Available online at www.sciencedirect.com
ScienceDirect
Gene 394 (2007) 13–24
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Genome-wide analysis of the *auxin response factors* (ARF) gene family in rice (*Oryza sativa*)

Dekai Wang [a], Kemei Pei [b], Yaping Fu [c], Zongxiu Sun [c], Sujuan Li [a], Heqin Liu [a], Kan Tang [a], Bin Han [d,*], Yuezhi Tao [a,*]

[a] *Crop Molecular Breeding Center, the Institute of Crop and Nuclear Technology Utilization, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, P.R. China*
[b] *Department of Chemistry, Zhejiang Sci-Tech University, Hangzhou 310018, P.R. China*
[c] *State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou 310006, P.R. China*
[d] *National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 500 Caobao Road, Shanghai 200233, P.R. China*

Received 2 October 2006; received in revised form 14 January 2007; accepted 19 January 2007
Available online 26 January 2007
Received by I. King Jordan

Abstract

Auxin response factors (ARFs) are transcription factors that bind with specificity to TGTCTC-containing auxin response elements (AuxREs) found in promoters of primary/early auxin response genes and mediate responses to the plant hormone auxin. The *ARF* genes are represented by a large multigene family in plants. A comprehensive genome-wide analysis was carried out in this study to find all ARFs in *Arabidopsis* (*Arabidopsis thaliana*) and rice (*Oryza sativa* subsp. *japonica*), 23 and 25 *ARF* genes, named as *AtARF*s and *OsARF*s, were identified, respectively. Chromosomal locations of all *OsARF*s were presented and it was found that the duplication of *OsARF*s was associated with only the chromosomal block duplications but not local tandem duplications. A phylogenetic tree was generated from alignments of the full-length protein sequences of 25 *OsARF*s and 23 *AtARF*s to examine the phylogenetic relationships of rice and *Arabidopsis* ARF proteins. All 48 members of *ARF* gene families fell into three major classes, a total of 13 sister pairs, including 9 *OsARF–OsARF*, 2 *AtARF–AtARF* and 2 *AtARF–OsARF* sister pairs were formed, showing different orthologous relationships between *AtARF*s and *OsARF*s. EST analysis and RT-PCR assays demonstrated that 24 of all 25 *OsARF* genes were active and the transcript abundance of some *OsARF* genes was affected by auxin treatment or light- and dark-grown conditions. The outcome of the present study provides basic genomic information for the rice *ARF* gene family and will pave the way for elucidating the precise role of *OsARF*s in plant growth and development in the future.

Keywords: *ARF* gene; *Oryza sativa*; Phylogenetic analysis

1. Introduction

The plant hormone auxin, typified by indole-3-acetic acid (IAA), regulates a variety of physiological processes, including tropic responses, apical dominance, lateral root initiation, vascular differentiation, embryo patterning, and shoot elongation (Davies, 1995).

Auxin enhances the transcription of several classes of early genes, such as the *Aux/IAA*, *Gretchen Hagen3* (*GH3*), and *SMALL AUXIN UP RNA* (*SAUR*) gene family members (Abel and Theologis, 1996), it regulates plant physiology by modulating the interaction of transcription factors with auxin response elements

Abbreviations: ARF, auxin response factors; cDNA, DNA complementary to RNA; KDa, kilodaltons; mRNA, messenger RNA; RT, reverse transcriptase; PCR, polymerase chain reaction; ORF, open reading frame; AuxREs, auxin response elements; EST, expressed sequence tag; IAA, indole-3-acetic acid; DBD, DNA-binding domain; TAIR, The *Arabidopsis* Information Resource; KOME, Knowledge-Based *Oryza* Molecular Biological Encyclopedia; NCBI, National Centre for Biotechnology Information; TIGR, The Institute for Genomic Research; My, million years.

* Corresponding author. Tel./fax: +86 571 86404203 (Y. Tao), Tel./fax: +86 21 64845260 (B. Han).
E-mail addresses: Taoyz@zaas.org (Y. Tao), Bhan@ncgr.ac.cn (B. Han).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.01.006

(AuxREs) of the affected genes. A number of putative AuxREs have been defined within upstream promoter regions of primary/early auxin responsive genes, such as one or more copies of a conserved motif, TGTCTC, or some variation of this motif (e.g., TGTCCC, TGTCAC), which confer auxin responsiveness (Hagen et al., 1991; Li et al., 1994; Liu et al., 1994; Ulmasov et al., 1995).

ARF1 was the first auxin transcription factor gene cloned in *Arabidopsis*, using a yeast one-hybrid screen with a core motif, TGTCTC element, as bait sequence (Ulmasov et al., 1997a). The ARF1 protein contains 665 amino acids and has two unique domains, namely, a DNA-binding domain (DBD) in the N-terminal and a protein–protein interaction domain in the C-terminal portion (Ulmasov et al., 1997b). A typical ARF protein contains a DNA-binding domain (DBD) (always B3-like DNA-binding domain) in the N-terminal region, and domains III and IV are similar to those found in the C-terminus of Aux/IAAs, which is a protein–protein interaction domain that allows the homo- and hetero-dimerization of ARFs and the hetero-dimerization of ARF and Aux/IAA proteins (Ulmasov et al., 1997a,b, 1999a; Kim et al., 1997; Ouellet et al., 2001). Domains III and IV in some ARF proteins increase the *in vitro* binding (i.e., gel mobility shift assays) of these ARFs to TGTCTC AuxREs by facilitating the formation of dimers that occupy the two half-sites in palindromic AuxREs (Ulmasov et al., 1999a). The amino acid composition of the middle region between the DNA-binding domain and domains III/IV determines whether an ARF protein functions as an activator or repressor (Ulmasov et al., 1999b; Tiwari et al., 2003).

It has been reported recently that the ARF proteins are encoded by a large gene family with 23 members in *Arabidopsis* (Hagen and Guilfoyle, 2002; Liscum and Reed, 2002; Remington et al., 2004; Okushima et al., 2005). Several rice ARF family transcriptional regulators were also identified homologous to *Arabidopsis* ARF1 (Sato et al., 2001; Waller et al., 2002).

The availability of the *Arabidopsis* and rice genome sequences (Arabidopsis Genomic Initiative, 2000; Feng et al., 2002; Goff et al., 2002; Sasaki et al., 2002; Yu et al., 2002; International Rice Genome Sequencing Project, 2005) makes it possible to carry out the genome-wide comparative analysis of gene families between not only these two species but also monocot and dicot plants. The structure and phylogenetic analyses of auxin responsive gene families *Aux/IAA*, *GH3* and *SAUR* have been preformed in rice (Terol et al., 2006; Jain et al., 2006a,b,c). The rice *ARF* gene family (*OsARF* genes), however, has not been analyzed in detail, and the phylogenetic relationship with other plant *ARF* genes remains poorly understood.

In this work, an attempt was made to find out all *ARF* genes existed in both *Arabidopsis* and rice through a systematic EST and genomic DNA sequence data mining, resulting in the identifications of 23 and 25 *ARF* genes respectively. Followed by the complete survey of genomic organization, chromosomal location and sequence homology of all *OsARF*s, the different gene expression patterns with or without light and auxin treatment in various organs/tissues of each *OsARF*s and, phylogenetic relationship between *ARF* genes in *Arabidopsis* and rice were also investigated and compared.

2. Materials and methods

2.1. Searching for ARF genes

Searching of multiple databases was performed for finding all members of the *ARF* family in *Arabidopsis* and rice. The strategy to obtain every gene of the *ARF* family in a genome is the following.

For *ARF* genes in *Arabidopsis* (*AtARF*), the key words of "auxin responsive factor" were firstly used as a query to search against the TAIR (The *Arabidopsis* Information Resource) database (Huala et al., 2001), the sequences obtained were then used as queries to search against the TIGR (http://tigrblast.tigr.org/er-blast/index.cgi?project=ath1) databases using the BLASTP program at the e-value of 1e−3 to avoid false positives. The Pfam database (http://www.sanger.ac.uk/Software/Pfam/search.shtml) was finally used to confirm each predicted AtARF protein sequence as an auxin response factor protein.

For *ARF* genes in rice (*OsARF*), all the predicted AtARF protein sequences were firstly used as query sequences to search against TIGR database using BLASTP program. The significant hits were then used as query sequences to search against the National Centre for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/BLAST) using the TBLASTN program. The e-value used in the BLASTP was 1e−5. The BLASTP program available on TIGR (http://tigrblast.tigr.org/euk-blast/index.cgi?project=osa1) was also employed to survey the additional OsARF proteins using all the predicted AtARF protein sequence. The e-value used in the BLASTP was 1e−3.

Table 1
ARF genes in *Arabidopsis*

Name	Generic name	Chromosome	TIGR locus ID
AtARF19		1	At1g19220
AtARF5	*MP*	1	At1g19850
AtARF6		1	At1g30330
AtARF13		1	At1g34170
AtARF12		1	At1g34310
AtARF22		1	At1g34390
AtARF21		1	At1g34410
AtARF20		1	At1g35240
AtARF15		1	At1g35520
AtARF14		1	At1g35540
AtARF23[a]		1	At1g43950
AtARF1		1	At1g59750
AtARF17		1	At1g77850
AtARF10		2	At2g28350
AtARF3	*ETTIN*	2	At2g33860
AtARF11		2	At2g46530
AtARF18		3	At3g61830
AtARF9		4	At4g23980
AtARF16		4	At4g30080
ARF7	*NPH4*	5	At5g20730
AtARF8		5	At5g37020
AtARF4		5	At5g60450
AtARF2	*HSS*	5	At5g62010

[a] AtARF23 contains a stop codon within the DNA-Binding domain (DBD) and lacks any sequence carboxy-terminal to be the DBD. It is likely to be a pseudogene.

Table 2
ARF gene in rice

Gene name[a]	Generic name[b]	ORF length[c] (bp)	Deduced polypeptide[d]			Chromosome[e]	Genomic locus[f]			Nearest marker[g]	Fl-cDNA accession no[h]	EST frequency[i]
			Length (aa)	MW (kDa)	PI		BAC/PAC name	Accession no.	Centimorgan position			
OsARF1	OsARF16	2100	699	77.50	5.45	1	P0708G02	AP001539	38.8	C955	AK102521, AK064007	15
OsARF2	OsETTIN2	2169	722	79.11	7.17	1	OJ1125_C04	AP003706	114.1–116.5	C50101S	AK103776	28
OsARF3		2196	731	79.58	7.94	1	P0503C12	AP003268	130.1	C1335	AK072330	10
OsARF4	OsARF2	2427	808	90.85	6.48	1	OSJNBa0093F16	AP004332	161.5	S1046	AK072309	113
OsARF5		3282	1093	121.97	6.82	2	OJ1679_B08	AP005294	10.8	C11227S	NF	2
OsARF6	OsARF6a	2727	908	100.85	6.21	2	OSJNBa0023I17	AP004863	15	C53963S	AK070569, AK066725	25
OsARF7		2037	678	74.79	6.39	2	OSJNBb0038F20	AP005808	81.7–83.6	E3634S	AK070026	25
OsARF8		2046	696	73.84	7.03	2	B1469H02	AP006168	101.5	C920	NF	33
OsARF9		2022	673	74.75	5.78	4	OSJNBa0064D20	AL662973	65.3	E60696S	AK064925	9
OsARF10		2088	695	75.24	6.61	4	OSJNBb0061C13	AL731629	78.2	C60078S	AK100795	21
OsARF11	OsARF5 (MP)	2868	955	105.22	6.32	4	OSJNBa0084K01	AL606999	120.3	S12653S	AK103452, AK103105	13
OsARF12	OsARF8	2469	823	91.31	6.32	4	OSJNBb0004A17	AL606652	123.8	R2231	AK071455, AK067150	47
OsARF13		1590	529	57.22	5.57	4	OSJNBa0039K24	AL606637	128.2	R2184S	AK109449	1
OsARF14		2100	699	75.52	6.87	5	P0022D06	AC132485	104.7	E31112S	AK067927, AK058446	5
OsARF15	OsETTIN1	2139	712	78.08	7.09	5	OSJNBb0053D02	AC124143	115.4	R521	AK099793, AK062170	24
OsARF16		3168	1055	116.86	6.60	6	P0528E04	AP003510	19.1	L1092	AK103327	2
OsARF17		2754	917	101.78	6.21	6	B1153E06	AP004989	109.5	C11635S	AK103280	17
OsARF18	OsARF10	2103	700	75.88	7.57	6	P0623A10	AP005395	113.1	E3288S	AK100322	36
OsARF19	OsARF7a	3348	1115	123.65	6.59	6	OJ1215_E11	AP004324	117.0–119.2	C69	AK103312	11
OsARF20		2187	728	79.66	7.50	7	OJ1046_F10	AP003861	31.0–35.7	R2401	NF	NF
OsARF21	OsARF7b	3351	1116	123.72	6.78	8	OJ1003_A09	AP005509	106.1	C10122S	AK111639	7
OsARF22		2283	760	82.64	6.79	10	OSJNBa0093B11	AC024594	48.8	S11014	NF	15
OsARF23	OsARF1	2562	853	93.98	6.79	11	B1356E08	AC150702	77.4	E11749	AK065936, AK071997	129
OsARF24		2526	841	92.31	6.77	12	OSJNBa0005P03	AL935066	61.6	E4418S	AK067061, AK065254	42
OsARF25	OsARF6b	2700	899	99.86	6.29	12	OJ1327_A12	AL713940	103.1	L714	AK121703, AK065025	19

[a] Systematic designation given to rice ARFs in this work.
[b] Generic name designated in previous publications (Sato et al., 2001).
[c] Length of open reading frame in base pairs.
[d] Length (number of amino acids), molecular weight (kilodaltons), and isoelectric point (p*I*) of the deduced polypeptide.
[e] Chromosomal localization of the *OsARF* gene.
[f] Name, accession number, and approximate centimorgan position of the BAC/PAC clone in which the *OsARF* gene is present.
[g] Nearest marker to the *OsARF* gene.
[h] Corresponding full-length cDNA is available or not. NF (not found).
[i] Number of ESTs present in GenBank release 092206.

The Pfam database (http://www.sanger.ac.uk/Software/Pfam/search.shtml) was finally used to confirm each predicted OsARF protein sequence as an auxin response factor protein.

2.2. Mapping OsARFs on rice chromosomes

Marker-based physical maps of each japonica rice chromosome were downloaded from the International Rice Genome Sequencing Project (http://rgp.dna.avrc.go.jp/IRGSP/download). BACs or PACs containing *OsARF* genes were searched. All the sequenced contigs of japonica cv Nipponbare have been physically constructed as pseudomolecules by the IRGSP (http://rgp.dna.affrc.go.jp/IRGSP), representing the 12 rice chromosomes, and are available in GenBank (Accession Nos. AP008207–AP008218) (release 3.0). Each of the OsARFs was positioned on these rice chromosome pseudomolecules by the BLASTN search. The distinctive name for each of the OsARFs identified in this study is given according to its position from the top to the bottom on the rice chromosomes 1 to 12.

2.3. Sequence analysis

The DNA and protein sequences of *Arabidopsis* and rice auxin response factors obtained from the TIGR database were analyzed by DNAMAN (version 5.2.2) and DNASTAR

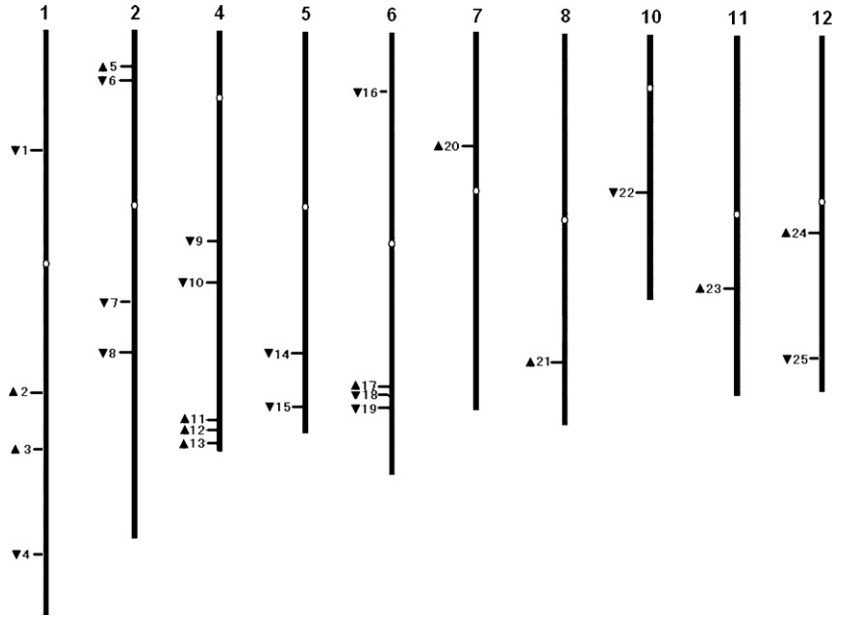

Fig. 1. Genomic distribution of ARF genes on rice chromosomes. White ovals on the chromosomes (vertical bar) indicate the position of centromeres. The arrows next to gene names show the direction of transcription. The chromosome numbers are indicated at the top of each bar.

programs. Multiple-sequence alignments were carried out using the ClustalX (version 1.83) program (Thompson et al., 1997).

The full-length cDNA (FL-cDNA) sequences of *OsARFs* were searched at the Knowledge-Based Oryza Molecular Biological Encyclopedia (KOME; http://cdna01.dna.affrc.go.jp/cDNA) (Kikuchi et al., 2003) by BLASTN search.

The EST analysis was performed using MEGABLAST tool against the EST database of rice available at the NCBI. Only top hits of the BLASTN search results for each *OsARF* gene showing a bit score of 500 or more were considered to be significant.

2.4. Phylogenetic analysis

Phylogenetic analysis was performed with MEGA3.1 program (Kumar et al., 2004) by neighbor-joining method (Saitou and Nei, 1987) and the bootstrap test was carried out with 1000 iterations.

2.5. Plant materials

Rice (*Oryza sativa* L subsp. Japonica var. Nipponbare) seeds were treated and grown as described (Terol et al., 2006). Briefly, seeds were disinfected with 0.1% $HgCl_2$ for 1 h and thorough washing with reverse osmosis (RO) water; soaked overnight in RO water to germinate and grown in Petri dishes for 10 days at 25 °C on filter paper moistened with water, under the conditions of either 12 h light or complete darkness. For auxin treatment, seedlings grown in light were incubated for 30 min in either water or 10 μM IAA solution. Harvested seedlings were frozen in liquid nitrogen immediately and stored at −70 °C until RNA isolation.

2.6. RT-PCR analysis

Total RNA was extracted from seedlings by using the Trizol reagent (Invitrogen, Germany) according to the manufacturer's instructions. The first cDNA strand was generated by reverse transcribing 5 μg of total RNA (50 μl reaction volume) using AMV reverse transcriptase (Takara Biotechnology, Japan) at 42 °C for 1 h.

All gene-specific primers were designed based on the sequences of the C-terminal regions of ARF proteins (Supplementary Table S1). The specific primer for the rice *Actin-1* gene was used as an internal control. Reactions were performed with Taq Polymerase (Takara Biotechnology, Japan) on a Peltier Thermal Cycler (PCT-200, MJ Research, USA), with the following profile: 5 min at 94 °C to start, followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 1 min at 72 °C, and a final extension period of 72 °C for 10 min to complete the reaction. Each PCR pattern was verified by triple replicate experiments, mixtures without template were employed as negative control and *OsActin1* DNA fragment as positive control for each gene investigated.

3. Results

3.1. Identification of OsARFs

From the first searching on the TAIR (The *Arabidopsis* Information Resource) database, 48 significant gene models that represent 25 loci annotated as "auxin responsive factor" were obtained. Using each of the then predicted 25 protein sequences as a query to search against the TAIR and TIGR databases with the BLASTP program resulted in a number of significant hits representing 31 *ARF* gene models or 25 loci. A total of 23 loci or 29 gene models were eventually confirmed as

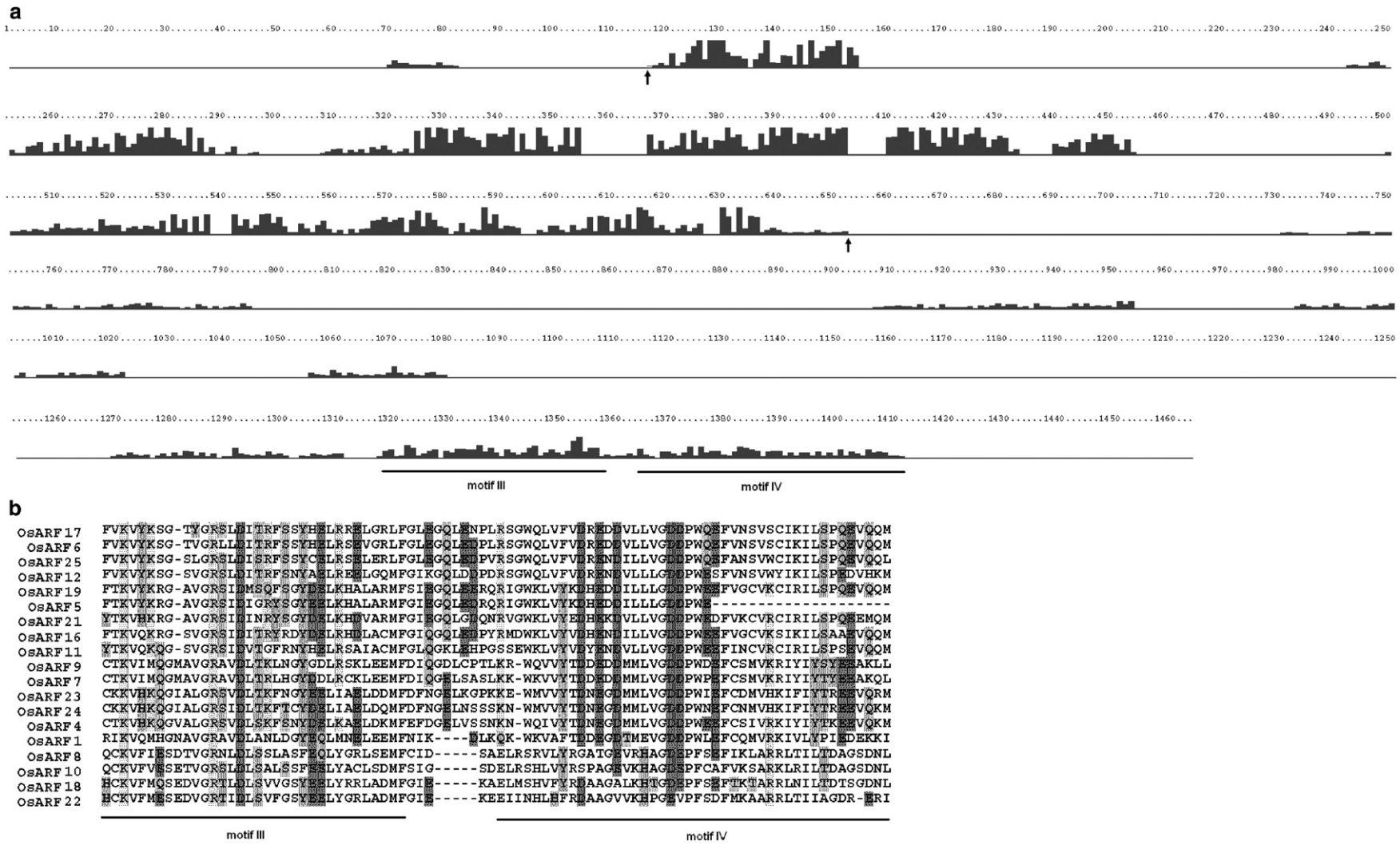

Fig. 2. a: Alignment profile of rice ARF proteins obtained with the ClustalX program. The height of the bars indicates the number of identical residues per position. The arrows indicate the region shows the core region with high-sequence similarity among DBDs regions; Motifs III and IV are consensus sequences shared by Aux/IAA proteins. b: Multiple alignments of the Motifs III and IV of the rice ARF proteins obtained with ClustalX. Conserved residues (present in more than 50% of aligned sequences) are highlighted in gray boxes. Amino acids considered as conserved are: K and V; G and R; V, L, P and Y. Gaps (marked with dashes) have been introduced to maximize the alignments. Conserved domains are underlined and indicated. The respective amino acid position is given on the left and right of each sequence.

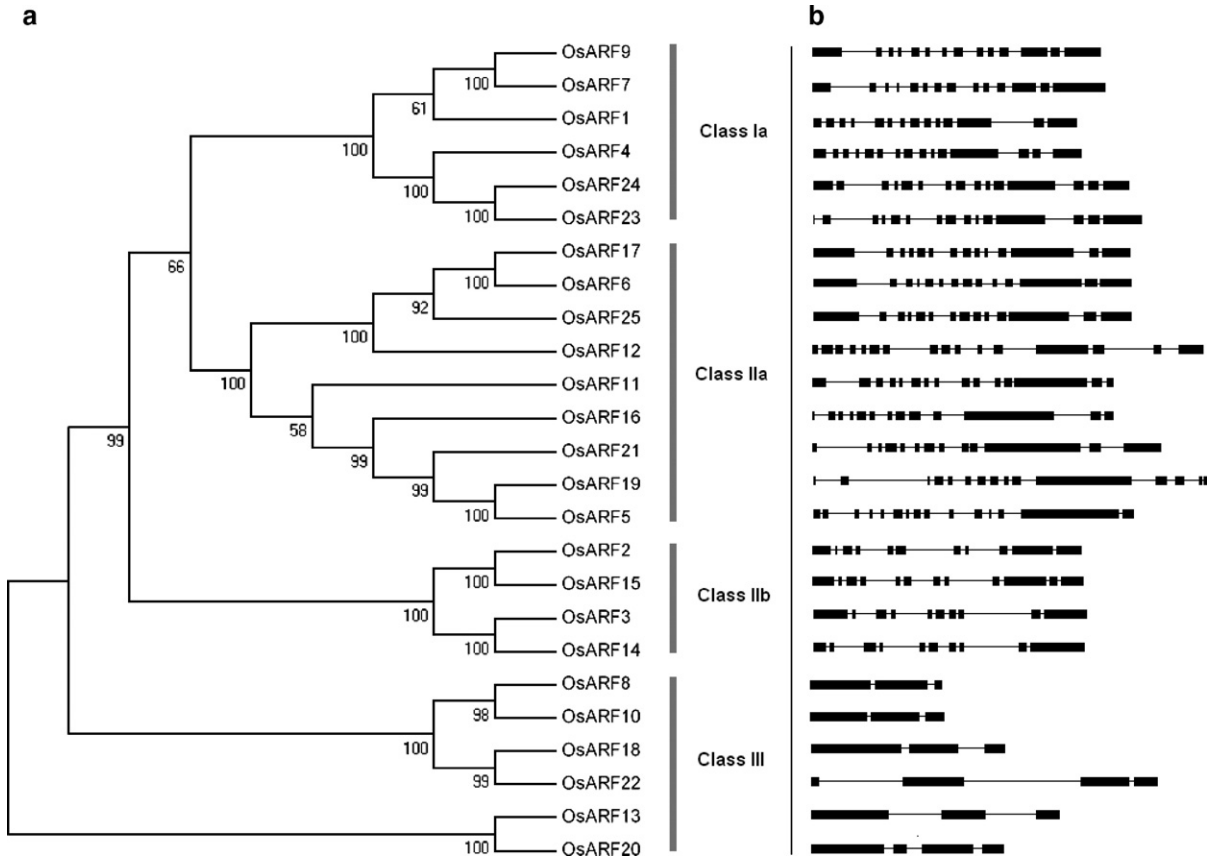

Fig. 3. a: Phylogenetic relationship among the rice ARF proteins. The unrooted tree was generated using MEGA3.1 program by neighbor-joining method. Bootstrap values (above 50%) from 1000 replicates are indicated at each branch. b: Exon–intron organization of corresponding ARF genes. The exons and introns are represented by black boxes and lines, respectively.

ARF genes in *Arabidopsis* by using the BLAST program available from Pfam database and all of these *AtARF*s were listed in Table 1. For sequence and phylogenetic analyses, only the longest frame was performed.

All 23 predicted AtARFs protein sequences were used as queries in the BLASTP program to search against TIGR rice annotated database to identify *OsARFs* and a total of 62 significant hits representing 47 rice loci were identified. By using the BLAST program available from the Pfam database, the total predicted 25 loci or 40 gene models were identified as *ARF* genes in rice. The survey for finding the *ARF* genes in rice was finally completed using 25 predicted rice protein sequences as queries. The overall analysis of the complete genome of rice revealed that the *OsARF* gene family is comprised of 25 members. These members of the *OsARF* gene family are summarized in Table 2.

3.2. Chromosomal locations of OsARFs

The BAC or PAC clones carrying the *OsARFs* were identified based on information from the rice chromosome pseudomolecules available at TIGR (Table 2). The approximate chromosome map positions of BACs/PACs in centimorgans from the top of the chromosome and their nearest markers obtained from the rice physical map are listed in Table 2. The chromosomal locations and directions of transcription of each *ARF* gene were, therefore, determined and demonstrated (Fig. 1; Supplementary Table S2).

The family of 25 *OsARF*s is distributed on 10 of the 12 rice chromosomes; no *OsARF*s are found to be located on chromosomes 3 and 9. Five *OsARF*s are presented on chromosome 4; 4 each on chromosomes 1, 2, and 6; 2 each on chromosomes 5 and 12 and, only one *OsARF* each on chromosomes 7, 8, 10 and 12 (Table 2; Fig. 1).

It is noteworthy that the nomenclature system for *OsARF*s used in the present study, a generic name from *OsARF1* to *OsARF25*, was provisionally given to distinguish each of the *ARF* genes according to its position from the top to the bottom on the rice chromosomes 1 to 12, which is different from the one previously used (Sato et al., 2001). Aside from providing a unique identifier for each member of a given gene family, this numbering system has been broadly used in genome-wide studies for the *GH3* and *Aux/IAA* gene families in rice and *ERF* in *Arabidopsis* and rice (Jain et al., 2006a,b,c; Terol et al., 2006; Nakano et al., 2006).

3.3. Sequence analysis of the OsARF proteins

Twenty-four putative OsARF protein sequences were found to have a typical DBDs domain as indicated by the outcomes of

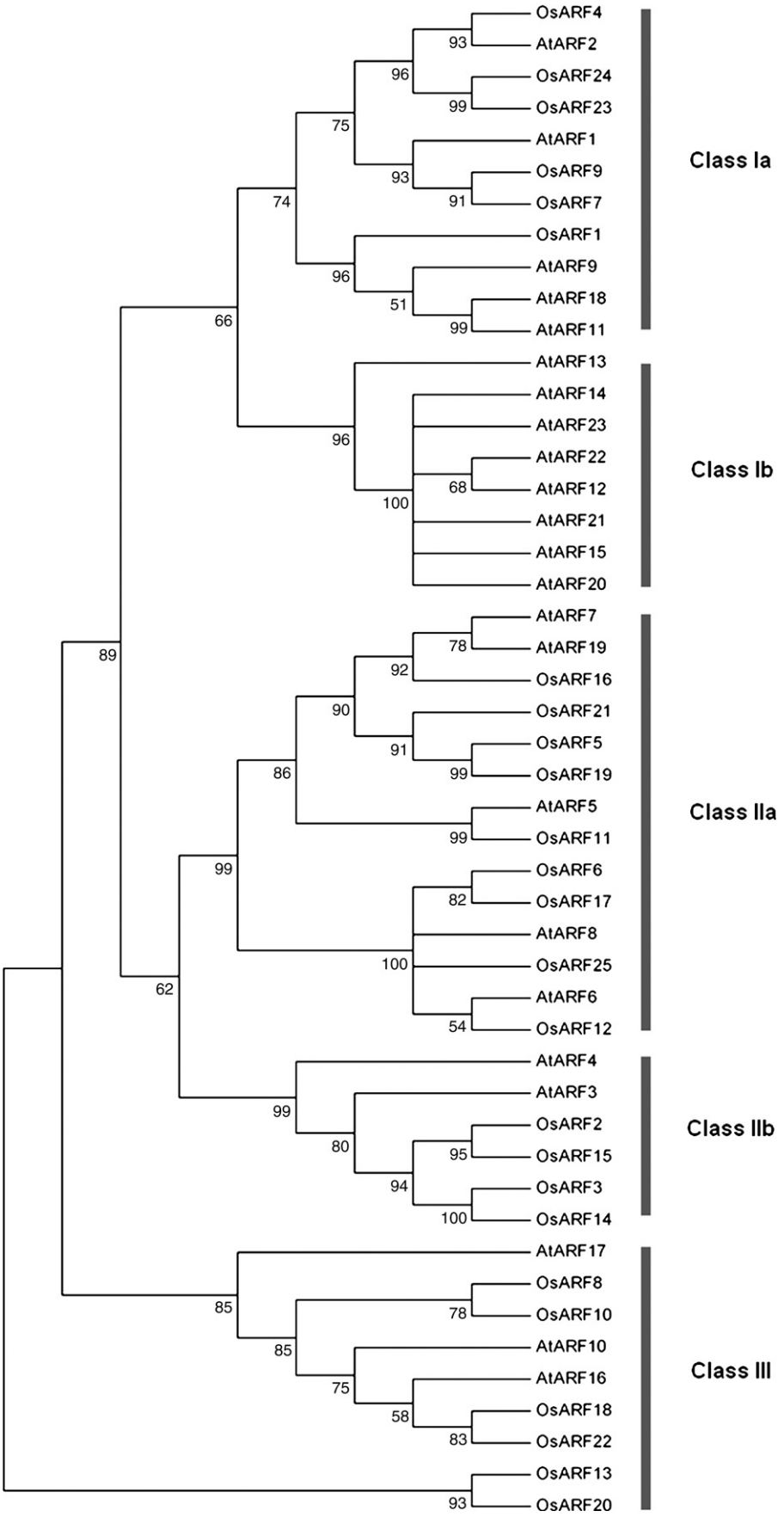

Fig. 4. Phylogenetic relationship of rice and *Arabidopsis* ARF proteins. The unrooted tree was generated using MEGA3.1 program by neighbor-joining method. Bootstrap values (above 50%) from 1000 replicates are indicated at each branch. For the AGI gene names of *Arabidopsis* ARF proteins, see Table 1.

Pfam analysis. OsARF20, the only exception among all 25 OsARF genes identified, has two DBDs domains. The amino acid sequences of these two domains were 31% similar, with 74 identical amino acids among the 230 amino acids of the basic region (data not showed).

The molecular weights of these deduced OsARF protein sequences generally ranged from 57 kDa (OsARF13) to 123 kDa (OsARF19 and OsARF21) (Table 2).

The multiple protein sequence alignments were performed using the ClustalX program to examine sequence features of these rice ARF domains (Supplementary Fig. 1). A pair-wise analysis of the full-length OsARF protein sequences indicated that the overall identities ranged from 8% (between OsARF5 and OsARF14) to 57.3% (between OsARF 8 and OsARF10) (Supplementary Table S3) and, the core region is highly conserved among these OsARFs (Fig. 2a).

All of the OsARF proteins contain a highly conserved region of about 320 amino acid residues in their N-terminal portion, which corresponds to the DNA-binding domain (DBD) of the *Arabidopsis* ARF family (Fig. 2a). Most of the OsARF proteins except OsARF2, 3, 13, 14, 15 and 20 contain a carboxy-terminal domain related to domains III and IV found in Aux/IAA proteins (Fig. 2b; Supplementary Fig. 1).

3.4. Gene structure and phylogenetic analysis of ARFs

A comparison of the full-length cDNA sequences with the corresponding genomic DNA sequences was made to determine the numbers and positions of exons and introns of each individual *OsARF* genes. It was showed that the coding sequences of all the ARF genes are disrupted by introns, the number of introns varied from 2 to 14 (Fig. 3b).

From the unrooted phylogenetic tree generated from alignments of the full-length protein sequences of all the OsARFs, it was showed that all the OsARFs fell broadly into three major classes: classes I, II and III, with well-supported bootstrap values (Fig. 3a). Class I contained 6 members. Class II was further subdivided into two subclasses, IIa (nine members) and IIb (four members). Class III contained six members that were the most divergent compared with those encoded by the other two classes. Twenty-five of the OsARFs formed 9 sister pairs, all of them had very strong bootstrap support (>99%).

To examine the phylogenetic relationships of rice and *Arabidopsis* ARF proteins, a phylogenetic tree was constructed from alignments of the full-length protein sequences of 25 OsARFs and 23 AtARFs (Fig. 4). All 48 members of the ARF gene families were also grouped into three major classes: Classes I, II and III, containing 19, 20 and 9 gene members, respectively. Both Classes I and II were further divided into two subgroups each, named as Class Ia and Class Ib and, Class IIa and Class IIb, respectively. This classification is very similar to the one of AtARFs (Supplementary Fig. 2).

A total of 13 sister pairs, including 9 OsARF–OsARF, 2 AtARF–AtARF and 2 AtARF–OsARF sister pairs were formed in the joint phylogenetic tree, subgroups Ia, IIa, IIb and III contained both AtARFs and OsARFs, but subgroup Ib contained only AtARFs.

3.5. Expression of OsARF genes

Twenty-two *OsARF*s were found to have corresponding FL-cDNA sequences from the KOME, indicating that majority of *ARF* genes are expressed in rice. A MEGABLAST search in EST database available at NCBI was performed and resulted in the identification of ESTs for most of the *OsARF* genes, but the frequency of ESTs for individual genes varies greatly (Table 2). For example, only one EST is deposited for *OsARF13* while

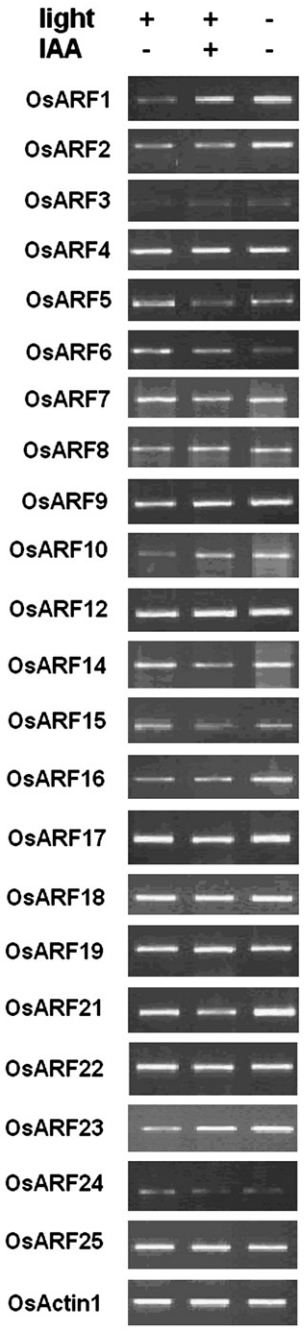

Fig. 5. RT-PCR analysis of OsARF gene family. Total RNA was extracted from ten-day-old seedlings germinated on wet filter paper soaked with distilled water and grown in either white light or darkness. Seedlings grown in light were further incubated for 30 min in either fresh water or IAA 10 μM. An amplified *OsActin1* gene was used as an internal control.

113 for *OsARF4* (Table 2). The transcription of *OsARF20* genes, however, remained unclear, as no corresponding full-length cDNA or EST could be found from all the databases available.

The data from RT-PCR analyses on RNA isolated from ten-day old seedlings grown in fresh water under either light (incubated in either water or not with 10 μM IAA for 30 min) or darkness showed that 24 of the predicted *OsARF*s displayed transcriptional activity, no expression was detected for *OsARF20* under any conditions tested in this work (Fig. 5).

Under light, auxin treatment enhanced transcription for *OsARF1* and *23*, while expression was decreased for *OsARF5*, *14* and *21*. Increased expressions were found for *OsARF1*, *2*, *16*, *21* and *23* in dark condition compared with light-grown plants. The evidence of the presence of associated ESTs and gene expression data as detected above indicated that a total of 24 out of all 25 predicted *OsARF* genes are more likely acting as the active genes.

To determine the organ-specific expression or in response to exogenous auxin and light stimuli, RT-PCRs were also performed with total RNA isolated from root, culm, leaf and young panicle. This analysis revealed that 24 of the total 25 *OsARF* genes are expressed in selected tissue/organs and the level of their mRNA was not altered (Supplementary Fig. 3). The results showed that rice *OsARF* genes are constitutive expressed, implying that there are functional redundancies among the ARF proteins.

It has been reported that *AtARF1–ARF10* mRNAs are ubiquitously expressed in most major organs of *Arabidopsis* plants, as well as in *Arabidopsis* suspension culture cells and exogenous auxin treatment (i.e. 50 μM NAA, 5 h) did not alter their mRNA levels (Ulmasov et al., 1999b). *OsARF1* (re-designated as *OsARF23* in this work) was identified and classified as a primary auxin responsive gene because its mRNA was up-regulated as treated with exogenous auxin. There is evidence that the *Arabidopsis ARF4* and *ARF19* gene expressions are induced by auxin (Okushima et al., 2005; Overvoorde et al., 2005).

The results of RT-PCR in this study demonstrated that some *ARF* expressions were induced by auxin, which means these genes, just as being named, function as *ARF* genes.

4. Discussion

4.1. Duplications of OsARFs and AtARFs

Twenty-five rice *ARF* genes were identified in this study, the number of members of this gene family is very similar to the one (23) of *Arabidopsis*. It gives rise to the questions of, firstly, where these multiple gene members originally came from and, secondly, why a similar number of *ARF* genes existed in the two different species, given that the rice genome size is about 3.7 times (450 Mb:130 Mb) and the gene number is about 1.5 times (37,000:25,000) of those of *Arabidopsis*, respectively.

It is believed that multiple members of a specific gene family of a particular organism are the natural products generated from the long evolutionary history the organism experienced. The number of members of a gene family reflects a succession of genomic rearrangements and expansions due to extensive duplication and diversification that frequently occurred in the course of evolution.

Several rounds of whole genome duplication have been found in both the *Arabidopsis* and rice genomes reported by previous studies (Lynch and Conery, 2000; Simillion et al., 2002; Raes et al., 2003; Wang et al., 2005). Two rounds of large-scale genome duplications have been taken place in most of the rice chromosomes, one was predicted around 40–50 My ago and another before the monocot–dicot divergence at about 120–150 My ago (Goff et al., 2002; Yu et al., 2005). Recently, 10 duplicated blocks accounting for 45% of rice genome sequences have been identified, indicating that rice is an ancient polyploidy (Wang et al., 2005).

Although the whole genome and segmental duplication events are logically the contributors of increasing the *ARF* gene numbers in rice and in *Arabidopsis* after their divergence, a recent series of tandem duplication was also reported as another force contributing to increase the number of AtARFs, as evidenced by the fact that 7 out of 8 very closely related AtARFs (12, 13, 14, 15, 20, 21, 23 and 23) in a single cluster are physically located near each other in a region proximal to the centromere of chromosome 1 in *Arabidopsis* (Hagen and Guilfoyle, 2002; Remington et al., 2004; Okushima et al., 2005).

From the outcome of the phylogenetic analysis in this study, nine sister pairs of OsARFs were formed but none of these pairs were genetically linked to each other as compared with their corresponding chromosomal locations. On the other hand, all closely linked OsARF loci such as OsARF 5, 6 on chromosome 2; OsARF 11, 12, 13 on chromosome 4 and OsARF 17, 18 and 19 on chromosome 6 are not paired together into the same sister groups.

Among the nine OsARF sister pairs, two in each were found occurring between chromosomes 1 and 5 (OsARF2 and 15, OsARF3 and 14), chromosomes 2 and 4 (OsARF7 and 9, OsARF8 and 10) and, chromosomes 2 and 6 (OsARF6 and 17, OsARF5 and OsARF19), one pair between chromosomes 11 and 12 (OsARF23 and 24). The chromosomal locations of these 7 sister pairs are the same duplicated chromosomal blocks as described previously by Paterson et al. (2004) and Yu et al. (2005). The sister pairs of OsARF18–22 and 13–20 may represent unidentified duplicated chromosomal blocks.

From the results mentioned above, it is clear that the contributors to the expansion of ARF genes are different in rice and *Arabidopsis*. Even though the whole genome and chromosomal segment duplications are mainly contributed to the expansion of both OsARF and AtARF, the tandem duplication, another possible contributor that occurred in *Arabidopsis*, was not found in rice.

The more rapidly evolved gene groups associated with tandem duplication resulted in a rather larger number of members of the gene family. Presuming the 7 duplicated AtARFs are originally derived from one AtARF gene, the total number of AtARFs is hence reduced from 23 to 16, the ratio of AtARFs (16) and OsARF (25) is, therefore, changed from nearly 1:1 (23:25) to

about 2:3 (16:25), which is much closer to the ratio of the number of genes annotated in *Arabidopsis* and rice.

4.2. Phylogenetic relationship between AtARF and OsARFs

Based on phylogenetic analysis, the organization of rice ARF proteins were very similar to that of the *Arabidopsis* ARF proteins, implying that rice and *Arabidopsis* ARF proteins were derived from a common ancestor. The fact that all three classes were formed by both *Arabidopsis* and rice *ARF* genes gave a suggestion that these *ARF* genes were existed before the divergence of monocots and dicots.

The fact that there were no OsARFs in class Ib suggested that class Ib ARF proteins were either lost in rice after divergence of monocots and dicots or evolved in dicots after the divergence from monocots. If the latter case is true, the *AtARF* genes in this group are more likely to have dicot-specific functions.

Three different kinds of grouping of both *AtARF*s and *OsARF*s were observed according to the outcome of phylogenetic analysis in this study, showing different orthologous relationships of *ARF* genes in these two species.

The 1:1 orthologous relationship was found between *AtARF2* and *OsARF24*, *AtARF5* and *OsARF11*, respectively, indicating these gene groups were descended from a common ancestor and corresponded to well-conserved functions. This relationship has been well proved by previous reports. *AtARF2* and *OsARF24*, *AtARF5* and *OsARF1* have been functionally characterized and mutations in *ARF5/MP* interfere with the formation of vascular strands and the initiation of the body axis in the early embryo (Hardtke and Berleth, 1998).

The groups containing single *OsARF* and multiple *AtARFs* (1:n) were found in *OsARF16/AtARF7/AtARF19*, while the groups containing multiple *OsARF* and single *AtARFs* were found in *OsARF7/OsARF9/AtARF1*. These 1:n and n:1 orthologue classes are presumed to represent conserved functions in rice and *Arabidopsis*, but these functions might have begun to diversify in one species as a result of gene duplication.

The n:n orthologue classes were also found from the phylogenetic tree. These might represent functions that have begun to diversify in both species.

4.3. Putative functions of ARFs

Since detailed information about the feature and number of domains imprinted in the protein sequences is always helpful for predicting functions of unknown genes, taking a close look of the protein sequences identified should be the first attempt to get useful clues about the functions of the corresponding *ARF* genes.

It has been reported that the middle regions of ARFs function as activation or repression domains. Transfection assays with protoplasts indicate that AtARF1 and 2, which contain middle regions rich in proline (P), serine (S) and threonine (T), are repressors and AtARF5, -6, -7, and -8, which contain middle regions rich in glutamine (Q), are activators (Ulmasov et al., 1999b).

The detailed sequence analysis of all 25 OsARF proteins revealed that proline (P), serine (S) and threonine (T) rich regions were found in the middle regions of protein sequences of OsARF1–4, OsARF7–10, OsARF13–15, OsARF18, OsARF20 and OsARF22–24, indicating these genes are more likely acting as repressors. While glutamine (Q), leucine (L) and serine (S) rich regions were found in the middle regions of protein sequences of OsARF5–6, OsARF11–12, OsARF16–17, OsARF19, OsARF21, and OsARF25, implying that these genes are possibly transcriptional activators (Supplementary Fig. 1). All OsARF proteins with Q-rich middle regions belong to class IIa, while PST-rich OsARFs are all in another class (Fig. 3a, b; Supplementary Fig. 1).

No ARF proteins have been reported to have two DBDs domains as observed in the protein sequence of OsARF20 identified in this study. The results that the OsARF20 gene was not detectably expressed according to RT-PCR, EST, and FL-cDNA in this report suggested that the *OsARF20* gene might be a pseudogene, or expressed at specific developmental stages or under special conditions. The biological functions of this ARF protein remain to be elucidated.

Checking phenotypic growth changes of gene knock-out mutants is the direct way to find out which functions are associated with a particular ARF gene. Based on the characterization of these mutants, the biological functions of these genes are relatively well characterized in comparison to those of other *ARF* genes.

It has been reported that mutations in the *ARF3/ETT* gene affect gynoecium patterning (Sessions et al., 1997; Nemhauser et al., 2000); loss-of-function mutations of *ARF7/MSG/NPH41/TIR5* result in impaired hypocotyls response to blue light and other differential growth responses associated with changes in auxin sensitivity (Watahiki and Yamamoto, 1997; Stowe-Evans et al., 1998; Harper et al., 2000); mutations in *ARF5/MP* interfere with the formation of vascular strands and the initiation of the body axis in the early embryo (Hardtke and Berleth, 1998); mutations in *ARF2/HSS* have been identified as suppressors of the hookless phenotype (Li et al., 2004); *ARF2* acts as a communication link between the ethylene and the auxin signaling pathways for regulating hypocotyls bending and, *ARF8* functions in hypocotyl elongation and is involved in auxin homeostasis (Tian et al., 2004).

Some of the *ARF* genes were independently isolated through the molecular genetic approach using morphological mutants, these were *ETTIN* (=*ARF3*, Sessions et al., 1997), *MONOPTEROS* (*MP*=*ARF5*, Hardtke and Berleth, 1998), *NPH4* (=*ARF7*, Harper et al., 2000) and *HSS* (=*ARF2*, Li et al., 2004). *Arabidopsis ARF1* and *ARF2* gene carrying T-DNA insertions were isolated and revealed that *ARF1* acts in a partially redundant manner with *ARF2* (Ellis et al., 2005). *ARF19* acts redundantly with *NPH4/ARF7* in controlling leaf expansion and lateral root growth (Okushima et al., 2005; Wilmoth et al., 2005) and, *ARF6* and *ARF8* act redundantly in flower maturation (Nagpal et al., 2005).

The activity of these ARFs is negatively regulated by heterodimerization with Aux/IAA proteins (Reed, 2001; Tiwari et al., 2003). In addition, ARF1 acted as a repressor of auxin-induced genes. *arf1* and *arf2* mutations had synergistic effects on some phenotypes and independent effects on others, indicating both

redundancy and specialization of *ARF1* and *ARF2* function. Furthermore, *nph4/arf7* and *arf19* mutations also enhanced the *arf2* senescence and abscission phenotypes, suggesting that ARFs of different functional classes can regulate common processes (Ellis et al., 2005).

Recently, 18 *ARF* gene family members have been identified and characterized in *Arabidopsis thaliana*. Most of the lines fail to show an obvious growth phenotype except for the previously identified *arf2/hss*, *arf3/ett*, *arf5/mp*, and *arf7/nph4* mutants, suggesting that there are functional redundancies among the ARF proteins. Subsequently, *arf7 arf19* double mutants were generated and showed a strong auxin-related phenotype not observed in the *arf7* and *arf19* single mutants, including severely impaired lateral root formation and abnormal gravitropism in both hypocotyl and root. Global gene expression analysis revealed that auxin-induced gene expression is severely impaired in the *arf7* single and *arf7 arf19* double mutants. The data suggest that the ARF7 and ARF19 proteins play essential roles in auxin-mediated plant development by regulating both unique and partially overlapping sets of target genes. These observations provide molecular insight into the unique and overlapping functions of *ARF* gene family members in *Arabidopsis* (Okushima et al., 2005).

To have an inkling about the functions of OsARFs, the phenotypes of rice *Tos17* retrotransposon insertion mutants (Miyao et al., 2003) of these genes with the aid of a *Tos17* rice mutant panel database (http://tos.nias.affrc.go.jp/) were checked and several insertion mutants corresponding to *OsARF5*, *OsARF11*, *OsARF12*, *OsARF19* and *OsARF24* were listed (Supplementary Table S4). The phenotypic variations of these insertion mutants were found to associate with several important traits such as fertility, height and grain yield. From these phenotypes, it can be speculated that these genes may play a critical role in different metabolic pathways and cellular processes in rice.

In conclusion, the very fact that most of the duplicated *OsARF*s are being expressed in different plant tissue/organs indicates that these genes have specific or redundant cellular functions. The unraveling of roles of individual members of this family in auxin signaling will require a concerted effort by adoption of diverse approaches, including molecular genetic analysis. The comparative and phylogenetic analyses of the *ARF* gene family in rice and *Arabidopsis* and the classification of *ARF* genes into orthologue groups will facilitate a comprehensive functional characterization of the *ARF* gene family. The results of the structure analyses of rice ARF proteins will pave the way for their functional analysis in the future.

Acknowledgements

We wish to thank two anonymous reviewers for providing much help and advice. This research was supported financially by the National Natural Science Foundation of China (Grant No. 30600349), the National Basic Research Program of China 973 Program (Grant No. 2005CB120801); the National High Technology Research and Development Program of China (863 Program) (Grant No. 2006AA10A102) and the Zhejiang Provincial Natural Science Foundation of China (Grant No. Y306149).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2007.01.006.

References

Abel, S., Theologis, A., 1996. Early genes and auxin action. Plant Physiol. 111, 9–17.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815.

Davies, P.J., 1995. Plant Hormones, Physiology, Biochemistry and Molecular Biology, 2nd ed. Kluwer, Dordrecht, The Netherlands.

Ellis, C.M., Nagpal, P., Young, J.C., Hagen, G., Guilfoyle, T.J., Reed, J.W., 2005. *AUXIN RESPONSE FACTOR1* and *AUXIN RESPONSE FACTOR2* regulate senescence and floral organ abscission in *Arabidopsis thaliana*. Development 132, 4563–4574.

Feng, Q., et al., 2002. Sequence and analysis of rice chromosome 4. Nature 420, 316–320.

Goff, S.A., et al., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296, 92–100.

Hagen, G., Guilfoyle, T., 2002. Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol. Biol. 49, 373–385.

Hagen, G., Martin, G., Li, Y., Guilfoyle, T.J., 1991. Auxin-induced expression of the soybean GH3 promoter in transgenic tobacco plants. Plant Mol. Biol. 17, 567–579.

Hardtke, C.S., Berleth, T., 1998. The *Arabidopsis* gene *MONOPTEROUS* encodes a transcription factor mediating embryo axis formation and vascular development. EMBO J. 17, 1405–1411.

Harper, R.M., et al., 2000. The *NPH4* locus encodes the auxin response factor ARF7, a conditional regulator of differential growth in aerial *Arabidopsis* tissue. Plant Cell 12, 757–770.

Huala, E., et al., 2001. The *Arabidopsis* Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res. 29, 102–105.

International Rice Genome Sequencing Project, 2005. The map-based sequence of the rice genome. Nature 436, 793–800.

Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K., Khurana, J.P., 2006a. Structure and expression analysis of early auxin-responsive *Aux/IAA* gene family in rice (*Oryza sativa*). Funct. Integr. Genomics 6, 47–59.

Jain, M., Kaur, N., Tyagi, A.K., Khurana, J.P., 2006b. The auxin-responsive *GH3* gene family in rice (*Oryza sativa*). Funct. Integr. Genomics 6, 36–46.

Jain, M., Tyagi, A.K., Khurana, J.P., 2006c. Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive *SAUR* gene family in rice (*Oryza sativa*). Genomics 88, 360–371.

Kikuchi, S., et al., 2003. Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science 301, 376–379.

Kim, J., Harter, K., Theologis, A., 1997. Protein–protein interactions among the Aux/IAA proteins. Proc. Natl. Acad. Sci. U. S. A. 94, 11786–11791.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Li, Y., Liu, Z.B., Shi, X., Hagen, G., Guilfoyle, T.J., 1994. An auxin-inducible element in soybean SAUR promoters. Plant Physiol. 106, 37–43.

Li, H., Johnson, P., Stepanova, A., Alonso, J.M., Ecker, J.R., 2004. Convergence of signaling pathways in the control of differential cell growth in *Arabidopsis*. Dev. Cell 7, 1–20.

Liscum, E., Reed, J.W., 2002. Genetics of Aux/IAA and ARF action in plant growth and development. Plant Mol. Biol. 49, 387–400.

Liu, Z.B., Ulmasov, T., Shi, X., Hagen, G., Guilfoyle, T.J., 1994. Soybean GH3 promoter contains multiple auxin-inducible elements. Plant Cell 6, 645–657.

Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.

Miyao, A., et al., 2003. Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15, 1771–1780.

Nagpal, P., et al., 2005. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development 132, 4107–4118.

Nakano, T., Suzuki, K., Fujimura, T., Shinshi, H., 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol. 140, 411–432.

Nemhauser, J.L., Feldman, L.J., Zambryski, P.C., 2000. Auxin and *ETTIN* in *Arabidopsis* gynoecium morphogenesis. Development 127, 3877–3888.

Okushima, Y., et al., 2005. Functional genomic analysis of the *AUXIN RESPONSE FACTOR* gene family members in *Arabidopsis thaliana*: unique and overlapping functions of *ARF7* and *ARF19*. Plant Cell 17, 444–463.

Ouellet, F., Overvoorde, P.J., Theologis, A., 2001. *IAA17/AXR3*: biochemical insight into an auxin mutant phenotype. Plant Cell 13, 829–841.

Overvoorde, P.J., et al., 2005. Functional genomic analysis of the *AUXIN/INDOLE-3-ACETIC ACID* gene family members in *Arabidopsis thaliana*. Plant Cell 17, 3282–2300.

Paterson, A.H., Bowers, J.E., Chapman, B.A., 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. U. S. A. 101, 9903–9908.

Raes, J., Vandepoele, K., Simillion, C., Saeys, Y., Van de Peer, Y., 2003. Investigating ancient duplication events in the *Arabidopsis* genome. J. Struct. Funct. Genomics 3, 117–129.

Reed, J.W., 2001. Roles and activities of Aux/IAA proteins in *Arabidopsis*. Trends Plant Sci. 6, 420–425.

Remington, D.L., Visionm, T.J., Guilfoylem, T.J., Reedm, J.W., 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.

Sasaki, T., et al., 2002. The genome sequence and structure of rice chromosome 1. Nature 420, 312–316.

Sato, Y., Nishimura, A., Ito, M., Ashikari, M., Hirano, H.Y., Matsuoka, M., 2001. Auxin response factor family in rice. Genes Genet. Syst. 76, 373–380.

Sessions, A., Nemhauser, J.L., McColl, A., Roe, J.L., Feldmann, K.A., Zambryski, P.C., 1997. *ETTIN* patterns the *Arabidopsis* floral meristem and reproductive organs. Development 124, 4481–4491.

Simillion, C., Vandepoele, K., Van Montagu, M.C., Zabeau, M., Van De Peer, Y., 2002. The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. 99, 13627–13632.

Stowe-Evans, E.L., Harper, R.M., Motchoulski, A.V., Liscum, E., 1998. *NPH4*, a conditional modulator of auxin-dependent differential growth responses in *Arabidopsis*. Plant Physiol. 118, 1265–1275.

Terol, J., Domingo, C., Talon, M., 2006. The GH3 family in plants: genome wide analysis in rice and evolutionary history based on EST analysis. Gene 371, 279–290.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Tian, C., et al., 2004. Disruption and overexpression of *auxin response factor 8* gene of *Arabidopsis* affect hypocotyl elongation and root growth habit, indicating its possible involvement in auxin homeostasis in light condition. Plant J. 40, 333–343.

Tiwari, S.B., Hagen, G., Guilfoyle, T.J., 2003. The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell 15, 533–543.

Ulmasov, T., Liu, Z.B., Hagen, G., Guilfoyle, T.J., 1995. Composite structure of auxin response elements. Plant Cell 7, 1611–1623.

Ulmasov, T., Hagen, G., Guilfoyle, T.J., 1997a. ARF1, a transcription factor that binds to auxin response elements. Science 276, 1865–1868.

Ulmasov, T., Murfett, J., Hagen, G., Guilfoyle, T.J., 1997b. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. Plant Cell 9, 1963–1971.

Ulmasov, T., Hagen, G., Guilfoyle, T.J., 1999a. Dimerization and DNA binding of auxin response factors. Plant J. 19, 309–319.

Ulmasov, T., Hagen, G., Guilfoyle, T.J., 1999b. Activation and repression of transcription by auxin-response factors. Proc. Natl. Acad. Sci. U. S. A. 96, 5844–5849.

Waller, F., Furuya, M., Nick, P., 2002. OsARF1, an auxin response factor from rice, is auxin-regulated and classifies as a primary auxin responsive gene. Plant Mol. Biol. 50, 415–425.

Wang, X., Shi, X., Hao, B., Ge, S., Luo, J., 2005. Duplication and DNA segmental loss in the rice genome: implications for diploidization. New Phytol. 165, 937–946.

Watahiki, M.K., Yamamoto, K.T., 1997. The massugu1 mutation of *Arabidopsis* identified with failure of auxin-induced growth curvature of hypocotyl confers auxin insensitivity to hypocotyl and leaf. Plant Physiol. 115, 419–426.

Wilmoth, J.C., et al., 2005. *NPH4/ARF7* and *ARF19* promote leaf expansion and auxin-induced lateral root formation. Plant J. 43, 118–130.

Yu, J., et al., 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296, 79–92.

Yu, J., et al., 2005. The genomes of *Oryza sativa*: a history of duplications. PLoS Biol. 3e, 38.