166,955	17,171,684	4729	674	74,338.78	6.06
VcARF49	28,538,840	28,549,308	10,468	773	86,055.44	5.83
VcARF50	295,030	303,309	8279	1117	124,713.39	6.64
VcARF51	17,097,454	17,102,070	4616	674	74,269.76	6.06
VcARF52	6,479,810	6,486,650	6840	822	91,863	6.14
VcARF53	11,275,837	11,283,809	7972	745	82,051.86	6.89
VcARF54	28,655,149	28,663,653	8504	667	73,970.28	8.61
VcARF55	14,535,659	14,540,819	5160	707	78,310.36	6.23
VcARF56	18,962,251	18,971,430	9179	405	46,288.47	4.76
VcARF57	16,109,186	16,113,795	4609	707	78,280.34	6.23
VcARF58	16,725,386	16,730,572	5186	707	78,306.42	6.23
VcARF59	21,377,071	21,385,465	8394	972	107,956.17	5.98
VcARF60	21,432,199	21,439,229	7030	580	65,502.63	5.99
VcARF61	26,970,680	26,980,225	9545	870	96,855.96	5.81
VcARF62	27,717,127	27,724,764	7637	636	70,995.13	8.82
VcARF63	19,890,774	19,899,622	8848	901	99,470.38	6.15
VcARF64	19,942,680	19,949,685	7005	580	65,561.79	6.08
VcARF65	25,169,999	25,180,560	10,561	827	92,046.12	5.75
VcARF66	3,134,356	3,144,615	10,259	816	90,885.94	5.82
VcARF67	1,635,026	1,639,892	4866	540	60,252.66	8.29
VcARF68	16,996,236	17,002,796	6560	213	24,700	5.6
VcARF69	19,316,087	19,323,149	7062	623	69,462.01	7.82
VcARF70	18,207,611	18,214,839	7228	623	69,508.02	7.84

3.2. Phylogenetic Analysis of VcARF Amino Acids and Structure of VcARF Genes

To explore the evolutionary relationships of ARF proteins among *Arabidopsis* and blueberry, we constructed a phylogenetic tree based on the alignments of 22 *Arabidopsis* ARFs (AtARFs) and 70 blueberry ARFs (VcARFs). All ARF proteins could be divided into

six groups, from clade I to clade VI, according to the classification pattern of *Arabidopsis*. Both *Arabidopsis* and blueberry ARFs were present in each group. Clade VI (Figure 1, green branches) contained most ARFs, including 12 AtARF and 17 VcARF members. Clade IV (Figure 1, orange branches) contained 23 ARFs, of which only three were AtARFs. Clade III (Figure 1, blue branches) included the smallest number of ARF transcription factors, with only three VcARFs and one AtARF (AtARF5).

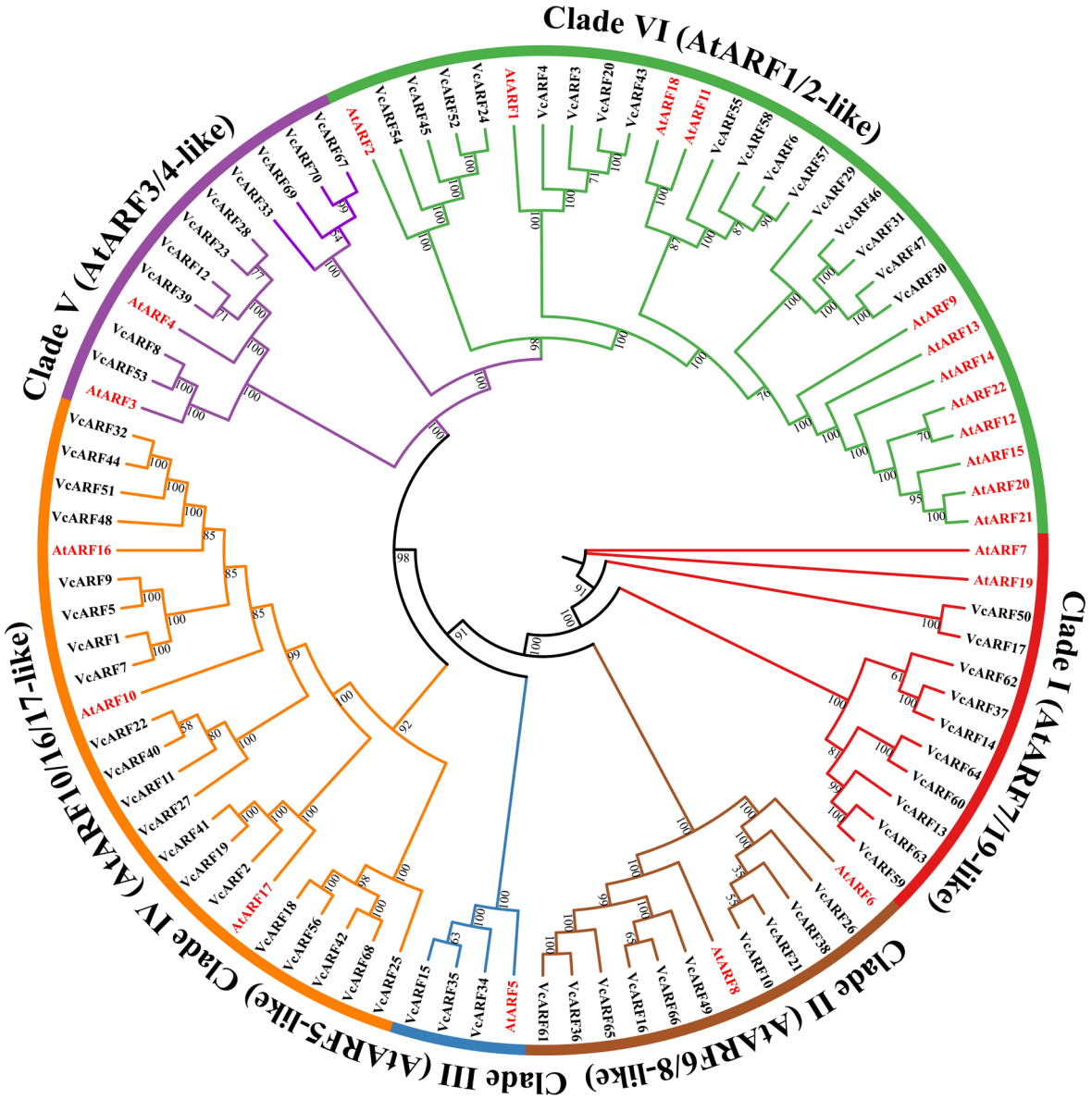

Figure 1. Phylogenetic relationships and classification of ARF proteins from *Arabidopsis* and blueberry. Numbers on the nodes indicate bootstrap values.

Gene structural analysis provided further evidence to support the phylogenetic topology groupings of multigene families. To gain further insights into the structural diversity of blueberry *ARF* genes, we analyzed the exon/intron organization of full-length cDNAs with corresponding genomic DNA sequences of individual *ARF* genes (Figure 2). Similar ARF classification patterns to the phylogenetic tree were observed, with *VcARF*s clustering into six groups according to their gene structure. Most closely related *VcARF*s within the same groups shared similar gene structures in terms of either intron numbers or exon lengths. Taking clade IV as an example, most genes in this group had two to four exons, with the

exception of *VcARF18*, *VcARF25*, *VcARF42*, *VcARF56*, and *VcARF68*. These genes contained no untranslated regions or longer introns. The gene structure appeared to be more variable in clades I, II, and VI, which had the largest number of exon/intron structural variants.

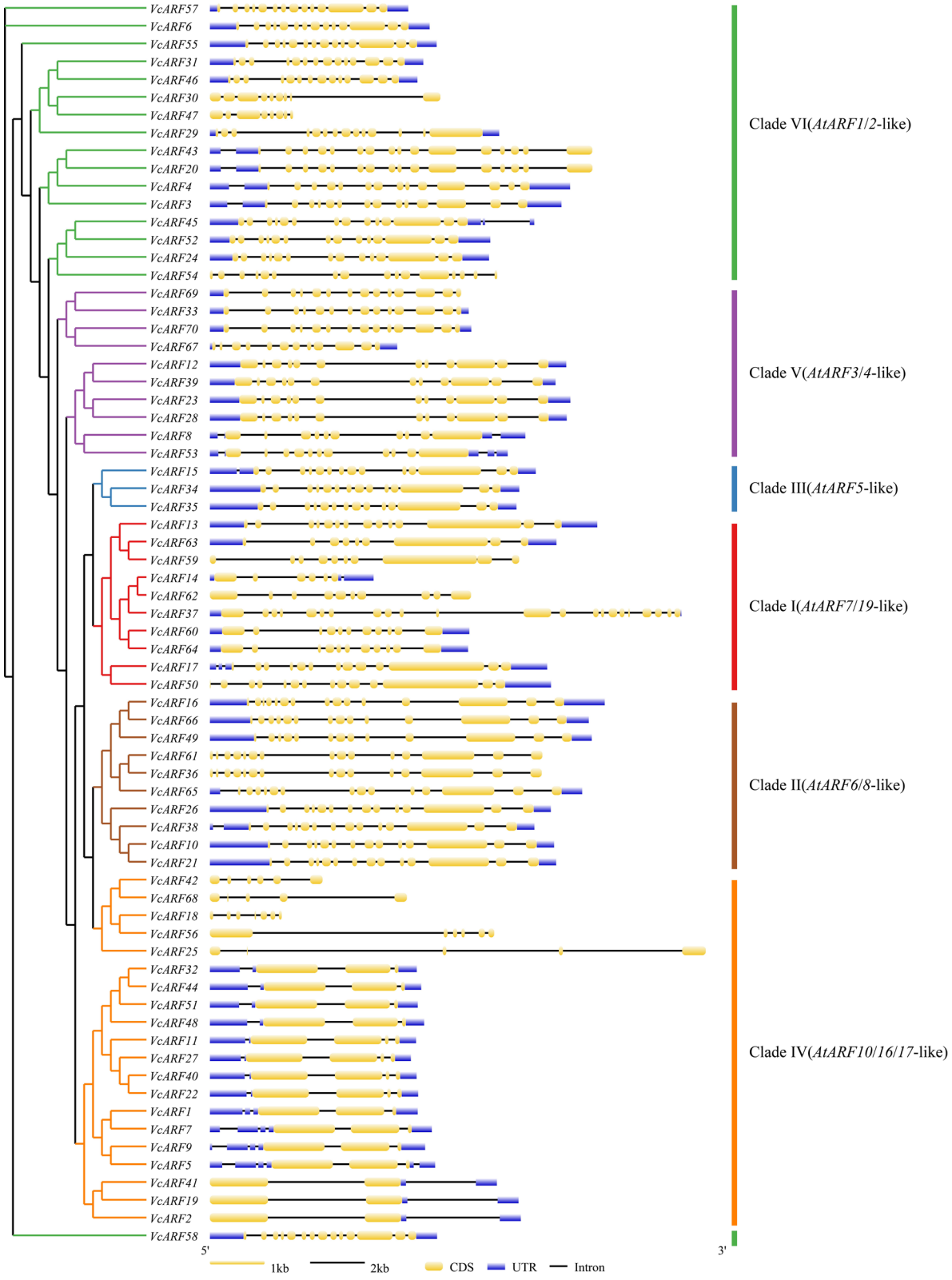

Figure 2. Intron and exon distribution of *VcARF* genes.

3.3. Chromosome Location and Conservative Motif Analysis

Next, we linked 70 *VcARF* genes onto a chromosome-scale genome assembly of the tetraploid highbush blueberry genome. All *VcARF* genes were mapped onto 40 linkage groups (Figure 3). The distribution of *VcARFs* on each chromosome was uneven, ranging from one to four. No *VcARF* was mapped onto chromosomes 4, 7, 16, 31, 32, 36, 45, or 48. Chromosomes 23 and 40 each had four *VcARFs*. *VcARFs* clustered in adjacent regions on chromosomes containing more than two *VcARF*s.

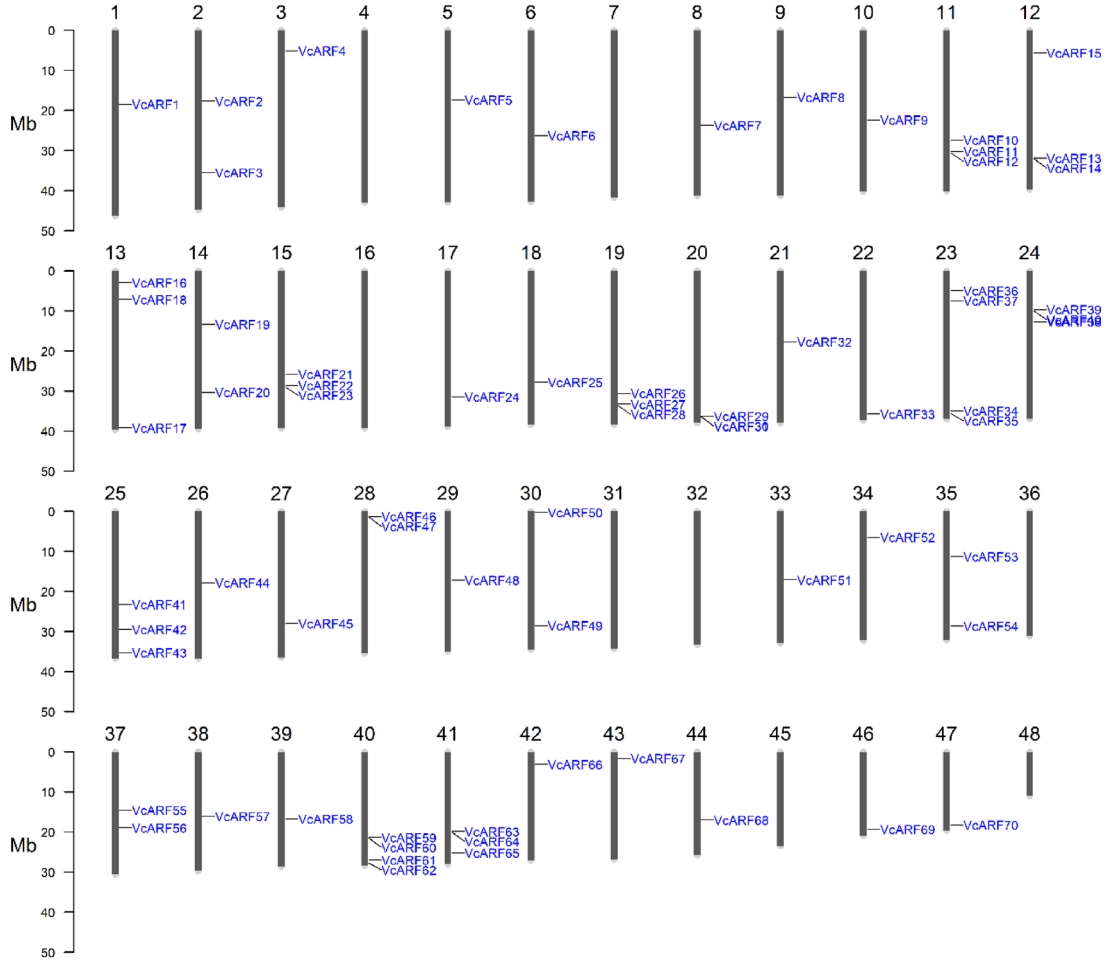

Figure 3. The chromosomal location of blueberry *VcARF* genes.

Conservative motifs of the blueberry VcARF protein were analyzed using the MEME online database. Identified motifs varied in length from 11–41 amino acid residues. A total of 14 conserved motifs were identified from VcARF members in clade II (Figure 4, brown branches) and clade III (Figure 4, blue branches), with VcARF49 as an exception. The same number of motifs was identified in clades II and III, with just a difference in their order. Neither motif 3 nor motif 13 were identified in VcARF49. Some conservative motifs were absent from clades I (Figure 4, red branches), IV (Figure 4, orange branches), and VI (Figure 4, green branches), especially for VcARF18, VcARF25, VcARF42, VcARF56, and VcARF68, in which only three conserved motifs were identified. The position and number of conserved motifs among VcARF members in clade I were highly varied, showing the most motif differences among all groups. Only five conserved motifs were found in VcARF62, which is less than that of other VcARFs in this group.

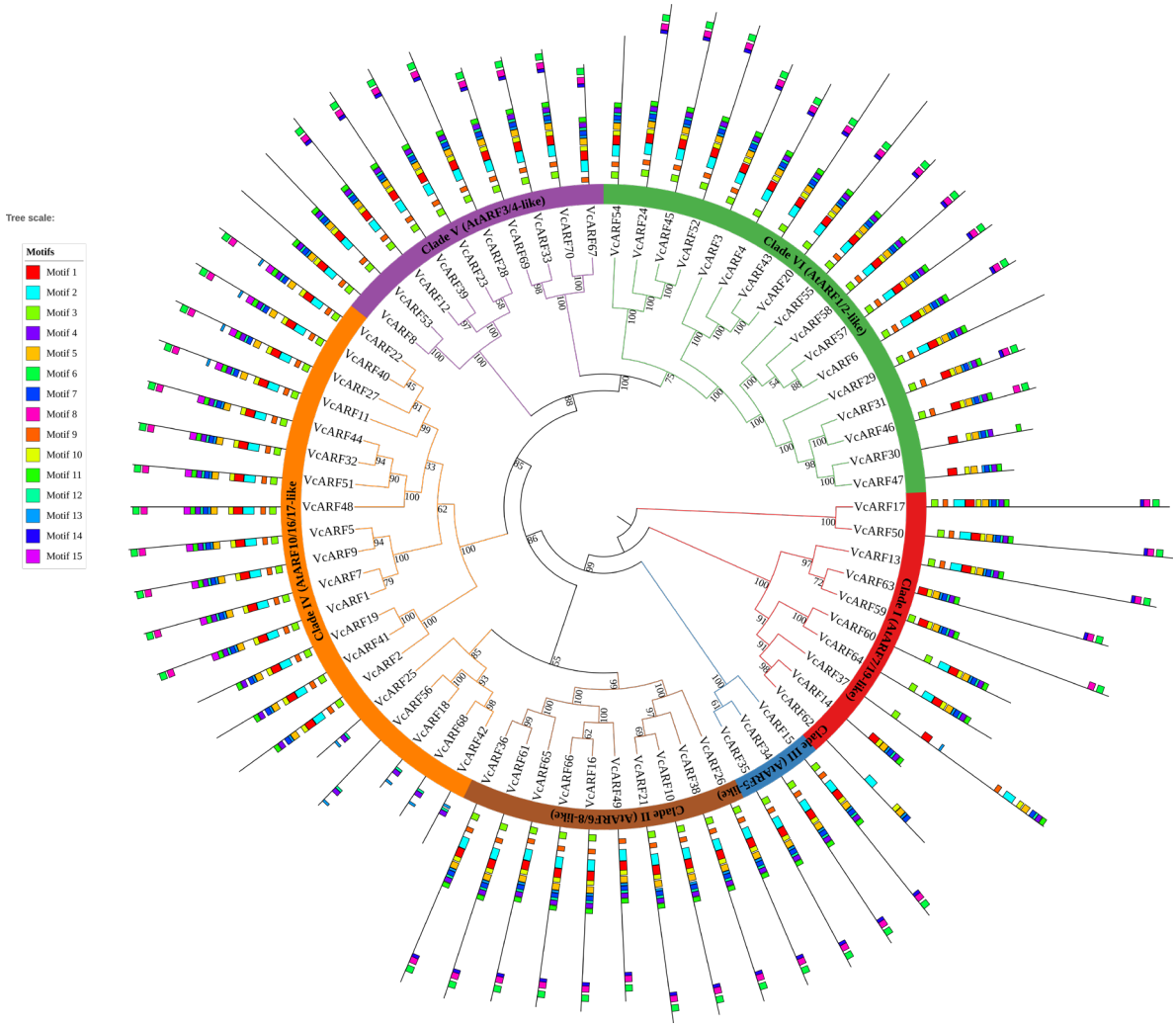

Figure 4. Conserved motif analysis of VcARF proteins in blueberry.

3.4. Prediction of Cis-Acting Elements in Promoters of Gene Family Members

To identify potential cis-acting elements of *VcARF*s, we analyzed 2 kbp promoter regions using bioinformatics analysis. This indicated that most cis elements of *VcARF* promoters belong to the responsive elements of plant hormones and environmental factors. In addition to two core elements (a TATA box and CAAT box), light responsive elements and MYB binding sites were identified in all *VcARF* genes (Figure 5). Oxygen-related metabolism and methyl jasmonate-responsive elements accounted for 88.6% and 75.7%, respectively, of *VcARF* cis-acting elements. Gibberellins and auxin-related responsive elements were also common in *VcARF* promoter regions, with 52.8% of *VcARF* genes containing the latter. Only *VcARF62* and *VcARF67* contained wound-responsive elements. Few cis elements related to development, i.e., cell cycle regulation, circadian control, and endosperm growth, were found in *VcARF* promoter regions. However, four cell cycle regulation and two circadian control responsive elements were found in the promoters of *VcARF15* and *VcARF63*, respectively.

3.5. Expression Pattern of VcARF Genes at Three Development Stages of 'Draper' Fruits

FPKM values of transcripts in different tissues of blueberry 'Draper' were retrieved from Colle et al. [34]. *VcARF* FPKM values at green mature (grnfrt), pink mature (pinkfrt), and mature (ripe) stages were extracted and used to construct a heat map (Figure 6). Dramatic differences were found among the expression patterns of those *VcARF* genes. A total of 58 *VcARF* genes showed low expression and few changes in expression during the

fruit ripening process. Genes that clustered into the same clade in Figures 1 and 2 displayed similar expression patterns. *VcARF4* and *VcARF52* had the highest expression level in green mature fruit, showing a transcript abundance that decreased with fruit ripening, leading to an obvious downregulation during the loss of fruit firmness. Other genes such as *VcARF14* and *VcARF37* in clade I (*AtARF7/19*-like), *VcARF16* and *VcARF66* in clade II (*AtARF6/8*-like), and *VcARF3* in clade VI (*AtARF1/2*-like) had higher transcript levels in green mature fruit compared with the other two stages.

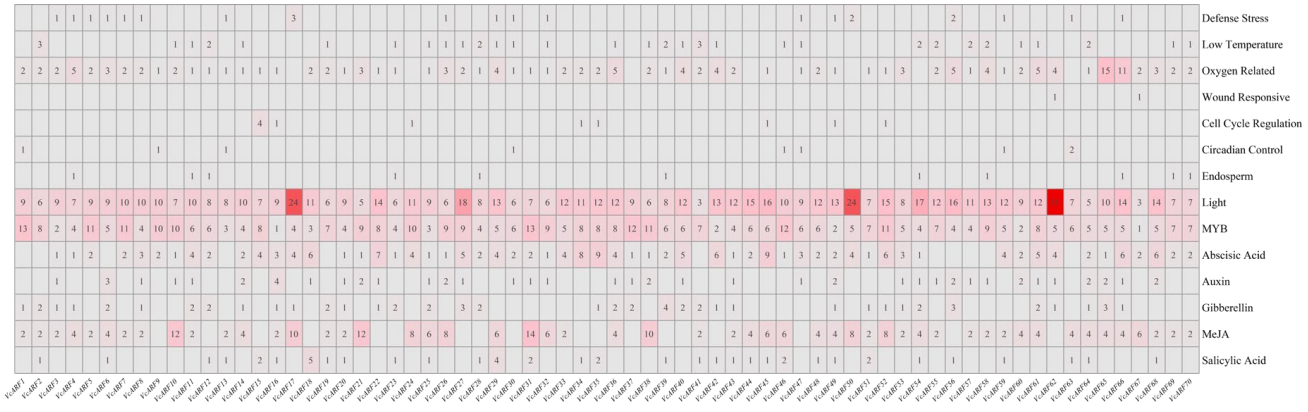

Figure 5. Types and numbers of cis-acting elements in the promoter regions of *VcARF* genes in blueberry.

3.6. Gene Expression during 'Star' and 'O'Neal' Fruit Development and Ripening

Because the expression level of *VcARF* genes varied according to different maturation stages, we speculated that some genes had an effect on fruit firmness. *VcARF* genes showing significant expression differences in the ripening process of blueberry 'Draper' were therefore selected to have their expression profiles validated at four stages of development of firm and soft flesh blueberries. Hence, the expression patterns of *VcARF3*, *VcARF4*, *VcARF14*, *VcARF37*, and *VcARF52* were evaluated in the fruit of 'Star' and 'O'Neal' using qPCR (Figure 7). Moderate changes were observed between the expression profiles of these five *VcARF* genes in 'Draper' transcriptomic data and qPCR results. The expression patterns of *VcARF3* and *VcARF37* in 'Star' and 'O'Neal' were similar to those in 'Draper'. The expression of *VcARF4* and *VcARF14* in firm flesh cultivar 'Star' decreased sharply from stage S5, which showed similar FPKM value changes to those observed in 'Draper'. However, the expression levels of these two genes were almost unchanged in the soft flesh cultivar 'O'Neal', particularly for *VcARF14*. By contrast, a slight change in the expression level of *VcARF52* was seen in 'Star' compared with 'O'Neal'.

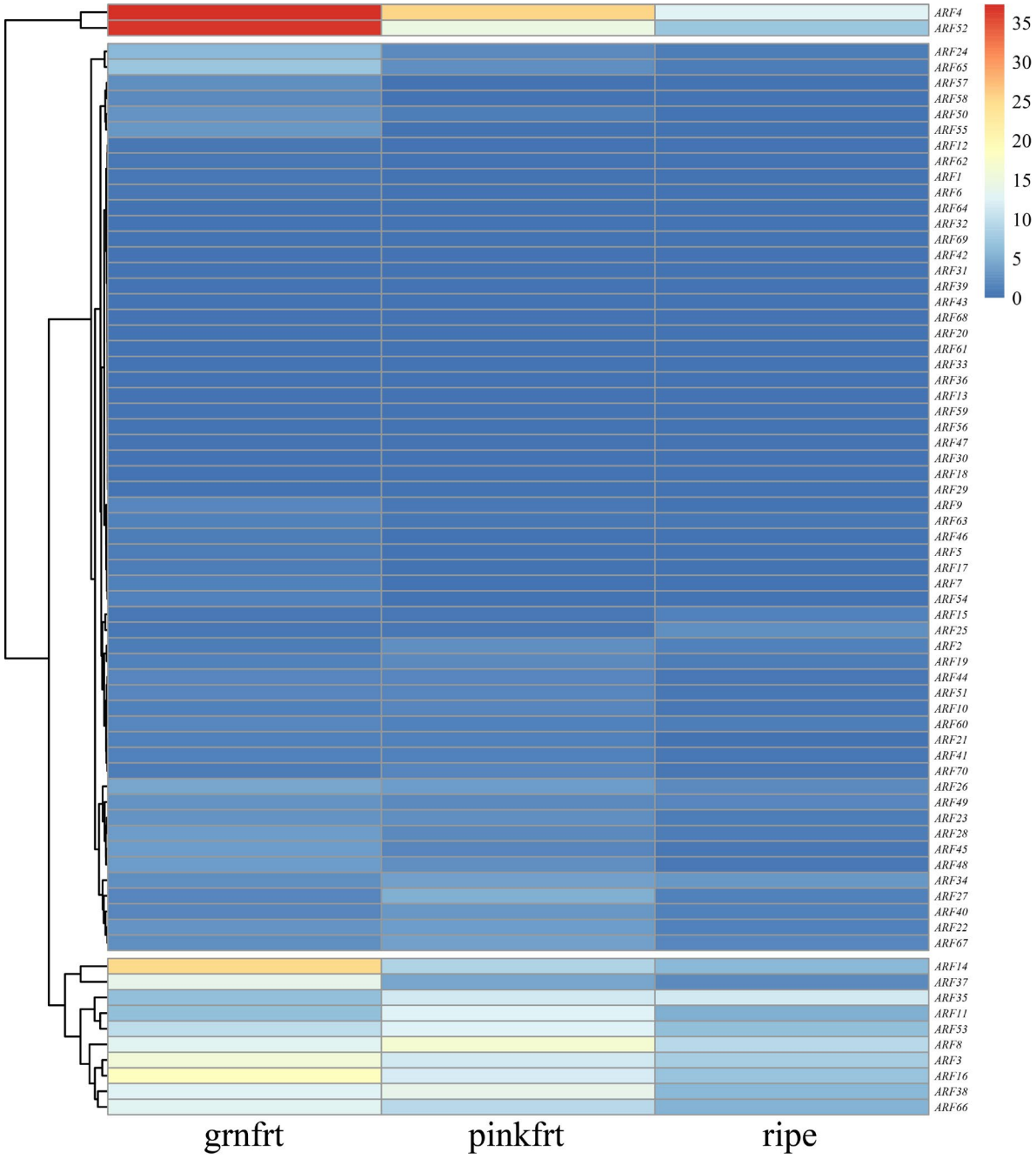

Figure 6. Heat map of RNA-seq data for *ARF* genes during fruit development.

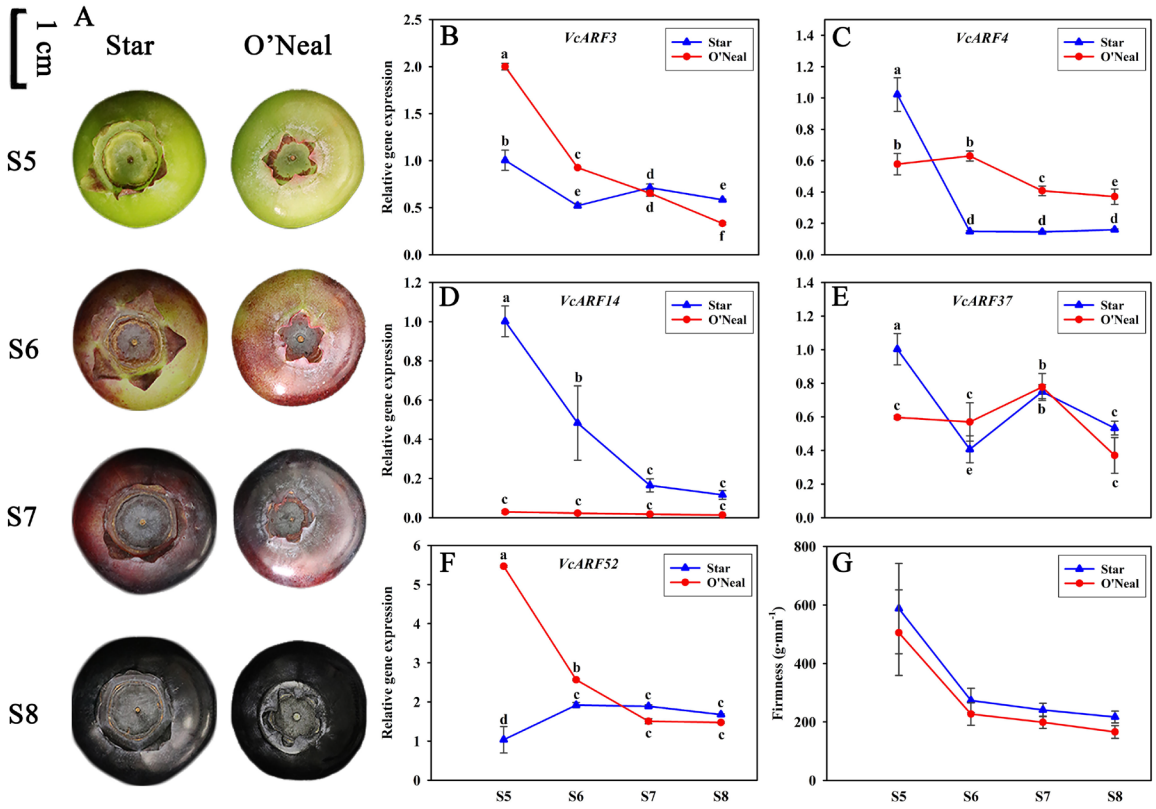

Figure 7. (**A**) Fruits of 'Star' and 'O'Neal' at different stages of development. (**B**–**F**) Relative expression of five *VcARF* genes at four development stages of 'Star' and 'O'Neal'. (**G**) Fruit firmness of 'Star' and 'O'Neal' at different stages of development.

4. Discussion

4.1. Expanded VcARF Family Members and High Sequence Divergence in Blueberry

ARF gene families have been identified in many sequenced plant genomes, with 22 *ARFs* identified in *A. thaliana*, 25 in *O. sativa* [14], 22 in *S. lycopersicum* [21], and 17 in *V. vinifera* [29]. In the current study, we identified 70 *VcARF* genes in a tetraploid blueberry genome, suggesting that the blueberry *ARF* gene family is expanded compared with other genomes. Long evolutionary periods typically lead to the presence of multiple members of a specific gene family [14,45]. Moreover, a recent whole-genome duplication event has been reported in the blueberry genome, in which tandem duplications may have contributed to metabolic diversity or gene functionalization [34]. Gene duplication has been shown to have a non-negligible effect on the formation of gene families [46], so we speculate that whole-genome duplication and tandem duplication are the main contributing events for the expansion of the blueberry *ARF* gene family.

Phylogenetic analysis indicated that many blueberry *VcARF* family members share high levels of similarity, while the adjacent chromosome locations of some *VcARF* genes provide further support for tandem duplication. Duplicated *ARF* genes exhibited different expression patterns, probably because of the lack of intense evolutionary selection pressure and the need for diversification [47,48]. Previous studies demonstrated that phylogenetically close *ARFs* may also be genetically linked, while phylogenetically distant ARFs may not be [48]. However, we found no correlation between phylogenetic distance and genetic link in blueberry *ARFs*. Because ARFs are transcription factors that regulate the expression of auxin response genes, those *VcARF* genes would be expected to contain AuxREs in their promoter regions. However, only 37 *VcARF* genes contained AuxREs in their promoter regions, of which 12 had more than two AuxREs. This indicates that the underlying mechanism for auxin inducibility of *VcARF* genes needs to be further elucidated in blueberry.

4.2. The Potential Contribution of VcARF Genes to Firmness Divergence between Firm Flesh and Soft Flesh Blueberries

Previous studies on the ripening process of tomato [21], grape [29], and apple [27] suggested that *ARF* genes were involved in the regulation of fruit development, especially with respect to floral organ and fruit development. *AtARF1* appeared to function as a transcriptional repressor in planta [16], while *ARF1* repressed auxin-induced gene expression in transient assays [11], and *arf1* mutations increased the transcription of *Aux/IAA* genes in *Arabidopsis* flowers [16]. Our phylogenetic analysis results showed that VcARF3 and VcARF4 had the highest amino acid similarity to AtARF1. Moreover, transcript abundances of *VcARF3* and *VcARF4* decreased greatly at stage S5 and changed slightly during blueberry fruit ripening, suggesting that *VcARF3* and *VcARF4* function during the early stages of fruit maturation. This indicates a similarity in the roles of *VcARF3* and *VcARF4* with that of *AtARF1*. We detected the significant downregulation of *VcARF3* expression during ripening in the 'O'Neal' cultivar, and a significant early expression decline of *VcARF4* in 'Star'. Therefore, both *VcARF3* and *VcARF4* have a likely repressor function during blueberry fruit ripening, with *VcARF3* having a dominant effect on fruit firmness.

VcARF14 and *VcARF37* are both phylogenetically related to *Arabidopsis AtARF7* and *AtARF19*. Because of the reported high level of similarity between ARF7 and ARF19 proteins, the expression of one ARF allows for functional compensation for the loss of the other in *arf7* and *arf19* single mutants [12]. In tomato, *SlARF7* acts as a negative regulator of fruit set until pollination and fertilization have taken place and moderates the auxin and gibberellin response during fruit growth. RNA interference-silenced *SlARF7* expression leads to parthenocarpic fruit growth, indicating that SlARF7 acts as a negative regulator of both fruit set and fruit development [24]. In our study, the expression of *VcARF14* in 'Star' showed significant differences with that in 'O'Neal', while an overall low transcript abundance and almost no expression changes were observed for *VcARF14* during fruit ripening in 'O'Neal'. *VcARF37* showed similar expression patterns between 'Star' and 'O'Neal'. Based on the protein sequence similarity between VcARF14 and AtARF7, we speculate that VcARF14 may also be a negative regulator of fruit ripening. The fruit firmness of 'Star' and 'O'Neal' decreased sharply from stage S5, but that of 'Star' was significantly greater than that of 'O'Neal' at all four stages, especially at stage S5 and S6. The higher expression levels of *VcARF14* in 'Star' compared with 'O'Neal' may account for this difference in firmness, but further studies should investigate the regulatory mechanisms of *VcARF14* during fruit softening.

VcARF52 was identified as an AtARF1/2-like transcription factor in clade VI (Figure 1) and has a similar sequence to that of *Arabidopsis* ARF2. ARF2 is most similar to ARF1, with these two proteins having both distinct and overlapping functions in *A. thaliana* [16]. ARF2 was reported to regulate leaf senescence and floral organ abscission independently of the ethylene and cytokinin response pathways [16]. ARF2A is a recognized auxin signaling component that may interconnect signals of ethylene and additional hormones to co-ordinate and initiate the complex ripening process of tomato [22]. Over-expressing *ARF2A* in tomato resulted in blotchy ripening in which certain fruit regions reddened and showed accelerated ripening [22]. This suggested that *SlARF2A* has a positive impact on tomato ripening. However, we found that the expression of *VcARF52* in 'O'Neal' decreased with fruit ripening, while its expression in 'Star' did not significantly change. Therefore, *VcARF52* probably accelerated the softening process at an early stage of blueberry fruit development but had little effect at later stages. Although ARF2 is thought to regulate fruit firmness by affecting fruit development and ripening, phylogenetic studies indicate that ARF1 and ARF2 diverged prior to the monocot–dicot split [49,50] so would have had ample time to evolve distinct biochemical activities. Thus, *VcARF52* may not conform to the canonical auxin response model.

Taken together, our findings suggest that *ARF* genes play essential roles in the ripening process of blueberry fruit. We propose that future studies should focus on *VcARF3*, *VcARF4*,

VcARF14, and *VcARF52* to elucidate their function in determining differences in fruit firmness between firm and soft flesh cultivars.

Author Contributions: Conceptualization, Y.Z., L.X. and W.G.; data curation, L.G. and H.K.; formal analysis, L.G., Z.S. and Y.Z.; funding acquisition; L.X. and W.G.; writing—original draft, Y.Z. and L.G.; writing—review and editing, Y.L. and F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Key Research Project of Science Technology Department of Zhejiang Province, grant number 2021C02066-9; the Key Research and Development Program of Zhejiang Science and Technology Department, grant number 2018C02007; and the Natural Science Foundation of Zhejiang Province, grant number LQ16C150001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Chandler, J.W. Auxin response factors. *Plant Cell Environ.* **2016**, *39*, 1014–1028. [CrossRef] [PubMed]
2. Guilfoyle, T.J.; Hagen, G. Auxin response factors. *Curr. Opin. Plant Biol.* **2007**, *10*, 453–460. [CrossRef] [PubMed]
3. Chapman, E.J.; Estelle, M. Mechanism of auxin-regulated gene expression in plants. *Annu. Rev. Genet.* **2009**, *43*, 265–285. [CrossRef]
4. Wang, R.; Zhang, Y.; Kieffer, M.; Yu, H.; Kepinski, S.; Estelle, M. HSP90 regulates temperature–dependent seedling growth in Arabidopsis by stabilizing the auxin co-receptor F-box protein TIR1. *Nat. Commun.* **2016**, *7*, 10269. [CrossRef] [PubMed]
5. Salehin, M.; Bagchi, R.; Estelle, M. SCFTIR1/AFB-Based Auxin Perception: Mechanism and Role in Plant Growth and Development. *Plant Cell* **2015**, *27*, 9–19. [CrossRef]
6. Guilfoyle, T.J.; Hagen, G. Getting a grasp on domain III/IV responsible for Auxin Response Factor–IAA protein interactions. *Plant Sci.* **2012**, *190*, 82–88. [CrossRef] [PubMed]
7. Wright, R.C.; Nemhauser, J.L. New tangles in the auxin signaling web. *F1000Prime Rep.* **2015**, *7*, 19. [CrossRef]
8. Li, S.; Xie, Z.; Hu, C.; Zhang, J. A review of Auxin Response Factors (ARFs) in plants. *Front. Plant Sci.* **2016**, *7*, 47. [CrossRef]
9. Boer, D.R.; Freire-Rios, A.; van den Berg, W.A.M.; Saaki, T.; Manfield, I.; Kepinski, S.; López-Vidrieo, I.; Franco-Zorrilla, J.M.; de Vries, S.C.; Solano, R.; et al. Structural basis for DNA binding specificity by the *a*uxin-dependent ARF transcription factors. *Cell* **2014**, *156*, 577–589. [CrossRef]
10. Tiwari, S.B.; Hagen, G.; Guilfoyle, T. The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. *Plant Cell* **2003**, *15*, 533–543. [CrossRef]
11. Ulmasov, T.; Hagen, G.; Guilfoyle, T.J. Activation and repression of transcription by auxin-response factors. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5844–5849. [CrossRef]
12. Okushima, Y.; Overvoorde, P.J.; Arima, K.; Alonso, J.M.; Chan, A.; Chang, C.; Ecker, J.R.; 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*, 444–463. [CrossRef]
13. Hardtke, C.; Ckurshumova, W.; Vidaurre, D.P.; Singh, S.A.; Stamatiou, G.; Tiwari, S.B.; Hagen, G.; Guilfoyle, T.J.; Berleth, T. Overlapping and non-redundant functions of the *Arabidopsis* auxin response factors *MONOPTEROS* and *NONPHOTOTROPIC HYPOCOTYL 4*. *Development* **2004**, *131*, 1089–1100. [CrossRef] [PubMed]
14. 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*, 13–24. [CrossRef] [PubMed]
15. Guilfoyle, T.J.; Hagen, G. Auxin response factors. *J. Plant Growth Regul.* **2001**, *20*, 281–291. [CrossRef]
16. Ellis, C.M.; Nagpal, P.; Young, J.C.; Hagen, G.; Guilfoyle, T.J.; Reed, J.W. AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* **2005**, *132*, 4563–4574. [CrossRef] [PubMed]
17. Sessions, A.; Nemhauser, J.L.; McColl, A.; Roe, J.L.; Feldmann, K.A.; Zambryski, P.C. ETTIN patterns the Arabidopsis floral meristem and reproductive organs. *Development* **1997**, *124*, 4481–4491. [CrossRef] [PubMed]
18. Pekker, I.; Alvarez, J.P.; Eshed, Y. Auxin Response Factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity. *Plant Cell* **2005**, *17*, 2899–2910. [CrossRef]
19. Wu, M.; Tian, Q.; Reed, J. Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development* **2006**, *133*, 4211–4218. [CrossRef]
20. Wilmoth, J.C.; Wang, S.; Tiwari, S.B.; Joshi, A.D.; Hagen, G.; Guilfoyle, T.J.; Alonso, J.M.; Ecker, J.R.; Reed, J.W. NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. *Plant J.* **2005**, *43*, 118–130. [CrossRef]

21. Wu, J.; Wang, F.; Cheng, L.; Kong, F.; Peng, Z.; Liu, S.; Yu, X.; Lu, G. Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum*. *Plant Cell Rep.* **2011**, *30*, 2059–2073. [CrossRef]
22. Breitel, D.A.; Chappell-Maor, L.; Meir, S.; Panizel, I.; Puig, C.P.; Hao, Y.; Yifhar, T.; Yasuor, H.; Zouine, M.; Bouzayen, M.; et al. AUXIN RESPONSE FACTOR 2 intersects hormonal signals in the regulation of tomato fruit ripening. *PLoS Genet.* **2016**, *12*, e1005903. [CrossRef]
23. Sagar, M.; Chervin, C.; Mila, I.; Hao, Y.; Roustan, J.-P.; Benichou, M.; Gibon, Y.; Biais, B.; Maury, P.; Latché, A.; et al. SlARF4, an Auxin Response Factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiol.* **2013**, *161*, 1362–1374. [CrossRef]
24. de Jong, M.; Wolters-Arts, M.; Feron RMariani, C.; Vriezen, W.H. The *Solanum lycopersicum* auxin response factor 7 (*SlARF7*) regulates auxin signaling during tomato fruit set and development. *Plant J.* **2009**, *57*, 160–170. [CrossRef]
25. Yuan, Y.; Mei, L.; Wu, M.; Wei, W.; Shan, W.; Gong, Z.; Zhang, Q.; Yang, F.; Yan, F.; Luo, Y.; et al. SlARF10, an auxin response factor, is involved in chlorophyll and sugar accumulation during tomato fruit development. *J. Exp. Bot.* **2018**, *69*, 5507–5518. [CrossRef]
26. Xing, H.; Pudake, R.N.; Guo, G.; Xing, G.; Hu, Z.; Zhang, Y.; Sun, Q.; Ni, Z. Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. *BMC Genom.* **2011**, *12*, 178. [CrossRef] [PubMed]
27. Luo, X.; Sun, M.; Xu, R.; Shu, H.; Wang, J.; Zhang, S. Genome wide identification and expression analysis of the ARF gene family in apple. *J. Genet.* **2014**, *93*, 785–797. [CrossRef]
28. 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. [CrossRef]
29. Wan, S.; Li, W.; Zhu, Y.; Liu, Z.; Huang, W.; Zhan, J. Genome-wide identification, characterization and expression analysis of the auxin response factor gene family in *Vitis vinifera*. *Plant Cell Rep.* **2014**, *33*, 1365–1375. [CrossRef] [PubMed]
30. Peng, Y.; Fang, T.; Zhang, Y.; Zhang, M.; Zeng, L. Genome-Wide identification and sxpression analysis of Auxin Response Factor (ARF) gene family in Longan (*Dimocarpus longan* L.). *Plants* **2020**, *9*, 221. [CrossRef] [PubMed]
31. Wang, S.; Shi, F.; Dong, X.; Li, Y.; Zhang, Z.; Li, H. Genome-wide identification and expression analysis of auxin response factor (ARF) gene family in strawberry (*Fragaria vesca*). *J. Integr. Agric.* **2019**, *18*, 1587–1603. [CrossRef]
32. Bian, Y.; Ballington, J.; Raja, A.; Brouwer, C.; Reid, R.; Burke, M.; Wang, X.; Rowland, L.J.; Bassil, N.; Brown, A. Patterns of simple sequence repeats in cultivated blueberries (*Vaccinium* section *Cyanococcus* spp.) and their use in revealing genetic diversity and population structure. *Mol. Breed.* **2014**, *34*, 675–689. [CrossRef]
33. Gupta, V.; Estrada, A.D.; Blakley, I.; Reid, R.; Patel, K.; Meyer, M.D.; Andersen, S.U.; Brown, A.F.; Lila, M.A.; Loraine, A.E. RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. *GigaScience* **2015**, *4*, 5. [CrossRef]
34. Colle, M.; Leisner, C.P.; Wai, C.M.; Ou, S.; Bird, K.A.; Wang, J.; Wisecaver, J.H.; Yocca, A.E.; I Alger, E.; Tang, H.; et al. Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *GigaScience* **2019**, *8*, giz012. [CrossRef]
35. Zifkin, M.; Jin, A.; Ozga, J.; Zaharia, L.I.; Schernthaner, J.P.; Gesell, A.; Abrams, S.R.; Kennedy, J.A.; Constabel, C.P. Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. *Plant Physiol.* **2011**, *158*, 200–224. [CrossRef]
36. Chang, S.; Puryear, J.; Cairney, J. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* **1993**, *11*, 113–116. [CrossRef]
37. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
38. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2020**, *49*, D412–D419. [CrossRef] [PubMed]
39. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* **2011**, *39*, W29–W37. [CrossRef]
40. Katoh, K.; Standley, D.M. MAFFT: Multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
41. Nguyen, L.; Schmidt, H.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2014**, *32*, 268–274. [CrossRef]
42. Anand, L.; Lopez, C.M.R. ChromoMap: An R package for interactive visualization and annotation of chromosomes. *bioRxiv* **2020**, 605600. [CrossRef]
43. Wickham, H. ggplot2: Elegant graphics for data analysis. *J. R. Stat. Soc. Ser. A* **2011**, *174*, 245–246. [CrossRef]
44. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}{}_{T}$ method. *Methods* **2001**, *25*, 402–408. [CrossRef] [PubMed]
45. Danilevskaya, O.N.; Meng, X.; Hou, Z.; Ananiev, E.V.; Simmons, C.R. A genomic and expression compendium of the expanded PEBP gene family from maize. *Plant Physiol.* **2007**, *146*, 250–264. [CrossRef]
46. Wang, Y.; Tan, X.; Paterson, A.H. Different patterns of gene structure divergence following gene duplication in Arabidopsis. *BMC Genom.* **2013**, *14*, 652. [CrossRef] [PubMed]
47. Lynch, M.; Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **2000**, *154*, 459–473. [CrossRef]

48. Kumar, R.; Tyagi, A.K.; Sharma, A.K. Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. *Mol. Genet. Genom.* **2011**, *285*, 245–260. [CrossRef]
49. Remington, D.L.; Vision, T.J.; Guilfoyle, T.J.; Reed, J. Contrasting modes of diversification in the *Aux/IAA* and *ARF* gene families. *Plant Physiol.* **2004**, *135*, 1738–1752. [CrossRef]
50. Sato, A.; Yamamoto, K.T. Overexpression of the non-canonical *Aux/IAA* genes causes auxin-related aberrant phenotypes in Arabidopsis. *Physiol. Plant.* **2008**, *133*, 397–405. [CrossRef]