Plant Cell Rep (2011) 30:2059–2073
DOI 10.1007/s00299-011-1113-z

ORIGINAL PAPER

Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum*

Jian Wu · Feiyan Wang · Lin Cheng · Fuling Kong · Zhen Peng · Songyu Liu · Xiaolin Yu · Gang Lu

Received: 24 March 2011 / Revised: 5 May 2011 / Accepted: 5 May 2011 / Published online: 7 July 2011

Abstract Auxin response factors (ARFs) encode transcriptional factors that bind specifically to the TGTCTC-containing auxin response elements found in the promoters of primary/early auxin response genes that regulate plant development. In this study, investigation of the tomato genome revealed 21 putative functional ARF genes (*SlARFs*), a number comparable to that found in *Arabidopsis* (23) and rice (25). The full cDNA sequences of 15 novel *SlARFs* were isolated and delineated by sequencing of PCR products. A comprehensive genome-wide analysis of this gene family is presented, including the gene structures, chromosome locations, phylogeny, and conserved motifs. In addition, a comparative analysis between ARF family genes in tomato and maize was performed. A phylogenetic tree generated from alignments of the full-length protein sequences of 21 *OsARFs*, 23 *AtARFs*, 31 *ZmARFs,* and 21 *SlARFs* revealed that these ARFs were clustered into four major groups. However, we could not find homologous genes in rice, maize, or tomato with *AtARF12-15* and *AtARF20*-23. The expression patterns of tomato ARF genes were analyzed by quantitative real-time PCR. Our comparative analysis will help to define possible functions for many of these newly isolated ARF-family genes in plant development.

Keywords *Solanum lycopersicum* · SlARF · ZmARF · Expression analysis · qRT-PCR · Phylogenetic analysis

Abbreviations

ARF	Auxin response factor
SlARF	*Solanum lycopersicum* auxin response factor
ZmARF	*Zea mays* auxin response factor
AtARF	*Arabidopsis thaliana* auxin response factor
OsARF	*Oryza sativa* auxin response factor
qRT-PCR	quantitative Real-time PCR
SOL	The International Solanaceae Project
PCR	Polymerase chain reaction
BlASTN	Search a nucleotide database using a nucleotide query
MaizeGDB	The Maize Genetics and Genomics Database
TBLASTN	Search translated nucleotide database using a protein query
ORF	Open reading frame

Communicated by C. Quiros.

Electronic supplementary material The online version of this article (doi:10.1007/s00299-011-1113-z) contains supplementary material, which is available to authorized users.

J. Wu · F. Wang · L. Cheng · F. Kong · Z. Peng · S. Liu · X. Yu · G. Lu
Key Laboratory of Horticultural Plant Growth, Development and Biotechnology, Agricultural Ministry of China, Hangzhou 310058, People's Republic of China

J. Wu · F. Wang · L. Cheng · F. Kong · Z. Peng · S. Liu · X. Yu · G. Lu (✉)
Department of Horticulture, Zhejiang University, Hangzhou 310058, People's Republic of China
e-mail: glu@zju.edu.cn

Introduction

Auxin regulates a host of plant developmental and physiological processes, including embryogenesis, organogenesis, tropic growth, and root and shoot architecture (Quint and Gray 2006). Two types of transcription factor families are required for controlling the expression of auxin response genes, auxin response factors (ARFs), and

Aux/IAA repressors (Guilfoyle and Hagen 2007). Members of the Aux/IAA family are generally regarded as repressors of auxin-induced gene expression (Ulmasov et al. 1997). Meanwhile, ARFs activate or repress the expression of auxin response genes by binding to auxin response elements (AuxREs) on promoters of auxin response genes (Tiwari et al. 2003). A number of putative AuxREs have been defined within the upstream promoter regions of primary/early auxin responsive genes, including one or more copies of the conserved motif TGTCTC (Ulmasov et al. 1999b). A typical ARF protein contains a conserved N-terminal B3-like DNA-binding domain (DBD) that regulates expression of auxin response genes, a conserved C-terminal dimerization domain (CTD) that resembles domains III and IV in Aux/IAA proteins, and a variable middle region (MR) (Ulmasov et al. 1997; Guilfoyle and Hagen 2007), located between the DBD and CTD, that determines whether the ARF functions as a transcriptional activator or repressor (Ulmasov et al. 1999a; Tiwari et al. 2003).

Recent advances have provided information on regulation of ARF gene expression, ARF roles in growth and developmental processes, and target genes regulated by ARFs (Liscum and Reed, 2002; Guilfoyle and Hagen 2007). It has been demonstrated that the ARF proteins participate in the transcriptional regulation of a variety of biological processes related to growth and development such as embryogenesis (Hamann et al. 2002; Weijers et al. 2006) leaf expansion (Wilmoth et al. 2005) leaf senescence (Lim et al. 2010), lateral root growth (Tatematsu et al. 2004; Okushima et al. 2007; Marin et al. 2010), and fruit development (Goetz et al. 2006, 2007; Guillon et al. 2008; Jong et al. 2009), as well as various responses to environmental stimuli. Recently, the involvement of ARF family members was reported in ethylene (Li et al. 2006), brassinosteroid (Vert et al. 2008), and ABA responses (Yoon et al. 2010).

Twenty-three ARF genes have been identified in the *Arabidopsis* genome, distributed over all five chromosomes (Wei and Cui 2006). Sequencing of the rice (*Oryza sativa*) genome (Rice Genome Initiative 2000), revealed 25 genes, distributed over 10 of the 12 rice chromosomes, that were postulated to encode proteins belonging to the ARF family (Wang et al. 2007; Shen et al. 2010). Phylogenetic analyses revealed that individual members of transcription factor families are clustered into subgroups of genes that are most closely related to other members of that same subgroup in Arabidopsis and rice. Recently, a total of 39 *PoptrARFs* and 24 *SvARF*s genes were also identified in *Populus trichocarpa* (Kalluri et al. 2007) and sorghum (*Sorghum vulgare*) genome (Andrew et al. 2009), respectively. In addition, the complete cDNA sequences of all 31 maize *ZmARFs* genes were also submitted to GenBank (Alper et al. 2009). However, in tomato, only 6 *SlARF* genes, including *SlARF2, SlARF3*, *SlARF4*, *SlARF6, SlARF7,* and *SlARF8,* have been identified and shown to be homologous to *AtARFs* (Alvarez et al. 2006; Goetz et al. 2007; Feng et al. 2009; Jong et al. 2009). To date, no systematic investigations of ARF family proteins have been reported in tomato until recently (Kumar et al. 2011). Moreover, Functional analysis of each transcription factor of the ARF family has not been performed, despite the importance of ARF proteins in multiple aspects of plant physiology.

The Genome Sequencing Project for tomato has been completed lately, and the ARFs of *Arabidopsis*, rice, and maize have also been published, so it is now feasible to carry out a genome-wide search for tomato homologues and to conduct a comparative analysis of ARFs for these four species. To elucidate the structure of *SlARF* and characterize expression during reproduction in tomato, 21 putative genes with ARF domains were identified through genomic data mining. The full cDNA sequences of 15 novel tomato ARFs were isolated by PCR-based approaches. The genomic structure, chromosomal location, and sequence homology of all *SlARFs* were then investigated, followed by comparative phylogenetic analysis, exon/intron mapping, and structural analysis of conserved protein motifs of the ARF family genes. Subsequently, the different temporal and spatial expression patterns during flowering and fruiting in tomato plants were determined for each *SlARF* gene by quantitative real-time PCR (qRT-PCR). The resulting classification of groups, identification of putative functional motifs, and characterization of the expression patterns will be useful for future analysis of the biological functions of ARF family genes in tomato.

Materials and methods

Searching for the ARF genes

Multiple database searches were performed to find all members of the ARF family in *Arabidopsis* (*Arabidopsis thaliana*), rice (*Oryza sativa* L. subsp. *japonica*), and maize (*Zea mays* L.). To find ARF genes in *Arabidopsis* (*AtARFs*) and rice (*OsARFs*), "auxin responsive factor" was used as a query to search the protein and nucleotide databases of The National Center for Biotechnology Information (NCBI), and the matching genes were confirmed by previous reports (Wang et al. 2007). Similarly, all 31 *ZmARFs* genes in maize were identified from the MaizeGDB Database (http://www.maizegdb.org).

To find previously identified and potential ARF family genes in tomato, multiple database searches were performed. First, "auxin responsive factor" was used as a query to search the SGN database (The tomato Information

Resource, http://solgenomics.net). Six known *SlARF* family genes were identified, including *SlARF2, SlARF3, SlARF4, SlARF6, SlARF7,* and *SlARF8.* To find other potential ARFs, we initially surveyed the SGN database using the amino acid sequences of the conserved ARF domains from all the known ARF families (including *AtARFs* and *OsARFs)* as queries. To increase the number of potential ARF proteins, we also performed the database searches using amino acid sequences of the ARF domains in some ARF family members in *Cucumis sativus*(3 members) and *Solanum lycopersicum* (6 members). Based on the combined results from all searches, we finally identified all members of tomato ARF family from the currently available genomic databases. After searching for ARF genes, bioinformatics tools, such as DNASTAR and FGENESH (http://linux1.softberry.com/berry) were used to analyze and predict those unknown *SlARFs*. NCBI ORF finder was used to find putative open reading frames and functional domains were determined by BLASTP of NCBI.

Isolation of the full-length cDNA sequence using RT-PCR

Total RNA was extracted from tomato ovaries using TRIZOL reagent (Invitrogen, Germany) according to the manufacturer's instructions. The first cDNA strand was generated using the Improm-TM Reverse Transcription system (Promega, Madison, WI, USA) following the manufacture's protocol. The full-length cDNA sequences of 15 novel *SlARFs* were amplified by PCR using primers designed based on the predicted results by FGENESH (listed in Supplementary Table 1). Since the predicted cDNA sequences were quite long (some even longer than 3.0 kb), it was difficult to design an adequate single pair of primers for the entire segment. Therefore, we designed two or more primer pairs for each ARF to amplify and clone the fragments. Then we assembled them into the whole target fragment, and verified the full-length cDNA by PCR with gene-specific primer sets and by BLASTN against SGN database. After optimization, the PCR conditions included denaturation at 94°C for 4 min, followed by 35 cycles of 30 s at 94°C, 30 s at 50–55°C (depending on the specific primers) and 70 s at 72°C, and a final 7-min elongation at 72°C. The amplifications were carried out using TGRADIENT Thermal Cycler machines (Biometra, Germany). The amplified cDNA fragments were cloned and sequenced using the ABI Prism 3730 sequencer (Invitrogen, Bioasia Biotech Co. Ltd). The DNA sequences were amplified utilizing gene-specific primer sets designed from the full length cDNA. Finally, the ORFs of the 15 unknown *SlARFs* were determined by ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and homologous alignment.

Mapping *SlARF* and *ZmARF* genes on chromosomes

To determine the location of *SlARF* and *ZmARF* genes on chromosomes, the *SlARF* and *ZmARF* sequences were further used as query sequences for the BLASTN search against SGN Tomato Whole genome Scaffolds data (2.30) (http://www.sgn.cornell.edu/tools/blast/) and Maize GDB B73 RefGen_v1 databases (http://www.maizegdb.org), respectively. Finally, the locations of all 21 *SlARFs* and 31 *ZmARFs* were detected.

Multiple-sequence alignments and phylogenetic analysis

Gene sequences were analyzed by DNAStar software and the net service ExPASy Proteomics Server (http://ca.expasy.org). Multiple-sequence alignments employed ClustalX v1.81 (Thompson et al. 1997). Phylogenetic analysis was performed using MEGA 4.1 program by the neighbor-joining (NJ) method (Saitou and Nei 1987). Conserved motifs were investigated by multiple alignment analyses using Clustal W.

Expression analysis of *SlARFs*

The plant materials used for expression analysis were sampled from tomato (*S. lycopersicum* L.) cv. Micro-Tom plants (Tomato Genetics Resource Center, University of California, Davis, USA) were grown until flowering in a temperature-controlled greenhouse at the experimental farm at Zhejiang University.

The leaves, stems, roots, and buds were collected from flowering tomato plants, and the various floral organs (sepal, petal, stamen, and ovary) were isolated from the flower buds (about 3 days before opening). To analyze the expression pattern of auxin response genes at different flower developmental stages, flower buds were collected at three stages of early floral development, which was roughly defined by the length of flower buds as follows: stage I: 3–4 mm, stage II: 5–6 mm, and stage III at 7–8 mm (Brukhin et al. 2003). In addition, the ovaries were sampled at 0, 3, 6, and 9 days after the flower fully opened. All the samples were frozen in liquid nitrogen immediately and stored at −75°C until RNA isolation.

Total RNA and the first cDNA strand were prepared as described earlier. QRT-PCR techniques were employed to determine characterize the gene expression profiles of new *SlARFs*. The quantitative real-time PCR (qRT-PCR) was carried out using the primer pairs listed in Supplementary Table S2. Because the four sister pairs, including SlARF6/SlARF6-1, SlARF8/SlARF8-1, SlARF13/SlARF13-1, and SlARF19/SlARF19-1 are so similar in nucleotide sequences, we only design appropriate primers for one member of

each pair. A sample of cDNA (1 μg) was subjected to real-time PCR in a final volume of 20 μl containing 12.5 μl SYBR Green Master Mix Reagent (Takara, Japan) and specific primers (3 pmol). Two biological and three technical replicates for each sample were performed in the real-time PCR machine (STRATAGENE, MX3500), programmed to heat for 30 s at 95°C, followed by 40 cycles of 5 s at 95°C and 45 s at 50°C, and at the end, one cycle of 1 min at 95°C, 30 s at 50°C, and 30 s at 95°C. To normalize the total amount of cDNA present in each reaction, the *Ubi3* gene was co-amplified as an endogenous control for calibration of relative expression. The $\Delta\Delta C_t$ method of relative gene quantification recommended by Applied Biosystems was used to calculate the expression level of different treatments.

Results

Identification and isolation of *SlARF* family genes in tomato

To identify the ARF family genes in tomato, BLAST searches of the SGN database were performed using the ARF domain of the *Arabidopsis* and rice protein as a query sequence. A total of 30 ARF-domain genome DNA sequences and two unigenes were obtained by TBLASTN at an *e* value of $1e^{-3}$ that were similar to ARF genes. All sequences were predicted by FGENESH (http://www.softberry.com/berry.phtml?topic=fgenesh). These predicted amino acid sequences were analyzed by blastp of NCBI to find their conserved domains, followed by homologous alignment with known *SlARF* genes. For these analyses, the tomato genome appeared to contain, in addition to the six previously known *SlARF* genes (Alvarez et al. 2006; Jong et al. 2009; Feng et al. 2009), 15 other putative novel ARF genes.

The PCR primers (Supplementary Table S1) were designed based on the predicted results of FGENESH. The potential full-length cDNA sequences of all 15 putative *SlARFs* were isolated through PCR-based approaches. The open reading frame (ORF) length of *SlARF* genes varied from 1,218 bp (*SlARF12*) to 3,371 bp (*SlARF7*), encoding polypeptides of 375–1,123 aa, with a predicted molecular mass range of 42.4–126.4 kD. The theoretical pI ranged from 5.48 to 8.58 (Table 1) and the calculated molecular masses of the deduced ORFs were almost identical with the sizes of ARF polypeptides previously determined in other plants (Wang et al. 2007).

It is noteworthy that the nomenclature system for *SlARFs* used in the present study, a generic name from *SlARF1* to *SlARF19-1*, was provisionally used to distinguish each of the ARF genes according to the homology between *AtARFs* and *SlARFs*. However, homologous genes for *SlARF11, SlARF12, SlARF13, SlARF13-1,* and *SlARF14* were not found in *Arabidopsis*, rice and maize, so they were named according to the order of submitted to the GenBank database.

Chromosomal locations of *SlARFs* and *ZmARFs*

The chromosomal locations and transcription directions of the 21 *SlARF* genes were determined and demonstrated using BLASTN analysis on Tomato WGS Chromosomes (Fig. 1). Similar to *Arabidopsis* and rice, SlARF family genes in tomato appeared to be distributed among all the linkage groups, except chromosome 9. The number of *SlARF* genes per chromosome ranged from one to three. Three *SlARFs* were present on chromosome 5, 6, and 11, two each were localized to chromosomes 2, 3, 8, and 12, and only one each to chromosomes 1, 4, 6, and 10 (Table 1; Fig. 1).

Interestingly, SlARF8 and SlARF8-1, SlARF19 and SlARF19-1, SlARF6, and SlARF6-1 were present in different chromosomes although they shared more than 90% amino acid sequence identity. The fact that the two genes on different strands were nearly identical suggested that they might be derived from recent gene duplication events. This finding is consistent with a previous report demonstrating that duplicated genes involved in signal transduction and transcription are preferentially retained compared with other functional gene categories (Blanc and Wolfe 2004).

Similarly, 31 *ZmARFs* distributed on 9 of 11 maize chromosomes. No *ZmARFs* was detected on chromosomes 7 or 11. Five *ZmARFs* were present on chromosome 5, four each on chromosomes 2, 4, 5, 6, and 10, two each on chromosomes 1 and 8, and only one *SlARF* on chromosomes 5. The location of *ZmARF8* by BLASTN was not found in present database (Fig. 1; Table 2).

Sequence analysis of the *SlARF* and *ZmARF* proteins

All the tomato SlARF protein sequences were found to contain DNA-binding domains (DBDs) and MR (middle region) domains (Table 1). All SlARF proteins contained a highly conserved region of about 390 amino acid residues in the N-terminal portion that corresponded to the DBD of the *Arabidopsis* ARF family (Fig. 2a). Fourteen deduced SlARF proteins contained a carboxyl-terminal domain (CTD) related to domains III and IV found in Aux/IAA proteins (Fig. 2b). Other seven of the SlARF proteins including SlARF3, SlARF6-1, SlARF13, SlARF13-1, SlARF14, and SlARF17 lacked a CTD.

Similarly, all ZmARF proteins contained a highly conserved N-terminal region of about 380 amino acid residues

Table 1 ARF gene in tomato

Gene name[a]	ORF length[b] (bp)	Deduced polypeptide[c]			Chromosome[d]	Pseudomolecule position (5′–3′)[e]	Domains[f]	Accession No.[g]
		Length (aa)	MW (kDa)	PI				
SlARF1	1,965	654	72.5	6.21	1	83496046–83499910	DBD MR CTD	HM061154.1
SlARF2	2,541	846	94	6.3	3	61269255–61275955	DBD MR CTD	DQ340255.1
SlARF3	2,244	747	81.4	7.03	2	37038600–37044855	DBD MR	DQ340254.1
SlARF4	2,436	811	90.3	5.86	11	50900407–50908834	DBD MR CTD	DQ340259.1
SlARF5	2,793	930	102.9	5.39	4	62817669–62822652	DBD MR CTD	HM195248.1
SlARF6	2,643	881	97.3	6.07	12	871710–879749	DBD MR CTD	HM594684.1
SlARF6-1	1,575	524	58.6	8.41	7	54886928–54892071	DBD MR	HM187579.1
SlARF7	3,372	1,123	126.4	6.19	7	52686671–52692503	DBD MR CTD	EF121545.1
SlARF8	2,535	844	94.6	5.93	2	21755763–21765341	DBD MR CTD	EF667342.1
SlARF8-1	2,535	844	94.4	5.85	3	8741104–8748438	DBD MR CTD	HM560979.1
SlARF9	1,785	594	66.7	8.57	8	62527488–62530848	DBD MR CTD	HM037250.1
SlARF10	2,100	699	77	8.45	11	51187714–51190849	DBD MR CTD	HM143941.1
SlARF11	2,043	680	76.3	8.58	12	42540651–42544219	DBD MR CTD	HM143940.1
SlARF12	1,218	405	46.1	8.49	8	2808202–2811405	DBD MR	HM565127.1
SlARF13	1,356	451	51.5	5.48	5	64570855–64574982	DBD MR	HM565128.1
SlARF13-1	1,419	472	54.4	5.78	5	64570767–64574981	DBD MR	HM565129.1
SlARF14	1,128	375	42.4	8.5	10	64394361–64395488	DBD MR	HM565131.1
SlARF16	2,016	671	74.9	5.77	6	43022652–43025267	DBD MR CTD	HM195247.1
SlARF17	1,869	622	68.35	5.66	11	6496429–6511309	DBD MR	HQ456923
SlARF19	3,036	1,011	123.4	6.21	7	6386964–6391364	DBD MR CTD	HM130544.1
SlARF19-1	3,273	1,090	120.6	6.02	5	58050452–58054715	DBD MR CTD	HM565130.1

DBD DNA binding domain, *MR* middle region, *CTD* carboxy-terminal dimerization domain

[a] Names referred to the known *SlARF* genes in tomato or *AtARFs* in Arabidopsis or given to tomato ARFs in this work

[b] Length of open reading frame in base pairs

[c] Length (number of amino acids), molecular weight (kDa), and isoelectric point (pI) of the deduced polypeptide

[d] Chromosomal localization of the SlARF gene

[e] The position of SlARFs on Chromosome

[f] Predicted by blastp of NCBI

[g] GenBank accession numbers present in NCBI

that corresponded to the DBD of the *Arabidopsis* ARF family (Fig. 3a). Twenty-two ZmARFs protein sequences contained three ARF domains, while 9 out of 31 ZmARFs only contained two domains and lacked a CTD domain (Table 2). In maize, the molecular mass of ZmARF protein sequences generally ranged from 50.56 kDa (ZmARF31) to127.49 kDa (ZmARF20) (Fig. 3b).

Gene structure and phylogenetic analysis of ARFs

A comparison of the full-length cDNA sequences with the corresponding genomic DNA sequences revealed the numbers and positions of exons and introns for each individual *SlARF* gene. The coding sequences of all the *SlARFs* except *SlARF14* were disrupted by introns. The number of introns varied from 1 (*SlARF15*) to 13 (*SlARF1, 2, 5, 8-1*) (Fig. 4a). It was suggested that *SlARF14* was the product of an mRNA inserted into the tomato genome (Babenko et al. 2004). Based on the presence of triplets containing SlARF10, 14, and 16 in the phylogenetic tree (Fig. 4a), we surmise that this mRNA might come from *SlARF10* mRNA, *SlARF16* mRNA, or both.

An unrooted phylogenetic tree was generated from the alignment of the full-length protein sequences of all SlARFs. The 21 SlARFs could be divided into three major classes (I–III, Fig. 4a) similar to those in rice (Wang et al. 2007). Class I was further divided into two sub-classes, Ia-1 with seven members and Ib with two members. Class II was also further divided into two sub-classes IIa and IIb, each containing four members. Class III contained four members that were the most divergent compared to those grouped into the other two classes. The *SlARFs* in class III

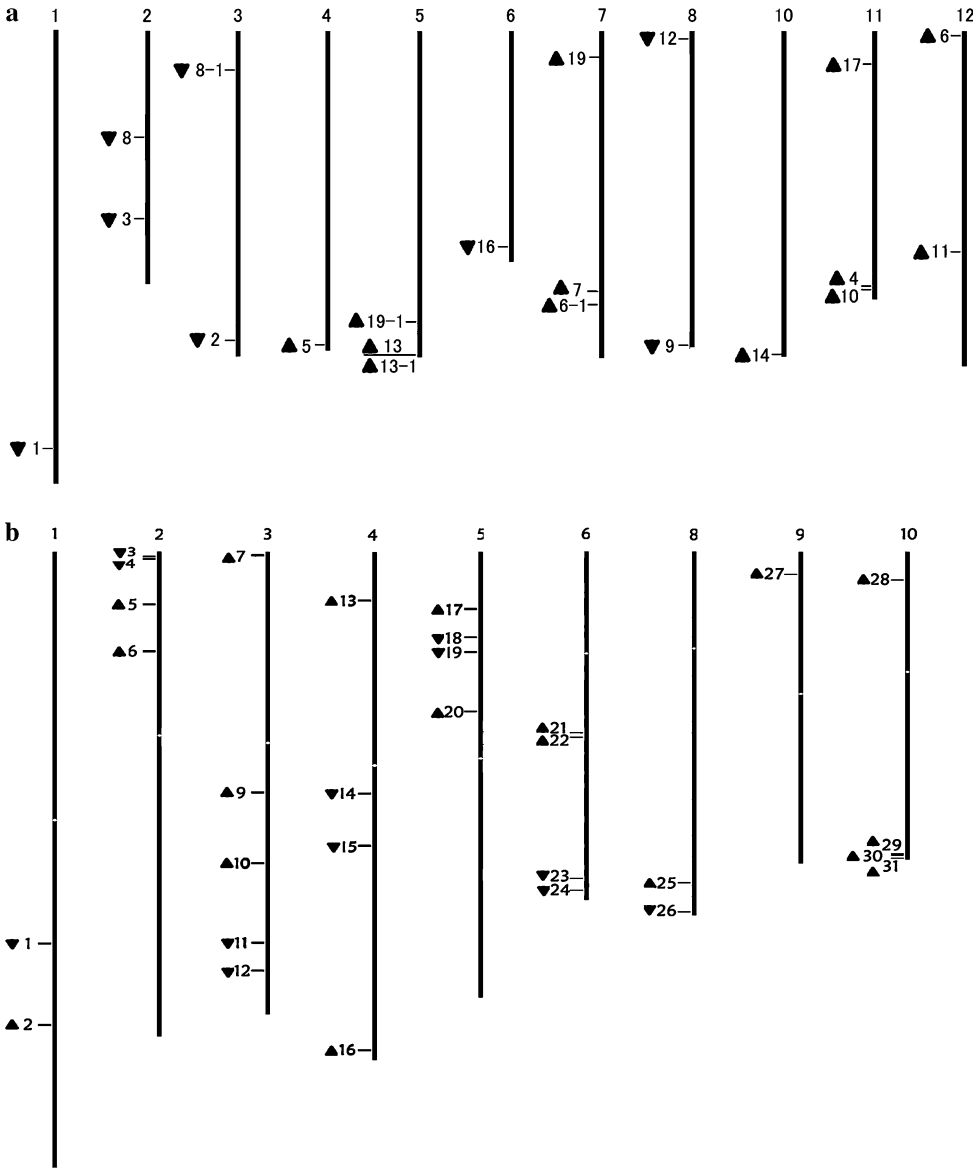

Fig. 1 **a** Genomic distribution of ARF genes on tomato chromosomes. **b** Genomic distribution of ARF genes on maize chromosomes. The *arrows* next to gene names show the direction of transcription. White ovals on the maize chromosomes (*vertical bar*) indicate the position of centromeres. The chromosome numbers are indicated at the *top* of each *bar*

contained fewer introns in their ORF regions than those in other groups, with one, *SlARF14,* even possessing no introns. The 21 *SlARFs* formed seven sister pairs (Fig. 4a), all with very strong bootstrap support (>99%).

The 31 ZmARFs were divided into four major classes, I–IV (Fig. 4b). Class I and class II were further subdivided into two subgroups, Ia-1 with seven members and Ib with five members, and IIa with six members, and IIb with five members. Classes III and IV contained six and two members, respectively. The 31 ZmARFs formed 13 sister pairs (Fig. 4b), while the ARF genes from *Arabidopsis* and rice formed 6 and 9 sister pairs, respectively (Wang et al. 2007).

To investigate the relationships of ARF proteins, the full-length protein sequences of the 23 AtARFs, 25 OsARFs, 31 ZmARFs, and 21 SlARFs were used to build the phylogenetic tree. All 100 ARF proteins could be classified into four major classes: class I contained 46 members, class II contained 33 members, class III contained 17, and class IV contained 4 gene members. Class I was divided into three subgroups, Ia-1 (24 members), Ia-2(8 members) and Ib (14 members). Class II was further divided into two subgroups, IIa-1 with 16 members and IIb with 17 members. This classification is very similar to that of AtARFs except for Class IV (Supplementary Fig. 1).

In the joint phylogenetic tree, a total of 51 sister pairs were formed, including 7 SlARF–SlARF pairs, 13 ZmARF–ZmARF pairs, 6 AtARF–AtARF pairs, 9 OsARF–OsARF pairs, 10 OsARF–ZmARF pairs, 5 SlARF–AtARF pairs, and one OsARF–AtARF pair. Interestingly, subgroups Ia-1, Ib, IIa, IIb, and class III contained ARF genes from all the four species, but only Arabidopsis AtARFs proteins were present in subgroup Ia-2, while in class IV, only ARFs from rice and maize (monocotyledon) were present (Fig. 5).

Table 2 ARF gene in maize

Gene name[a]	ORF length[b] (bp)	Deduced polypeptide[c]			Chromosome[d]	Pseudomolecule position (5′-3′)[e]	Domains[f]
		Length (aa)	MW (kDa)	PI			
ZmARF1	3,273	1,090	120.61	6.00	1	190425934–190440385	DBD MR CTD
ZmARF2	2,070	689	75.10	8.24	1	230050814–230054396	DBD MR CTD
ZmARF3	2,451	816	90.86	5.95	2	2254440–2260389	DBD MR CTD
ZmARF4	2,808	935	102.80	5.78	2	3326838–3331570	DBD MR CTD
ZmARF5	1,542	513	55.45	5.97	2	25344439–25348528	DBD MR
ZmARF6	1,974	657	72.95	5.85	2	47923807–47928519	DBD MR CTD
ZmARF7	2,061	686	76.72	5.92	3	1549577–1553240	DBD MR CTD
ZmARF8	2,124	707	75.36	6.63	–	142909530–142911874	DBD MR CTD
ZmARF9	2,646	881	97.05	5.68	3	120046239–120051115	DBD MR CTD
ZmARF10	2,400	799	89.23	5.91	3	155258715–155263028	DBD MR CTD
ZmARF11	2,067	688	74.99	6.89	3	194771595–194776621	DBD MR CTD
ZmARF12	2,127	708	77.97	6.67	3	208704580–208710364	DBD MR
ZmARF13	2,568	855	93.70	6.66	4	23392694–23397733	DBD MR CTD
ZmARF14	2,019	672	74.86	5.89	4	117505349–117510219	DBD MR CTD
ZmARF15	2,136	711	75.95	6.77	4	142909530–142911874	DBD MR
ZmARF16	2,718	905	100.25	5.74	4	242084605–242088841	DBD MR CTD
ZmARF17	1,935	644	70.96	6.52	5	27805185–27808975	DBD MR CTD
ZmARF18	2,757	918	101.47	6.16	5	41636156–41641236	DBD MR CTD
ZmARF19	2,151	716	77.52	6.67	5	48559282–48561684	DBD MR
ZmARF20	3,450	1,149	127.49	5.85	5	77679754–77687038	DBD MR CTD
ZmARF21	2,097	698	75.01	8.41	6	88003320–88005774	DBD MR
ZmARF22	2,778	925	102.25	6.01	6	90218419–90222597	DBD MR CTD
ZmARF23	2,043	680	73.95	6.40	6	158770425–158776238	DBD MR
ZmARF24	2,211	736	80.46	7.87	6	164617532–164621982	DBD MR
ZmARF25	2,406	801	89.76	6.09	8	158808291–158812635	DBD MR CTD
ZmARF26	2,061	686	74.35	6.68	8	172945404–172950741	DBD MR
ZmARF27	3,162	1,053	116.75	6.20	9	11007475–11018124	DBD MR CTD
ZmARF28	2,442	813	89.71	6.69	28	13837429–13842601	DBD MR CTD
ZmARF29	2,838	945	103.84	5.93	28	146657465–146662383	DBD MR CTD
ZmARF30	2,430	809	90.31	5.85	28	147260154–147267078	DBD MR CTD
ZmARF31	1,389	462	50.56	5.38	28	148713848–148715450	DBD MR

DBD DNA binding domain, *MR* middle region, *CTD* carboxy-terminal dimerization domain

[a] The name of ZmARF present in NCBI

[b] Length of open reading frame in base pairs

[c] Length (number of amino acids), molecular weight (kDa), and isoelectric point (pI) of the deduced polypeptide

[d] Chromosomal localization of the *ZmARF* gene

[e] The position of ZmARFs on chromosome

[f] Predicted by blastp of NCBI

Expression characterization of *SlARF* genes

A previous report demonstrated that ARF genes were constitutive expressed (Wang et al. 2007). In our study, we also found that most of the *SlARFs* could be detected in root, stem, buds, and ovary using qRT-PCR (Fig. 6). The *SlARF5* mRNA was more highly expressed in stem and leaf than in root, flower, and ovary. Stem exhibited higher expression of *SlARF6*, *SlARF13*, and *SlARF19-1* than other organs, while *SlARF1*, *SlARF2*, and *SlARF3* were mainly expressed in leaf. Meanwhile, higher mRNA level of *SlARF9, SlARF16,* and *SlARF17* was detected in root.

qRT-PCR analysis demonstrated that expression of most *SlARF* genes could be detected in all parts of the flower (Supplementary Fig. 4). *SlARF6*, *SlARF7,* and *SlARF8* were expressed at a higher level in sepal and petal than in

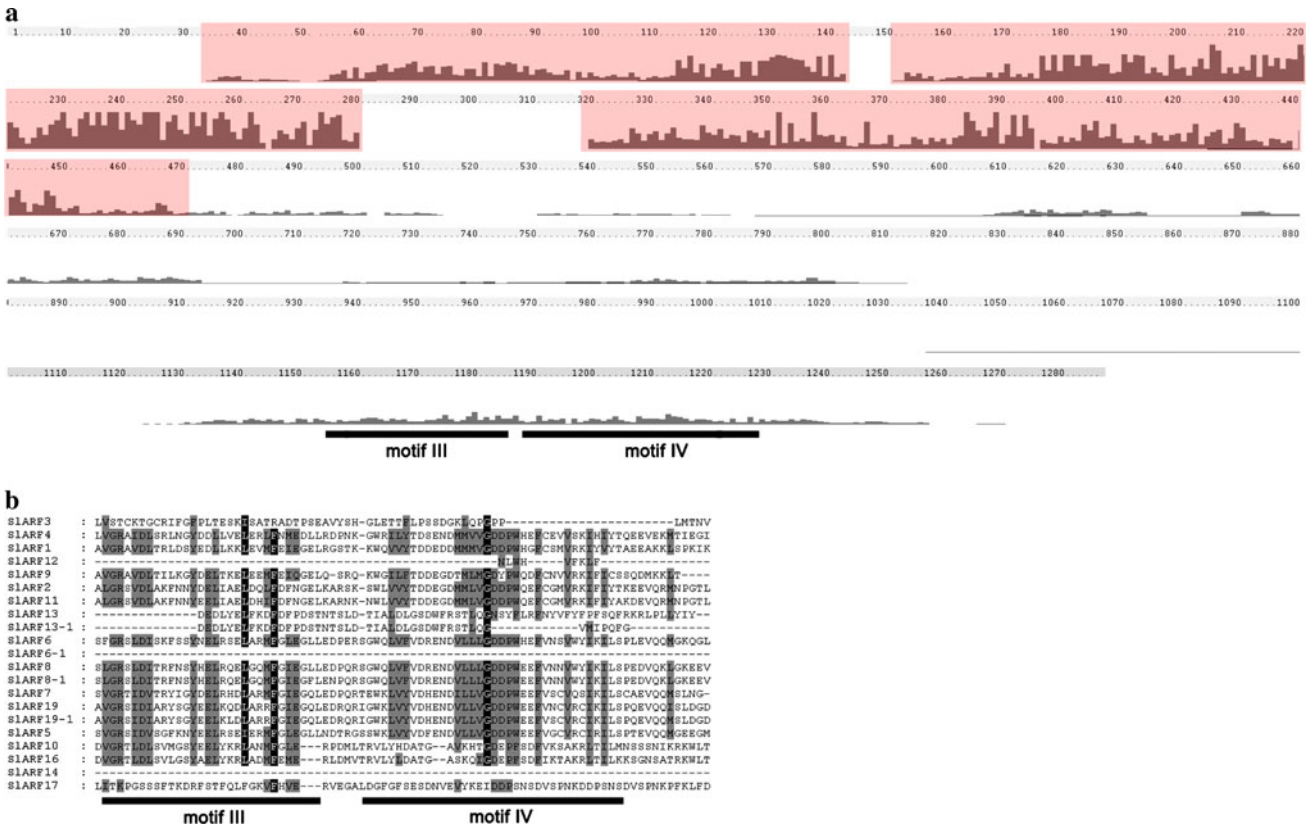

Fig. 2 **a** Alignment profile of tomato ARF proteins obtained with the ClustalX program. The height of the *bars* indicates the number of identical residues per position. The *shaded regions* indicate the high sequence similarity among DBDs regions. Motifs III and IV are consensus sequences shared by Aux/IAA proteins. **b** Multiple alignments of Motifs III and IV in tomato ARF proteins obtained with ClustalX. *Black* and *light gray* shading indicate identical and conversed amino acid residues, respectively. Conserved domains are also *underlined* and correspond to part **a**

stamen and ovary, while *SlARF7, SlARF14, SlARF16, SlARF17,* and *SlARF19-1* were expressed mainly in petal. The *SlARF9, SlARF10,* and *SlARF11* mRNAs were detected in ovary at a higher level than in other tissues.

During the different developmental periods of tomato flowers, most of the *SlARF* genes detected exhibited a similar expression pattern (Supplementary Fig. 5). Most *SlARF* mRNAs increased during tomato flower development, while only *SlARF4*, *SlARF9,* and *SlARF17* mRNA levels significantly decreased during flower development. Similar expression patterns were also detected during the different periods of ovary development (Supplementary Fig. 6). With the development of ovary and young fruit, the expression level increased to a peak on the third day after flower opening and then markedly decreased on the nineth day. Only *SlARF16* mRNA reached its highest level on sixth day after flower opening, and then decreased significantly on the ninth day. Noticeably, the expression level of *SlARF4* was quite different with the other ARF genes, constantly increasing even after pollination during young fruit development.

Discussion

In this study, 15 novel tomato ARF genes were identified and the full-length cDNA sequences of these *SlARFs* were isolated. The total number of 21 SlARFs detected in present study is more than the 17 SlARFs identified by Kumar et al. (2011) using publically available tomato EST databases. Our gene isolation and sequencing strategy was based on TBLASTN and PCR-based methods which allowed us to find more ARF genes and get exact and comprehensive data. Comparing the deducing amino acid sequences, 14 corresponding SlARF genes can be found in the report of Kumar et al. (2011). The number of SlARF members from tomato is comparable to that of *Arabidopsis* (23) and rice (25) (Okushima et al. 2005; Wang et al. 2007), although <39 in populus (Kalluri et al. 2007). Meanwhile, a total of 31 putative maize ARF genes were also predicted and analyzed from the MaizeGDB Database. The genome sizes of tomato, maize, Arabidopsis, and rice are quite different (tomato ~950 Mb, maize ~2,300 Mb, *Arabidopsis* ~125 Mb, rice ~450 Mb), as are the

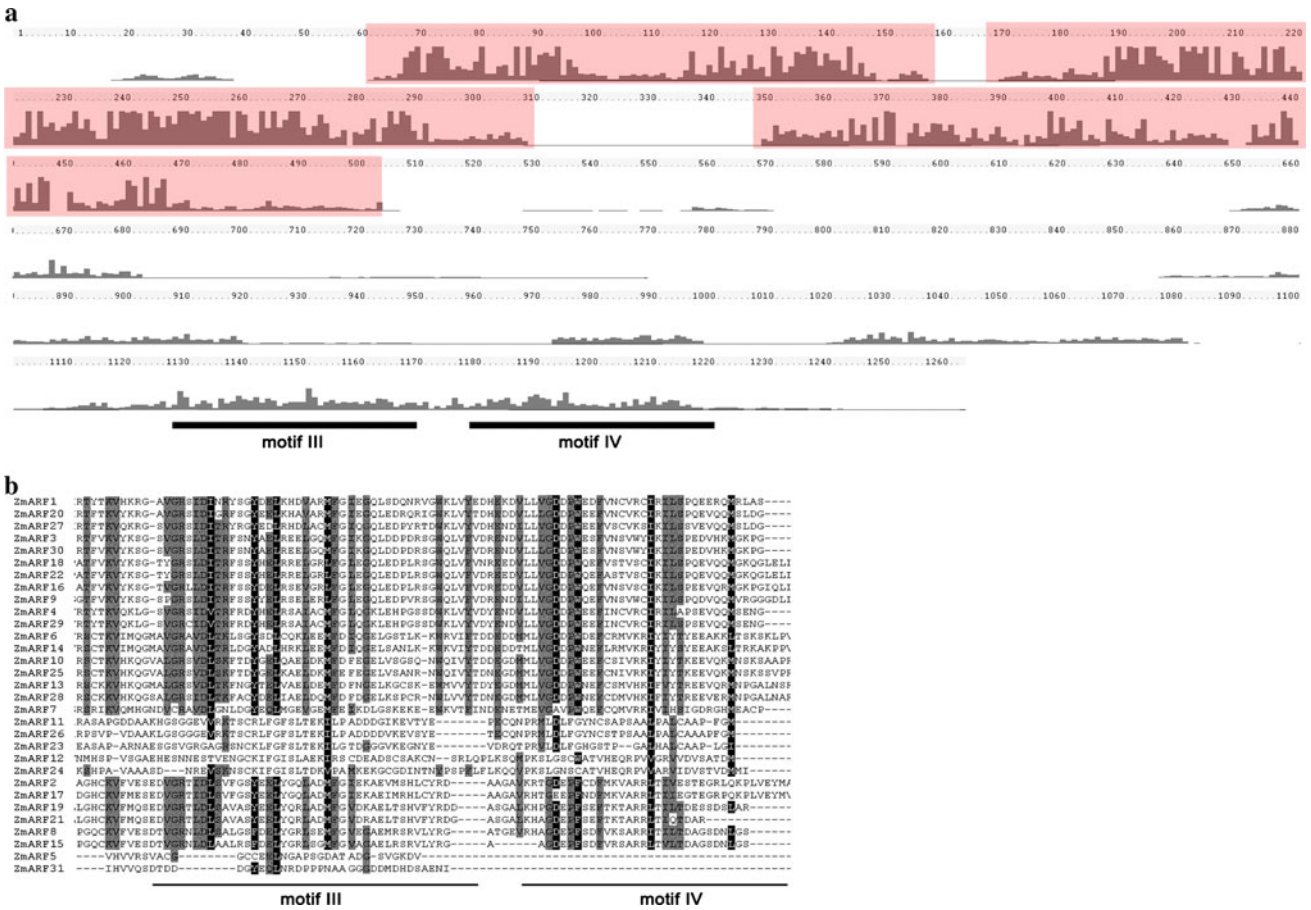

Fig. 3 **a** Alignment profile of maize ARF proteins obtained with the ClustalX program. The height of the *bars* indicates the number of identical residues per position. The *shaded regions* indicate the high sequence similarity among DBDs regions. Motifs III and IV are consensus sequences shared by Aux/IAA proteins. **b** Multiple alignments of Motifs III and IV of in maize ARF proteins obtained with ClustalX. Conserved residues are highlighted in *gray boxes*. Conserved domains are also *underlined* and correspond to part **a**

estimated total number of genes (tomato ~35,000, maize ~53,760, *Arabidopsis* ~25,000, rice ~37,000), so it is interesting to find roughly a similar number of ARF genes in these four different species.

Large-scale duplication of the tomato genome has been reported (Ku et al. 2000). It was suggested that tomato was likely a paleopolyploid (Hoeven et al. 2002) in which large-scale genome duplication occurred approximately 50–52 million years ago (Schlueter et al. 2004). The maize genome is replete with chromosomal duplications and repetitive sequences, the result of an ancient polyploid event that occurred over 11 million years ago (Gaut et al. 2000).

In this study, 7 sister pairs of *SlARFs* and 13 sister pairs of *ZmARFs* were identified by phylogenetic analysis. All sister pairs were compared with their corresponding chromosomal locations. Except for *SlARF13* and *SlARF13-1*, which were likely the products of alternative splicing of mRNA, none of these sister pairs were genetically linked, as was also observed in the *OsARFs* (Wang et al. 2007). Conversely, all closely linked *SlARF* and *ZmARF* loci, such as *SlARF7* and *SlARF6-1* on chromosome 7, *SlARF4* and *SlARF10* on chromosome 11, *ZmARF3* and *ZmARF4* on chromosome 2, *ZmARF2* and *ZmARF22* chromosome 6, *ZmARF23* and *ZmARF24* on chromosome 6, and finally *ZmARF29, ZmARF30,* and *ZmARF31* on chromosome 10, were not grouped in sister pairs. Similarly, none of the six sister pairs in *Arabidopsis* were genetically linked (Okushima et al. 2005). Sister pairs were located on chromosomes 12 and 7 (*SlARF6* and *6-1*), chromosomes 2 and 3 (*SlARF8* and *8-1*), chromosomes 7 and 5 (*SlARF19* and *19-1*), chromosomes 11 and 6 (*SlARF10* and *16*), chromosomes 2 and 11 (*SlARF3* and *4*), and chromosomes 3 and 12 (*SlARF2* and *11*). Only *SlARF13* and *SlARF13-1* were found on the same chromosome. Among the 12 *ZmARF* sister pairs, three were found on chromosomes 2 and 10 (ZmARF3 and 30, ZmARF4 and 29, ZmARF5 and 31), two on chromosomes 1 and 5 (SlARF1 and 20, SlARF2 and 17), and two on chromosomes 3 and 8 (SlARF10 and 25, SlARF11 and 26). The chromosomal

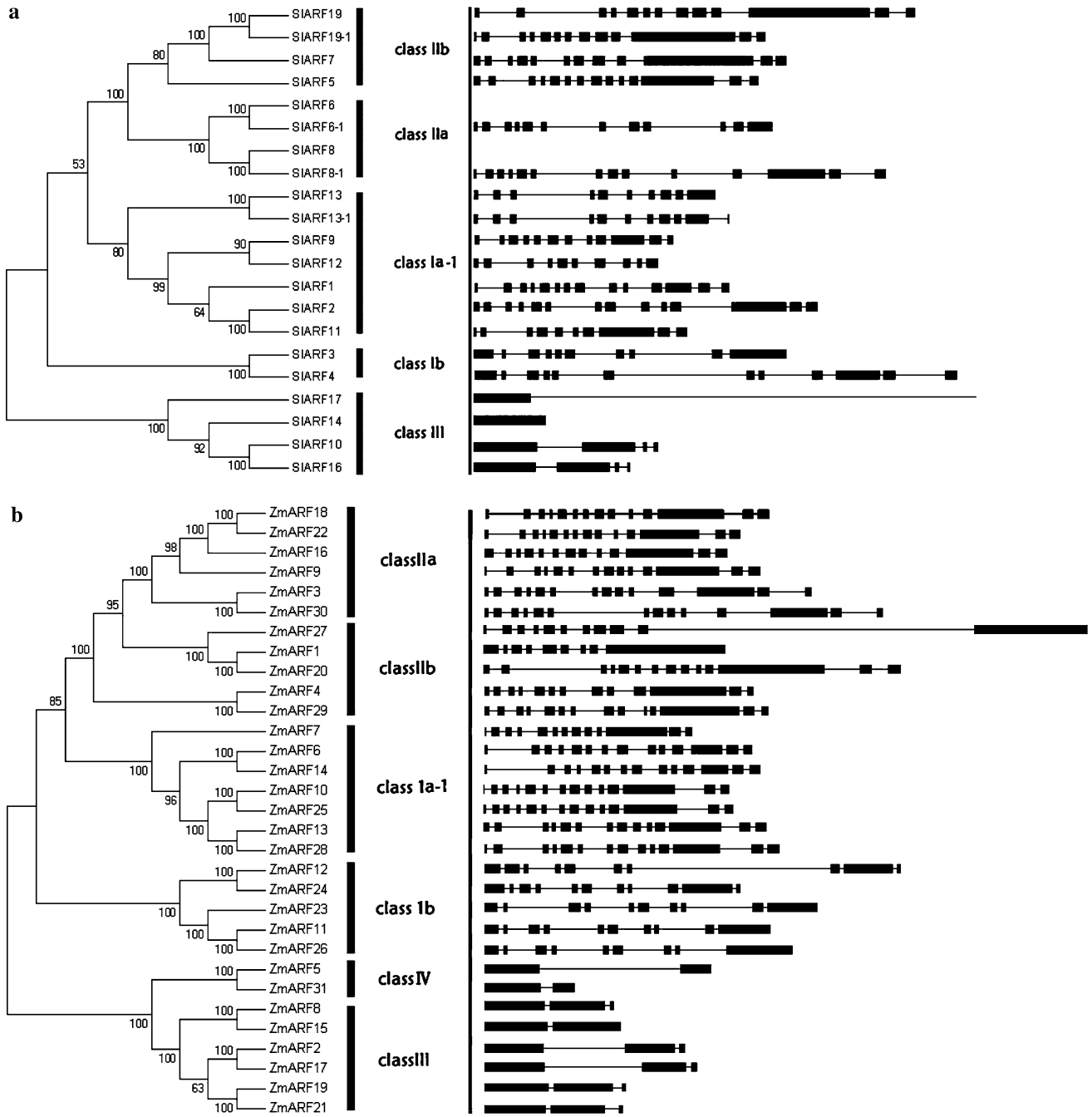

Fig. 4 **a** Left part illustrates the phylogenetic relationships among the tomato ARF proteins. The unrooted tree was generated using MEGA4.1 program by the neighbor-joining method. Bootstrap values (above 50%) from 1,000 replicates are indicated at each branch. *Right part* illustrates the exon–intron organization of corresponding ARF genes. The exons and introns are represented by *black boxes* and *lines*, respectively. **b** The same information for maize ARF proteins as shown in part **a**

locations of these SlARFs and ZmARFs sister pairs may represent duplicated chromosomal blocks.

Based on the above results, we conclude that whole genome and chromosomal segment duplications are the main factors responsible for the expansion of *SlARFs*, and especially *ZmARFs*. In *Arabidopsis* and rice, tandem duplications played a more important role in *AtARF* duplication (Wang et al. 2007), as evidenced by the fact that seven very closely related AtARFs (12, 13, 14, 15, 20, 21, and 23) in a single cluster are physically located near each other in a region of chromosome 1 in *Arabidopsis* (Remington et al. 2004; Okushima et al. 2005).

Phylogenetic analysis revealed that the organization of *Arabidopsis*, maize, tomato, and rice ARF proteins was

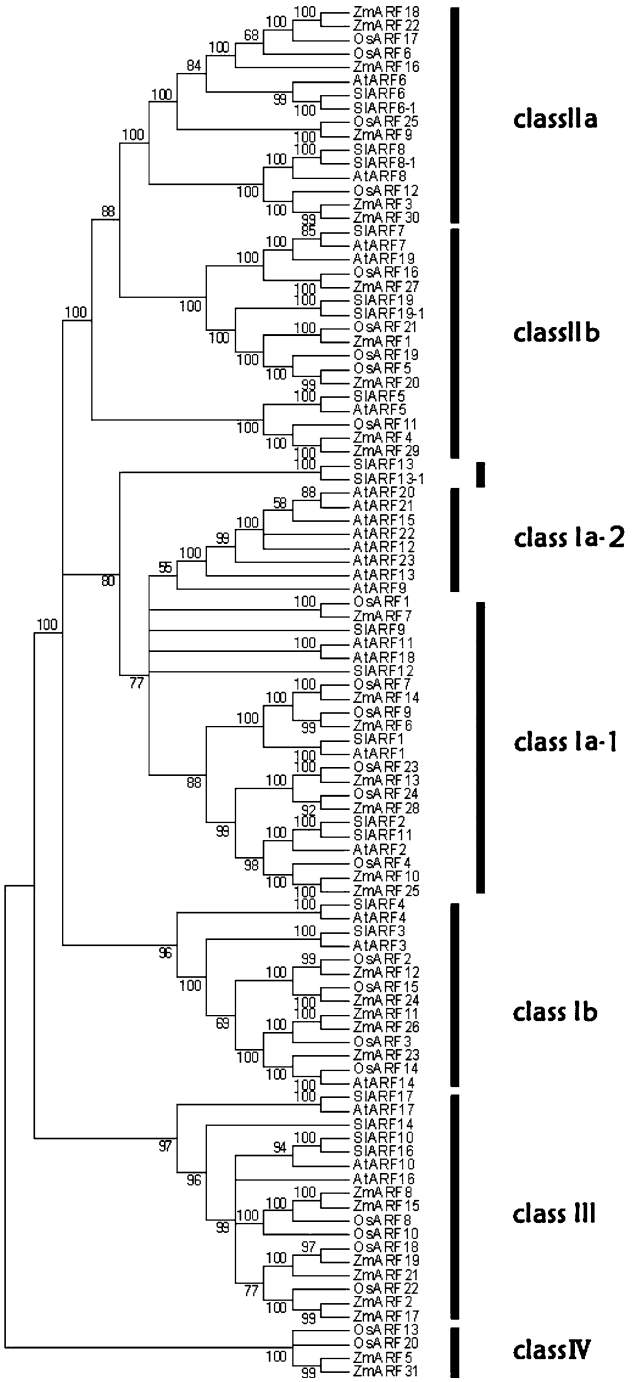

Fig. 5 Phylogenetic relationships among tomato, rice, maize, and Arabidopsis ARF proteins. The unrooted tree was generated using MEGA4.1 program by the neighbor-joining method. Bootstrap values (above 50%) from 1,000 replicates are indicated at each branch

very similar to each other in classes I, II, and III, implying that ARFs within these classes derived from a common ancestor. The 15 interspecies sister pairs, including OsARF25 and ZmARF9, SlARF1 and AtARF1, and OsARF14 and AtARF14, indicate these gene groups were descended from a common ancestor and possess well-conserved functions. The ten sister pairs between OsARFs and ZmARFs detected by phylogenetic analysis indicate that OsARFs and ZmARFs have a close evolutionary relationship, as may SlARFs and AtARFs with four interspecies sister pairs. In contrast, only one sister pair was found between monocots and dicots (OsARF14–AtARF14), indicating that the common ancestor of this sister pair appeared before the divergence of monocots and dicots. Remington et al. (2004) also suggested that ARF lineages originated before the monoco–eudicot divergence. The separation between dicot and monocot ARF sequences within each of the clades arose from this gene duplication, indicating that the ARF family expanded and diversified after the divergence between the two lineages. Most of the duplications in the *Arabidopsis* genome occurred shortly after the divergence between asterids (tomato) and rosids (*Arabidopsis*) 112–156 million years ago (Baumberger et al. 2003).

Class Ia-2 is a special subclass that only contained AtARFs, suggesting that these AtARFs were generated over the long-term evolution of *Arabidopsis,* but after the divergence of monocots and dicots. Moreover, segregation to a separate subclass suggests that these proteins have species-specific functions. The ARFs in class IV were all from maize and rice, the two representative monocots, suggesting that class IV proteins were either lost in dicots after divergence of monocots and dicots or evolved solely in monocots after the divergence.

Groups containing multiple ARFs from all four species, such as class IIa, were also found in the phylogenetic tree, which may indicate that a diversification of functions has occurred in all four species. Furthermore, six triplets containing one OsARF and multiple ZmARFs were found, while only one triplet containing multiple OsARFs and one ZmARF was found (OsARF5/OsARF19/ZmARF20). Groups containing one AtARF and multiple SlARFs were found in AtARF6/SlARF6/SlARF6-1, and groups containing one SlARF and multiple AtARFs were found in SlARF7/AtARF7/AtARF19. These classes are presumed to represent conserved functions in rice, tomato, maize, and Arabidopsis, but these functions might have begun to diversify in corresponding species as a result of gene duplication. Compared with OsARFs, the diversification of ZmARFs occurred more frequently, leading to a larger ARF family (31 members) than in rice (25 members).

The features and number of domains present in the protein sequences is very useful information for predicting the function of a new gene. Indeed, careful analysis of protein sequence is the first step when postulating the functions of novel ARF genes. The middle regions of ARFs function as activation domains (ADs) or repression domains (RDs) (Ulmasov et al. 1999a). Protoplast transfection assays indicated that AtARF1, AtARF2, AtARF4,

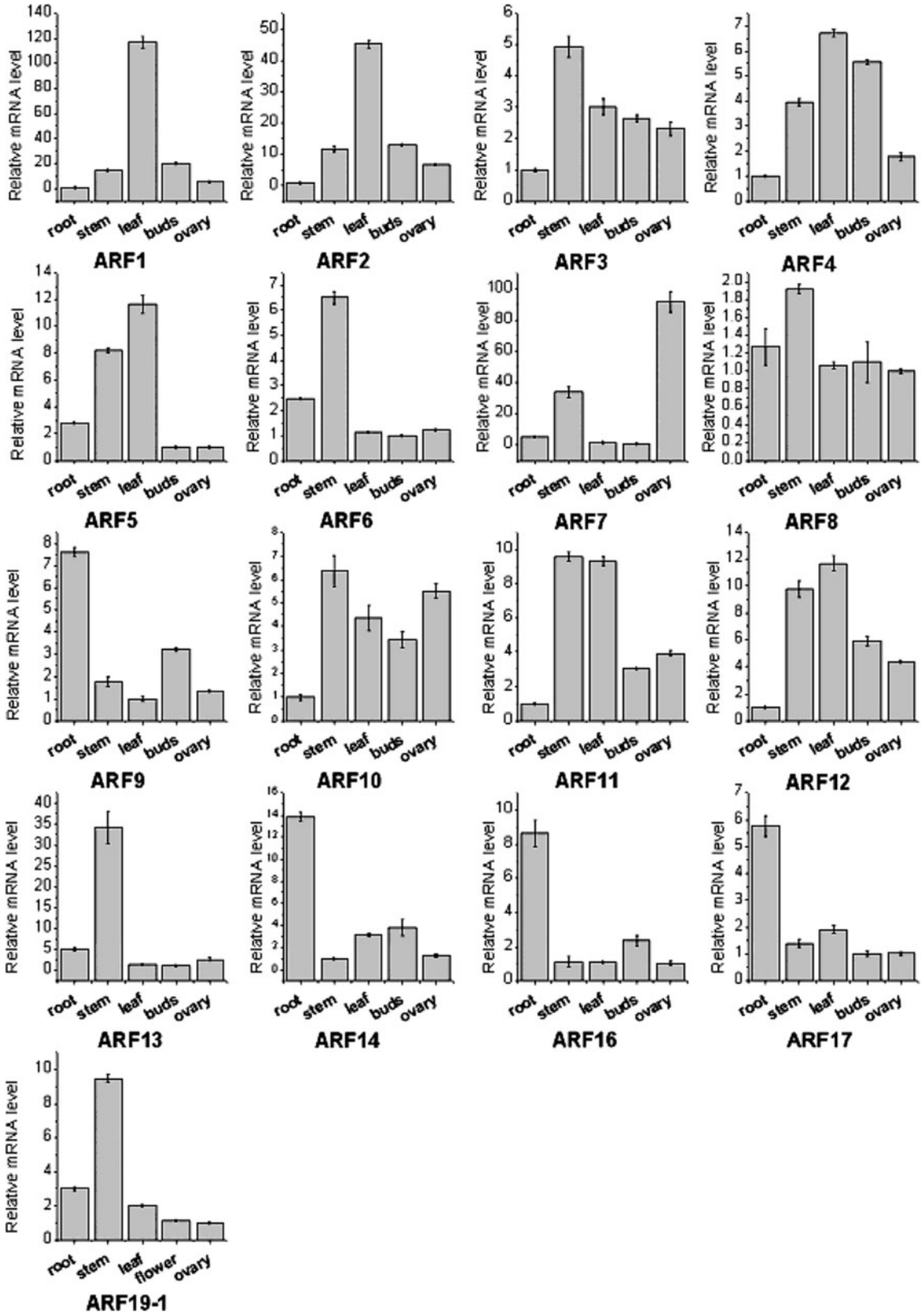

Fig. 6 qRT-PCR analyses of 17 *SlARF* genes in different organs (root, stem, leaf, buds, and ovary) of the tomato plant

and AtARF9, which contain middle regions rich in proline (P), serine (S) and threonine (T), are repressors, while AtARF5, AtARF6, AtARF7, and AtARF8, which contain middle regions rich in glutamine (Q), are activators (Tiwari et al. 2003; Ulmasov et al. 1999a). Interestingly, all four species encode some CTD-truncated ARFs, which is consistent with previous report that flowering plant tend to encode more CTD-truncated ARFs (Paponov et al. 2009). Compared with CTD-truncated ARFs of AtARFs (4 out of 23 ARFs) in Arabidopsis, more numbers of ARFs lacking a CTD were found in rice (6 out of 25 ARFs) and tomato (6 out of 21 ARFs), especially in maize (9 out of 31 ARFs). Kumar et al. (2011) also proved that tomato SlARF 2, 3, 6, 7, and 13 (5 out of 17 ARFs) showed absence of C-terminal Aux/IAA domains. So tomato, as a relatively advanced dicotyledon, has a higher percentage of CTD-truncated ARFs than Arabidopsis (28.6 and 17.4%, respectively). This situation is also been found in maize and rice as monocot crops (29.0 and 24%, respectively). The CTD is required for ARF-IAA dimerization, it seems that more CTD-truncated ARFs appeared during evolution, which may regulate gene expression in an auxin-independent manner (Shen et al. 2010). In previous studies, CTD-truncated ARFs are all putative repressors (Shen et al. 2010; Guilfoyle and Hagen 2007). ARF6-1, as a putative activator, also was found to be lack a CTD. So this gene might function in a different way, a more in-depth study is needed to further explain this phenomenon.

The protein sequences of all 21 SlARFs were analyzed, and the proline (P), serine (S), and threonine (T) rich regions were found in the MR domain sequences of *SlARF1*, *SlARF2*, *SlARF3*, *SlARF4*, *SlARF9*, *SlARF10*, *SlARF11*, *SlARF12*, *SlARF13*, *SlARF13-1*, *SlARF16*, *SlARF14*, and *SlARF17*, indicating these genes are more likely acting as repressors. In contrast, glutamine (Q)-rich regions, which are also somewhat rich in leucine (L) and serine (S), were found in the MR domain sequences of *SlARF5*, *SlARF6*, *SlARF6-1*, *SlARF7*, *SlARF8*, *SlARF8-1*, *SlARF19*, and *SlARF19-1*, implying that these genes are likely to be transcriptional activators (Supplementary Fig. 2, Supplementary Fig. 3). Although the MR of *SlARF5* was enriched in Q, it differed from other Q-rich ARFs such as *ARF6*, *ARF7,* and *ARF8* in having no homopolymeric Q stretches. Nevertheless, its MR has activation potency nearly equivalent to that of ARFs with homopolymeric Q stretches (Ulmasov et al. 1999a). All *SlARFs and ZmARFs* proteins with Q-rich MRs belong to class I, while PST-rich SlARFs and ZmARFs belong to other classes (Fig. 5).

Among 21 *SlARFs*, 17 *SlARF* genes were detected in all sampling tissues and organs in tomato, implying that most tomato *SlARF* genes exhibited constitutive expression. The mRNA levels of *SlARF1* and *SlARF2* were significantly higher in leaf than other organs (Fig. 6), implying that they, regarded as a repressor, might play an important role in the development of leaf, and this conjecture has been confirmed in Arabidopsis (Ellis et al. 2005; Lim et al. 2010). *AtARF2* functions in the auxin-mediated control of *Arabidopsis* leaf longevity by acting as a repressor of auxin signaling (Lim et al. 2010). The loss of expression of *AtARF7* genes is directly responsible for the reduced gene expression observed in mesophyll cells (Wang et al. 2005).

In *Arabidopsis*, *arf6/arf8* double-null mutant flowers were arrested as infertile closed buds with short petals, short stamen filaments, undehisced anthers that did not release pollen, and immature gynoecia (Nagpal et al. 2005). The *AtARF2* gene regulated floral organ abscission independently of the ethylene and cytokinin response pathways, and *AtARF1* was partially redundant with *ARF2* (Ellis et al. 2005). Similar expression patterns for several SlARF genes also indicate about their possible overlapping functions during various developmental processes in plants (Fig. 6, Supplementary Fig. 4, Supplementary Fig. 5, Supplementary Fig. 6).

In tomato, SlARF7 acts as a negative regulator of fruit set before pollination and fertilization and moderates the auxin response during fruit growth (Jong et al. 2009). The AtARF8 gene acts as an inhibitor to stop further carpel development in the absence of fertilization and the generation of signals required to initiate fruit and seed development (Goetz et al. 2006). DR12 (auxin response factor 4), can finely modify tomato fruit texture; when this gene was down-regulated, the pericarp tissue of tomato fruit became thicker (Guillon et al. 2008). In contrast to other genes, SlARF4 mRNA levels increased during fruit development, consistent with a previous study (Jones et al. 2002), implying that SlARF4 might be essential for fruit development. Other SlARFs may function in early fruit development, despite the difference in expression level. Kumar et al. (2011) proved that SlARF1 and 9 exhibited maximum expression level at open flower stage, and the expression level of SlARF16 was high at flower bud stage, indicating that these genes might be involved in flower development in tomato, while the mRNA levels of SlARF3, 5, 6, 13, 15, and 17 exhibited low expression during floral development and an increase at either 30 DAP (days after pollination) or mature green fruit stage, indicating that these ARF genes could be involved in the regulation of aspects of plant development. In sum, these studies indicate that some ARFs are indispensable for tomato flower and fruit development (Jones et al. 2002; de Jong et al. 2009).

In conclusion, the full cDNA sequences of 15 novel *SlARFs* were identified using PCR-based method. A comprehensive genome-wide analysis of *SlARF* gene family is presented, including the gene structures, chromosome

locations, phylogeny, and conserved motifs. The expression characteristics of all 17 *SlARFs* were also analyzed. The major challenge for the future is to define the specific functions of each individual ARF gene during plant growth and development.

Accession numbers: Sequence data from this article can be found in the GenBank/EMBL data libraries under the following Accession numbers: HM061154(SlARF1), HM19 248.1(SlARF5), HM187579.1(SlARF6-1), HM560979.1 (SlARF8-1), HM037250.1(SlARF9), HM143941.1(SlARF 10), HM143940.1(SlARF11), HM565127.1(SlARF12), HM565128.1(SlARF13), HM565129.1(SlARF13-1), HM56 5131.1(SlARF14), HM195247.1(SlARF16), HM456923 (SlARF17), HM130544.1(SlARF19), HM565130.1 (SlARF 19-1).

Acknowledgments This work was supported by the State Key Basic Research and Development Plan of China (2009CB119000), the National Natural Science Foundation of China (30771470 and 31071804) and the Zhejiang Province Research Project of China (2009C32025).

References

Alper Y, Milton YN Jr, Bernardo GF, Glaucia MS, Daniel J, John G, Erich G (2009) GRASSIUS: a platform for comparative regulatory genomics across the grasses. Plant Physiol 149:171–180

Alvarez JP, Pekker I, Goldshmidt A, Blum E, Amsellem Z, Eshed Y (2006) Endogenous and synthetic microRNA stimulate simultaneous, efficient, and localized regulation of multiple targets in diverse species. Plant Cell 18:1134–1151

Andrew HP, John EB, Remy B (2009) The Sorghum bicolor genome and the diversification of grasses. Nature 457:551–556

Babenko VN, Rogozin LG, Mekhedov SL, Koonin EV (2004) Prevalence of intron gain over intron loss in the evolution of paralogous gene families. Nucl Acids Res 32:3724–3733

Baumberger N, Doesseger B, Guyot R, Diet A, Parsons RL, Clark MA, Simmons MP, Bedinger P, Goff SA, Ringli C, Keller B (2003) Whole-genome comparison of leucine-rich repeat extensins in Arabidopsis and rice. A conserved family of cell wall proteins form a vegetative and a reproductive clade. Plant Physiol 131:1313–1326

Blanc G, Wolfe KH (2004) Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. Plant Cell 16:1679–1691

Brukhin V, Hernould M, Gonzalez N, Chevalier C, Mouras A (2003) Flower development schedule in tomato *Lycopersicon esculentum* cv Sweet Cherry Sex. Plant Reprod 15:311–320

de Jong M, Mariani C, Vriezen WH (2009) The role of auxin and gibberellin in tomato fruit set. J Exp Bot 60:1523–1532

Ellis CM, Nagpal P, Jeffery YC, Gretchen H, Thomas J, Jason RW (2005) AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. Development 132:4563–4574

Feng J, Wang K, Liu X, Chen S, Chen J (2009) The quantification of tomato microRNAs response to viral infection by stem-loop real-time RT-PCR. Gene 437:14–21

Gaut BS, D'Ennequin MLT, Peek AAS, Sawkins MC (2000) Maize as a model for the evolution of plant nuclear genomes. Proc Natl Acad Sci USA 97:7008–7015

Goetz M, Vivian-Smith A, Johnson SD, Koltunow AM (2006) AUXIN RESPONSE FACTOR8 is a negative regulator of fruit initiation in Arabidopsis. Plant Cell 18:1873–1886

Goetz M, Hooper LC, Johnson SD, Rodrigues JCM, Vivian-Smith A, Koltunow AM (2007) Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in *Arabidopsis* and tomato. Plant Physiol 145:351–366

Guilfoyle TJ, Hagen G (2007) Auxin response factor. Curr Opin Plant Biol 10:453–460

Guillon F, Philippe S, Bouchet B, Devaux MF, Frasse P, Jones B, Bouzayen M, Lahaye M (2008) Down-regulation of an auxin response factor in the tomato induces modification of fine pectin structure and tissue architecture. J Exp Bot 61:1419–1430

Hamann T, Benkova E, Baurle I, Kientz M, Jurgens G (2002) The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. Genes Dev 16:1610–1615

Hoeven RV, Ronning C, Giovannoni J, Martin G, Tanksley S, van der Hoeven R (2002) Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. Plant Cell 14:1441–1456

Jones B, Frasse P, Olmos E, Zegzouti H, Li ZG, Latché A, Pech JC, Bouzayen M (2002) Down regulation of DR12, an auxin response factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. Plant J 32:603–613

Jong MD, Wolters-Arts M, Feron R, Mariani C, Vriezen WH (2009) The *Solanum lycopersicum* auxin response factor 7 (SlARF7) regulates auxin signaling during tomato fruit set and development. Plant J 57:160–170

Kalluri UC, Difazio SP, Brunner AM, Tuskan GA (2007) Genome-wide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*. BMC Plant Biol 7:1–14

Ku HM, Vision T, Liu JP, Tanksley SD (2000) Comparing sequenced segments of the tomato and Arabidopsis genomes: large-scale duplication followed by selective gene loss creates a network of synteny. Proc Natl Acad Sci USA 97:9121–9126

Kumar R, Tyagi AK, Sharma AK (2011) Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. Mol Genet Genomics 285:245–260

Li J, Dai X, Zhao Y (2006) A role for auxin response factor 19 in auxin and ethylene signaling in *Arabidopsis*. Plant Physiol 140:899–908

Lim PO, Kim LeeIC, JY KimHJ, Ryu JS, Woo HR, Nam HG (2010) Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity. J Exp Bot 61:1419–1430

Liscum E, Reed JW (2002) Genetics of Aux/IAA and ARF action in plant growth and RF8 promote jasmonic acid production and flower maturation. Development 132:4107–4118

Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheretd H, Nussaumea L, Crespib MD, Maizel A (2010) miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. Plant Cell 22:1104–1117

Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, Guilfoyle TJ, Hagen G, Alonso JM, Cohen JD, Farmer EE, Ecker JR, Reed JW (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development 132:4107–4118

Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Lui A, Nguyen D, Onodera C, Quach H,

Smith A, Yu G, Theologis A (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19. Plant Cell 17:444–463

Okushima Y, Fukaki H, Onoda M, Theologis A, Tasaka M (2007) ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. Plant Cell 19:118–130

Paponov IA, Teale W, Lang D, Paponov M, Rski R, Rensing SA, Palme K (2009) The evolution of nuclear auxin signalling. BMC Evol Biol 9:126–141

Quint M, Gray WM (2006) Auxin signaling. Curr Opin Plant Biol 9:448–453

Remington DL, Vision TJ, Guilfoyle TJ, Reed JW (2004) Contrasting modes of diversification in the Aux/IAA and ARF gene families. Plant Physiol 135:1738–1752

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Schlueter JA, Dixon P, Granger C, Grant D, Clark L (2004) Mining EST databases to resolve evolutionary events in major crop species. Genome 47:868–876

Shen CJ, Wang SK, Bai YH, Wu YR, Zhang SN, Chen M, Guilfoyle TJ, Wu P, Qi YH (2010) Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa* L.). J Exp Bot 61:3971–3981

Tatematsu K, Kumagai S, Muto H, Sato A, Watahiki MK, Harper RM, Liscum E, Yamamoto KT (2004) MASSUGU2 encodes Aux/IAA19, an auxin-regulated protein that functions together with the transcriptional activator NPH4/ARF7 to regulate differential growth responses of hypocotyl and formation of lateral roots in *Arabidopsis thaliana*. Plant Cell 16:379–393

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 25:4876–4882

Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell 15:533–543

Ulmasov T, Hagen G, Guilfoyle TJ (1997) *ARF1*, a transcription factor that binds to auxin response elements. Science 276: 1865–1868

Ulmasov T, Hagen G, Guilfoyle TJ (1999a) Activation and repression of transcription by auxin. Proc Natl Acad Sci USA 96: 5844–5849

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

Vert G, Walcher CL, Chory J, Nemhauser JL (2008) Integration of auxin and brassinosteroid pathways by auxin response factor 2. Proc Natl Acad Sci USA 105:9829–9834

Wang SC, Tiwari SB, Hagen G, Guilfoyle TJ (2005) AUXIN RESPONSE FACTOR7 restores the expression of auxin-responsive genes in mutant Arabidopsis leaf mesophyll protoplasts. Plant Cell 17:1979–1993

Wang DK, Pei KM, Fu YP, Sun ZX, Li SJ, Liu H, Tang K, Han B, Tao YZ (2007) Genome-wide analysis of the auxin response factors (ARF) gene family in rice (*Oryza sativa*). Gene 394:13–14

Wei HB, Cui BM (2006) Research progresses on auxin response factors. J Integ Plant Biol 48:622–627

Weijers D, Schlereth A, Ehrismann JS, Schwank G, Kientz M, Jürgens G (2006) Auxin triggers transient local signaling for cell specification in Arabidopsis embryogenesis. Dev Cell 10:265–270

Wilmoth JC, Wang S, Tiwari SB, Joshi AD, Hagen G, Guilfoyle TJ, Alonso JM, Ecker JR, Reed JW (2005) NPH4/ARF7 and ARF19 promote leaf expansion and auxin induced lateral root formation. Plant J 43:118–130

Yoon EK, Yang JH, Lee WS (2010) Auxin and abscisic acid responses of auxin response factor 3 in Arabidopsis lateral root development. J Plant Biol 53:150–154