ORIGINAL PAPER

Mukesh Jain · Navneet Kaur · Rohini Garg ·
Jitendra K. Thakur · Akhilesh K. Tyagi ·
Jitendra P. Khurana

Structure and expression analysis of early auxin-responsive *Aux/IAA* gene family in rice (*Oryza sativa*)

Abstract Auxin exerts pleiotropic effects on plant growth and development by regulating the expression of early auxin-responsive genes of auxin/indoleacetic acid (*Aux/IAA*), small auxin-up RNA, and *GH3* classes. These genes have been studied extensively in dicots like soybean and *Arabidopsis*. We had earlier characterized a cDNA of the first monocot member of *Aux/IAA* family from rice. The achievement of the large scale rice genome sequencing combined with the availability of full-length cDNA sequences from Knowledge-based Oryza Molecular Biological Encyclopedia provided us the opportunity to draw up the first comprehensive list of *Aux/IAA* genes in a monocot. By screening the available databases, we have identified 31 *Aux/IAA* genes having high sequence identity within the conserved domains I, II, III, and IV. The genomic organization as well as chromosomal location of all the *Oryza sativa indoleacetic acid* (*OsIAA*) genes is reported. The rice Aux/IAA proteins can be classified in two groups (A and B) on the basis of their phylogenetic relationship with *Arabidopsis* Aux/IAA proteins. An evolutionary pattern of the rice *Aux/IAA* genes has been discussed by analyzing their structure (exon/intron organization) and duplications. Interestingly, the duplication of rice *Aux/IAA* genes was found to be associated with chromosomal block duplication events in rice. The in-silico analysis has been complemented with real-time polymerase chain reaction analysis to quantify transcript levels of all *Aux/IAA* family members. *OsIAA* genes showed differential and over-lapping organ-specific expression patterns in light- and dark-grown seedlings/plants. Although auxin enhanced the transcript abundance of most of the *OsIAA* genes, the effect was more pronounced on *OsIAA9*, *14*, *19*, *20*, *24*, and *31*. These results provide a foundation for future studies on elucidating the precise role of rice *Aux/IAA* genes in early steps of auxin signal transduction.

Keywords *Aux/IAA* · Auxin · Phylogenetic analysis · *Oryza sativa* (rice) · Real-time PCR

M. Jain · N. Kaur · R. Garg · J. K. Thakur · A. K. Tyagi ·
J. P. Khurana (✉)
Interdisciplinary Centre for Plant Genomics and Department of Plant Molecular Biology, University of Delhi South Campus, Benito Juarez Road,
New Delhi, 110021, India
e-mail: khuranaj@genomeindia.org
Tel.: +91-011-24115126
Fax: +91-011-24115270

Introduction

The phytohormone auxin plays a critical role in regulating several plant responses, including cell elongation, cell division, differentiation, root initiation, apical dominance, and tropic responses. Auxin rapidly and specifically alters transcript levels of numerous genes (Abel and Theologis 1996). In an attempt to understand the molecular mechanism of auxin action, several auxin-responsive genes have been identified and characterized from different plant species. These auxin-responsive genes have been broadly grouped into three major classes: auxin/indoleacetic acid (*Aux/IAA*), *GH3*, and small auxin-up RNA (*SAUR*) gene families (Guilfoyle 1999).

Following the initial identification of *Aux/IAA* genes from soybean (Walker and Key 1982; Ainley et al. 1988), members of this class were isolated from pea (Theologis et al. 1985), *Arabidopsis* (Conner et al. 1990; Abel et al. 1995), mung bean (Yamamoto et al. 1992), and rice (Thakur et al. 2001). The *Aux/IAA* genes are rapidly induced by auxin in the presence of a translational inhibitor, cycloheximide, indicating that these represent a class of primary auxin-responsive genes (Theologis et al. 1985; Abel et al. 1995; Thakur et al. 2005).

The *Aux/IAA* genes encode short-lived nuclear proteins comprising four highly conserved domains, designated as domain I, II, III, and IV (Abel et al. 1994). Recently, domain I has been shown to act as a strong transcriptional repressor (Tiwari et al. 2004). Domain II is responsible for

48

Table 1 *Aux/IAA* gene family in rice

Gene[a]	Accession no.[b]	TIGR locus ID[c]	ORF length[d] (bp)	No. of introns[e]	Chromosome no.[f]	Genomic locus[g] BAC/PAC name	Accession no.	cM position	Nearest marker[h]	Deduced polypeptide[i] Length (aa)	Mol wt (kDa)	pI	EST frequency[j]
OsIAA1	AK109373	LOC_Os01g08320	600	4	1	P0509B06	AP002903	19.9	C470	199	21.58	5.21	12
OsIAA2	AK100314	LOC_Os01g09450	717	4	1	P0710E05	AP002743	25.4	C101	238	24.37	8.07	5
OsIAA3	AK104654	LOC_Os01g13030	792	4	1	P0485D09	AP001859	36.9	R665	263	28.22	6.79	51
OsIAA4	AK103865	LOC_Os01g18360	270	2	1	P0511C01	AP002070	54.3	C916A	89	10.20	5.53	1
OsIAA5	AK106121	LOC_Os01g48450	816	4	1	OJ1117_G01	AP003374	116.5	S2717	271	28.43	8.81	9
OsIAA6	AK068600	LOC_Os01g53880	1008	5	1	P0439E07	AP003768	127.3	G393	335	36.69	7.77	37
OsIAA7	AK121870	LOC_Os02g13520	903	5	2	OSJNBb0035N08	AP005756	33.6–36.3	C630	300	32.21	6.62	3
OsIAA8	AK066518	LOC_Os02g49160	618	3	2	P0685G12	AP005113	128.3	S11127	205	22.18	8.68	3
OsIAA9	AK073365	LOC_Os02g56120	549	2	2	OJ1548_F12	AP004240	147.2–150.5	S1730	182	19.35	5.13	5
OsIAA10	AK069892	LOC_Os02g57250	843	4	2	P0643F09	AP005111	152.7	G1234	280	30.04	6.14	18
OsIAA11	NF	LOC_Os03g43400	702	4	3	OSJNBa0010N03	AC145379	94.9–96.6	S13580	233	25.31	6.12	2
OsIAA12	AK073044	LOC_Os03g43410	681	2	3	OSJNBa0010N03	AC145379	94.9–96.6	S13580	226	23.81	7.73	5
OsIAA13	AK059838	LOC_Os03g53150	711	4	3	OJ1365_D05	AC096855	137.6	S770	236	25.10	8.95	55
OsIAA14	AK059619	LOC_Os03g58350	588	2	3	OSJNBa0094F01	AC093713	149.1–151.1	R2224	195	20.12	8.67	2
OsIAA15	AK100080	LOC_Os05g08570	639	4	5	OSJNBa0029B02	AC144738	33.6	S16612S	212	22.88	5.55	13
OsIAA16	NF	LOC_Os05g09480	687	5	5	OJ1097_A12	AC093954	37.2	R3572	228	23.59	6.06	3
OsIAA17	AK106192	LOC_Os05g14180	774	4	5	B1402B06	AC145477	50.2	S10091	257	27.47	7.09	102
OsIAA18	AK071192	LOC_Os05g44810	984	5	5	OSJNBa0075A10	AC144740	107.4	C11084	327	36.05	5.37	5
OsIAA19	AK109363	LOC_Os05g48590	846	4	5	OSJNBa0001A14	AC144735	112.4–115.7	S974	281	29.96	5.98	15
OsIAA20	AK102541	LOC_Os06g07040	552	3	6	P0680A03	AB023482	12.9–13.5	R2634	183	19.90	6.01	2
OsIAA21	AK121989	LOC_Os06g22870	804	4	6	OSJNBa0012F14	AP004784	65.8	C62815S	267	28.35	6.99	26
OsIAA22	NF	LOC_Os06g24850	450	1	6	P0526E12	AP003574	65.8	S20459S	149	16.92	7.28	NF
OsIAA23	AK069376	LOC_Os06g39590	582	3	6	P0417D05	AP004236	84.5	C12560S	193	20.47	9.03	1
OsIAA24	AK103483	LOC_Os07g08460	660	2	7	OJ1506_G02	AP003835	31.0–35.7	R2401	219	23.05	7.63	15
OsIAA25	AK068232	LOC_Os08g01780	741	3	8	P0007D08	AP004584	3.0	C1017	246	27.32	8.42	1
OsIAA26	AK111128	LOC_Os09g35870	423	4	9	OJ1439_F07	AP005681	79.1	R3330	140	15.46	4.73	1
OsIAA27	NF	LOC_Os11g11410	432	3	11	OSJNBa0052C03	AC123524	33.4	S2712	143	16.28	6.76	1
OsIAA28	NF	LOC_Os11g11420	489	3	11	OSJNBa0052C03	AC123524	33.4	S2712	162	18.51	4.39	NF
OsIAA29	NF	LOC_Os11g11430	516	3	11	OSJNBa0052C03	AC123524	33.4	S2712	171	18.73	3.82	1
OsIAA30	AK068213	LOC_Os12g40890	834	4	12	OSJNBb0062H20	AL837528	99.7	S861	277	28.99	5.00	28

Table 1 (continued)

Gene[a]	Accession no.[b]	TIGR locus ID[c]	ORF length[d] (bp)	No. of introns[e]	Chromosome no.[f]	Genomic locus[g] BAC/PAC name	Accession no.	cM position	Nearest marker[h]	Deduced polypeptide[i] Length (aa)	Mol wt (kDa)	pI	EST frequency[j]
OsIAA31	AK073361	LOC_Os12g40900	594	2	12	OSJNBb0062H20	AL837528	99.7	S861	197	20.90	8.79	7

NF Not found

[a]Systematic designation given to rice *Aux/IAA* genes
[b]Accession numbers of full-length cDNA sequence available at KOME (http://www.cdna.01.dna.affrc.go.jp/cDNA/)
[c]Locus ID of each *OsIAA* gene on rice chromosome pseudomolecules available at TIGR (release 3)
[d]Length of open reading frame in base pairs
[e]Number of introns present within ORF
[f]Chromosomal localization of *OsIAA* gene
[g]Name, accession number, and approximate cM position of the BAC/PAC clone in which *OsIAA* gene is present
[h]Nearest marker to the *OsIAA* gene
[i]Length (number of amino acids), molecular weight (kiloDaltons), and isoelectric point (pI) of the deduced polypeptide
[j]Number of ESTs present in Genbank release 111204

rapid degradation of Aux/IAA proteins (Worley et al. 2000; Ouellet et al. 2001). Domain III is part of an amphipathic βαα-DNA recognition motif found in β ribbon of DNA binding domain of prokaryotic repressors such as MetJ and Arc (Abel et al. 1994; Phillips 1994). However, its role in DNA binding has not been demonstrated yet. Domains II and IV also contain nuclear localization signals (NLSs) (Abel et al. 1994; Abel and Theologis 1996). Domains III and IV mediate homo- and hetero-dimerization among the Aux/IAA proteins and auxin response factors (ARFs) (Kim et al. 1997; Ulmasov et al. 1997; Ouellet et al. 2001). The DNA-binding domain of ARFs binds to auxin-responsive elements (AuxREs) present within the promoters of auxin-responsive genes and regulates their expression (Ulmasov et al. 1995; Kim et al. 1997; Ulmasov et al. 1997; Tiwari et al. 2003).

Molecular genetic analyses of several mutants of *Aux/IAA* genes have demonstrated that they play a central role in regulating plant growth and development (Rouse et al. 1998; Tian and Reed 1999; Gray and Estelle 2000; Nagpal et al. 2000; Reed 2001; Rogg et al. 2001; Liscum and Reed 2002). Biochemical studies showed that the conserved region of domain II of Aux/IAA proteins is responsible for rapid degradation. Single amino acid change in domain II resulted in altered auxin response due to increased protein accumulation, suggesting that rapid degradation of Aux/IAA proteins is necessary for a normal auxin response (Worley et al. 2000). The degradation of Aux/IAA proteins was found to be proteasome dependent (Gray et al. 2001; Ramos et al. 2001; Thakur et al. 2005). Auxin treatment in fact promotes degradation of Aux/IAA proteins by enhancing the interaction between SCF[TIR1] complex and Aux/IAA proteins by affecting the SCF component, TIR1, or its associated proteins (Gray et al. 2001; Kepinski and Leyser 2004). The F-box protein, TIR1, a component of the SCF[TIR1] complex, and a few other F-box proteins have been identified recently as the auxin receptors (Dharmasiri et al. 2005a,b; Kepinski and Leyser 2005).

Earlier, we reported the isolation and characterization of an *OsiIAA1* cDNA, representing the first monocot *Aux/IAA* gene from rice (Thakur et al. 2001) and demonstrated that the nuclear-localized OsiIAA1 protein is degraded rapidly on auxin application via proteasome (Thakur et al. 2005). In the present study, we have identified and comprehensively analyzed the entire early auxin-responsive *Aux/IAA* gene family from rice (*Oryza sativa*). The work involved the identification of *Aux/IAA* gene family from rice, analysis of their chromosomal distribution, gene structure, gene duplications, and phylogenetic relationship. Their expression has been analyzed in terms of the frequency of expressed sequence tags (ESTs) (available in databases). Real-time polymerase chain reaction (PCR) analysis demonstrated that *Aux/IAA* genes are expressed differentially in various organs/tissues grown in light or dark and by auxin treatment in rice, indicating that they may perform specific as well as redundant functions in different cells/tissues.

Materials and methods

Identification of *Aux/IAA* gene family in rice

For the identification of *Aux/IAA* homologs in rice, the Knowledge-based Oryza Molecular Biological Encyclopedia (KOME, http://www.cdna01.dna.affrc.go.jp/cDNA) (Kikuchi et al. 2003), the National Centre for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/BLAST), and The Institute for Genomic Research (TIGR) database (http://www.tigrblast.tigr.org/euk-blast) resources were used. The amino acid sequences of 29 Aux/IAA proteins from *Arabidopsis*, downloaded from The *Arabidopsis* Information Resource (TAIR), were used to search for their homologs in rice in the KOME using the TBLASTN program (Altschul et al. 1997). The NCBI and TIGR databases were searched for additional members of rice *Aux/IAA* gene family by TBLASTN and BLASTN. The EST analysis was performed using MEGABLAST tool against the EST database of rice (*O. sativa*) available at the NCBI.

Sequence analysis

Each of the *Aux/IAA* genes was positioned on rice chromosome pseudomolecules available at TIGR (release 3) by the BLASTN search. The number and position of exons and introns for individual *O. sativa indoleacetic acid* (*OsIAA*) genes were determined by comparison of the cDNAs with their corresponding genomic DNA sequences. Multiple sequence alignments were done using the ClustalX (version 1.83) program (Thompson et al. 1997). The phylogenetic analysis was carried out by neighbor-joining method, and the unrooted tree was displayed using the NJPLOT program. The Gene Runner program (version 3.04) was used for the DNA and protein sequence analysis.

Plant material and growth conditions

Seeds of rice (*O. sativa* L. ssp. *indica* var. Pusa Basmati 1), after disinfection with 0.1% $HgCl_2$ for 1 h and thorough washing, were soaked overnight in reverse-osmosis (RO) water. The different tissues were harvested from 6-day-old seedlings grown on cotton saturated with RO water, either in dark or a 14-h light and 10-h dark cycle, in a culture room maintained at 28±1°C. Floral tissue was collected from rice plants grown under the field conditions. The callus tissue was raised as described previously (Jain et al. 2004). For auxin treatment, the coleoptile apical segments (10 mm) from the 3-day-old etiolated rice seedlings were incubated in KPSC buffer (10 mM potassium phosphate, pH 6.0, 2% sucrose, 50 μM chloramphenicol) for 16 h to deplete endogenous auxin. The buffer was changed every 2 h, and the coleoptile segments transferred to a fresh buffer with 30 μM concentration of 2,4-dichlorophenoxy-acetic acid (2,4-D) and incubated for 3 h.

RNA isolation and real-time PCR analysis

Total RNA was extracted using the RNeasy Plant mini kit (Qiagen, Germany) according to the manufacturer's instructions, followed by DNase I treatment to remove any genomic DNA contamination. The quantitative real time PCR analysis was performed as described (Jain et al. 2005). In brief, the cDNA samples synthesized from 3 μg of the total RNA using High Capacity cDNA Archive kit (Applied Biosystems, USA) were used as template and mixed with 200 nM of each primer and SYBR Green PCR Master Mix (Applied Biosystems) for real-time PCR analysis, using ABI Prism 7000 Sequence Detection System and Software (PE Applied Biosystems) according to the manufacturer's instructions. Each pair of primers designed by using Primer Express 2.0 software (PE Applied Biosystems) was checked by the BLAST program in rice genomic sequence available in TIGR database to ensure that the primers amplify a unique and desired cDNA segment. The primer sequences are listed in Supplementary Table S1. The specificity of the reactions was verified by melting curve analysis. The relative mRNA levels for each of the 31 *OsIAA* genes in RNA isolated from various tissue samples were quantified with respect to the internal standard, *UBQ5*. At least two independent RNA isolations were used for cDNA synthesis, and each cDNA sample was subjected to real-time PCR analysis in triplicate.

Results and discussion

Aux/IAA gene family in rice

The *Aux/IAA* genes are present as multigene families in soybean (Ainley et al. 1988), pea (Oeller et al. 1993), mung bean (Yamamoto et al. 1992), tobacco (Dargeviciute et al. 1998), and tomato (Nebenfuhr et al. 2000). The *Arabidopsis* genome contains 29 *Aux/IAA* genes (Liscum and Reed 2002). The hybridization of multiple DNA fragments in Southern analysis of rice genomic DNA using *OsiIAA1* cDNA as probe, at low stringency, indicated that *Aux/IAA* genes may be represented as a multigene family in the rice genome too (Thakur et al. 2001). In an attempt to identify Aux/IAA protein coding genes in rice, the TBLASTN search of full-length cDNA clones of rice available at KOME was performed using 29 AtIAA proteins as query. In this search, 25 nonredundant clones (among a total of 48) having high sequence similarity with AtIAA proteins could be identified and were designated as *OsIAA* genes. Six additional members, not represented in KOME cDNA collection, were identified by TBLASTN or BLASTN search of the whole rice genomic and annotated sequences (Goff et al. 2002; Yu et al. 2002; http://www.tigr.org/tdb/e2k1/osa1/; http://www.rgp.dna.affrc.go.jp/). Thus, the overall analysis of the complete genome of rice revealed that auxin-inducible *Aux/IAA* gene family is comprised of 31 members, and they were designated as *OsIAA1* to *31*

according to their position on rice chromosomes (Table 1, Fig. 1). The corresponding TIGR gene locus IDs (release 3) for all the 31 members are listed in Table 1. Most of these have also been annotated as *Aux/IAA* family members by TIGR, and others as auxin-responsive genes.

Chromosomal distribution of rice *Aux/IAA* genes

The bacterial artificial chromosome (BAC) or P1 phage-derived artificial chromosome (PAC) clones carrying the *Aux/IAA* genes were identified (Table 1). The approximate chromosome map positions of BACs/PACs given in centi-Morgans (cM) from top of the chromosome and their nearest marker are indicated in Table 1. In addition, the position (in megabases) and direction of transcription (arrows) of each gene were determined on rice chromosome pseudomolecules available at TIGR (release 3) as shown in Fig. 1.

The 31 rice *Aux/IAA* genes were found to be distributed on 10 of the 12 rice chromosomes. No *OsIAA* gene could be located on chromosome 4 and 10. Six *Aux/IAA* genes are present on chromosome 1; five on chromosome 5; four on chromosome 2, 3, and 6 each; three on chromosome 11; two on chromosome 12; and one each on chromosomes 7, 8, and 9 (Fig. 1). The distribution of 31 members of this multigene family did not reveal evident clusters. However, fine mapping analysis revealed the presence of adjacent genes on chromosome 3 (*OsIAA11* and *12*), 11 (*OsIAA27*, *28*, and *29*), and 12 (*OsIAA30* and *31*) possibly due to tandem duplication either in inverse or same orientation (Fig. 1). Among the nine nonoverlapping duplicated blocks described by Paterson et al. (2004), four occurring between chromosomes 1 and 5 (*OsIAA1* and *15*, *OsIAA2* and *16*, *OsIAA3* and *17*, *OsIAA5* and *19*, *OsIAA6* and *18*), 2 and 6 (*OsIAA9* and *20*), 3 and 7 (*OsIAA14* and *24*), and 3 and 12 (*OsIAA12* and *31*; *OsIAA11* and *30*) gave rise to *Aux/IAA* gene duplications. The duplicated block between chromosome 3 and 12 contains two adjacent duplicated *OsIAA* genes, suggesting the occurrence of local duplications prior to the chromosomal segment duplication. The *Aux/IAA*
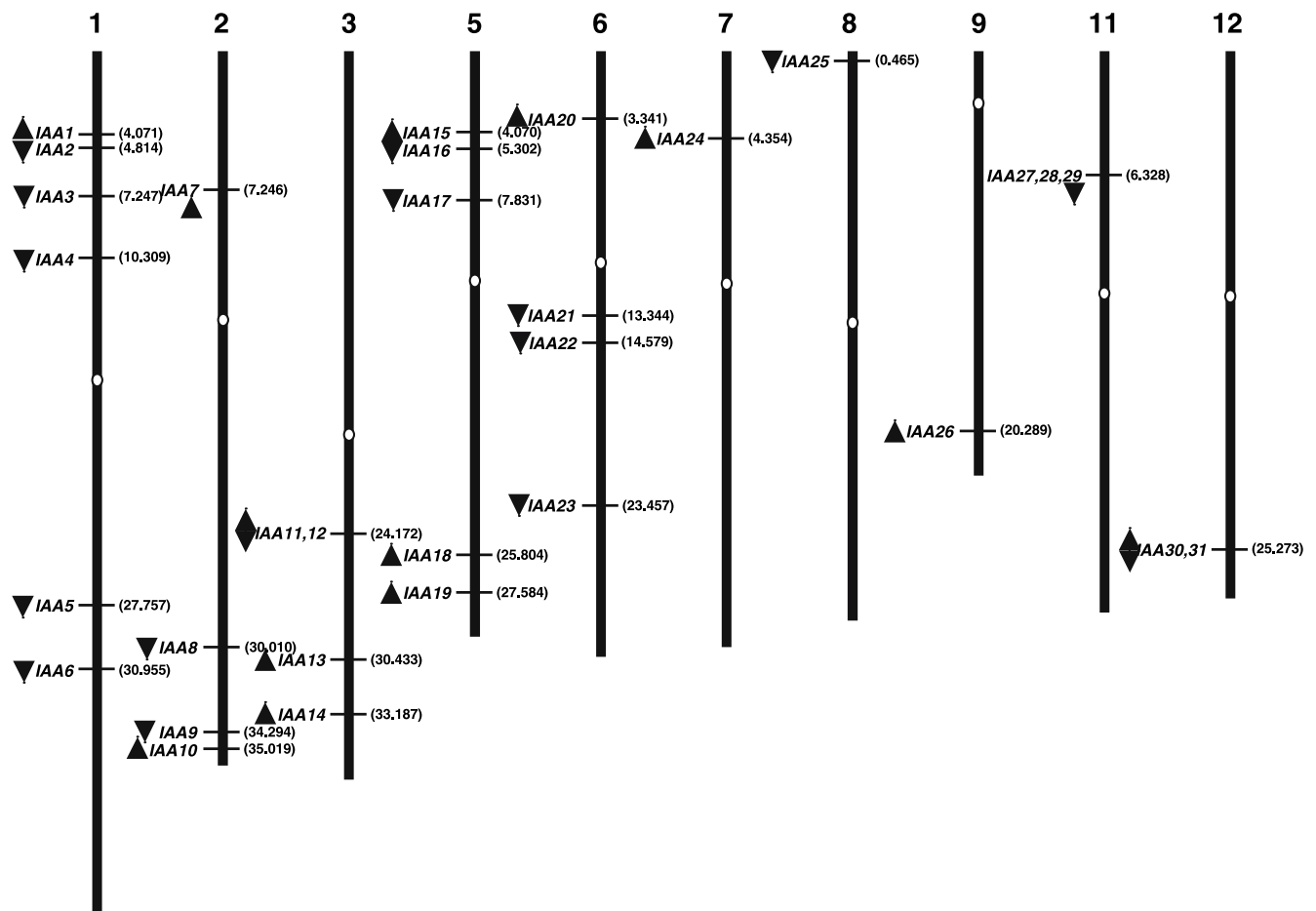

Fig. 1 Genomic distribution of *Aux/IAA* genes on rice chromosomes. White ovals on the chromosomes (*vertical bar*) indicate the position of centromeres. The *arrows* next to gene names show the direction of transcription. The number in parentheses designate the position of the first exon of each *Aux/IAA* gene in megabases (Mb) on rice chromosome pseudomolecules at TIGR (release 3). The chromosome numbers are indicated at the top of each bar

genes in *Arabidopsis* also have been found to be located on duplicated chromosomal segments (Remington et al. 2004). Similar examples of large gene families in eukaryotes have been reported, which may in part be explained due to extensive gene duplications.

Sequence analysis of OsIAA proteins

All deduced Aux/IAA protein sequences described thus far generally range from 20 to 35 kDa in size and harbor the four conserved domains (Hagen and Guilfoyle 2002;
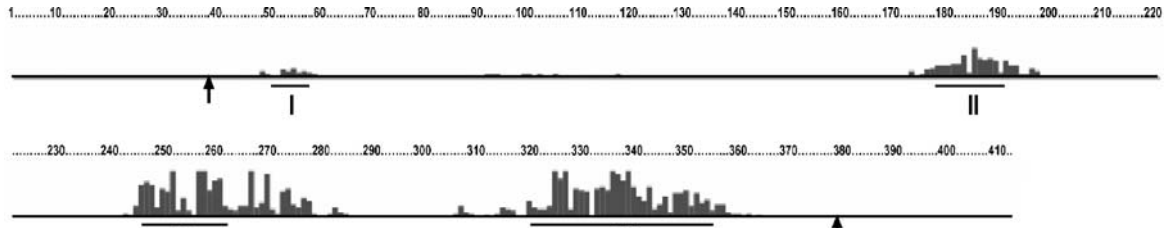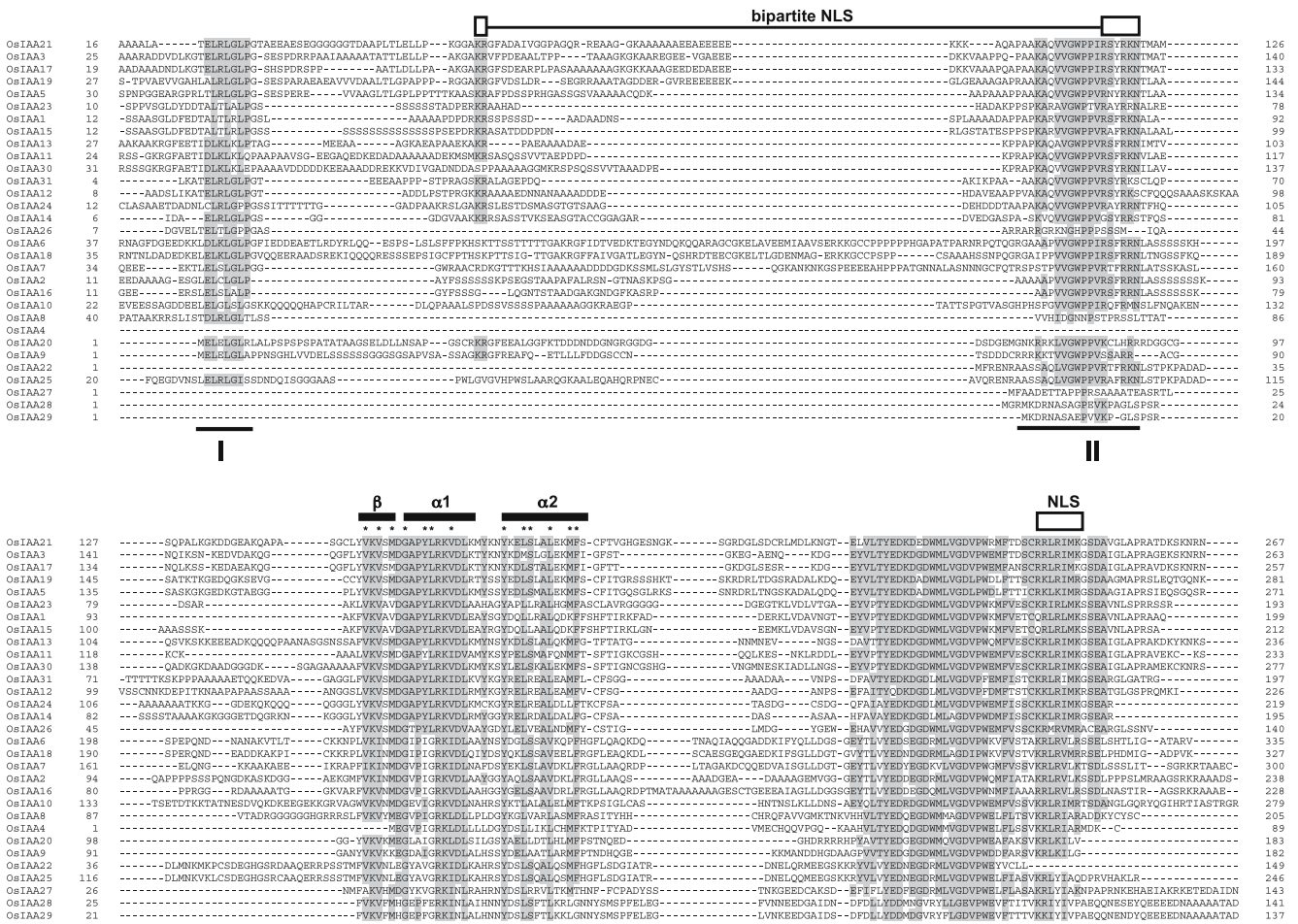

Fig. 2 a Alignment of rice Aux/IAA proteins obtained with the ClustalX program. The height of the bars indicates the number of identical residues per position. The *arrows* indicate the region shown in the multiple alignment below. **b** Multiple alignments of the domains I to IV (region marked with *arrows* in Fig. 2a) of the rice Aux/IAA proteins obtained with ClustalX and manual correction. Conserved residues (present in more than 50% of aligned sequences) are highlighted in *gray boxes*. Amino acids considered as conserved are: K and R; D and E; I, L, and V. Gaps (*marked with dashes*) have been introduced to maximize the alignments. Conserved domains are *underlined* and indicated by roman numerals. Hydrophobic amino acid residues in the predicted conserved amphipathic βαα motif (*black bar*) are indicated by *asterisks*. NLS are represented by *open boxes*. The respective amino acid position is given on the left and right of each sequence

Liscum and Reed 2002). The deduced molecular mass of the Aux/IAA polypeptides ranges from 10 kDa for OsIAA4 to 37 kDa for OsIAA6 (Table 1). To examine in detail the domain organization of OsIAA proteins, the multiple sequence alignments of the full-length protein sequences were done using the ClustalX program. Most of the OsIAA proteins showed the presence of all the four conserved domains (domains I, II, III, and IV; Fig. 2). A pair-wise analysis of the full-length OsIAA protein sequences indicated that the overall identities range from 14 to 76%. However, the amino acid identity within the conserved domains reaches up to 96%. Domain I represents leucine-
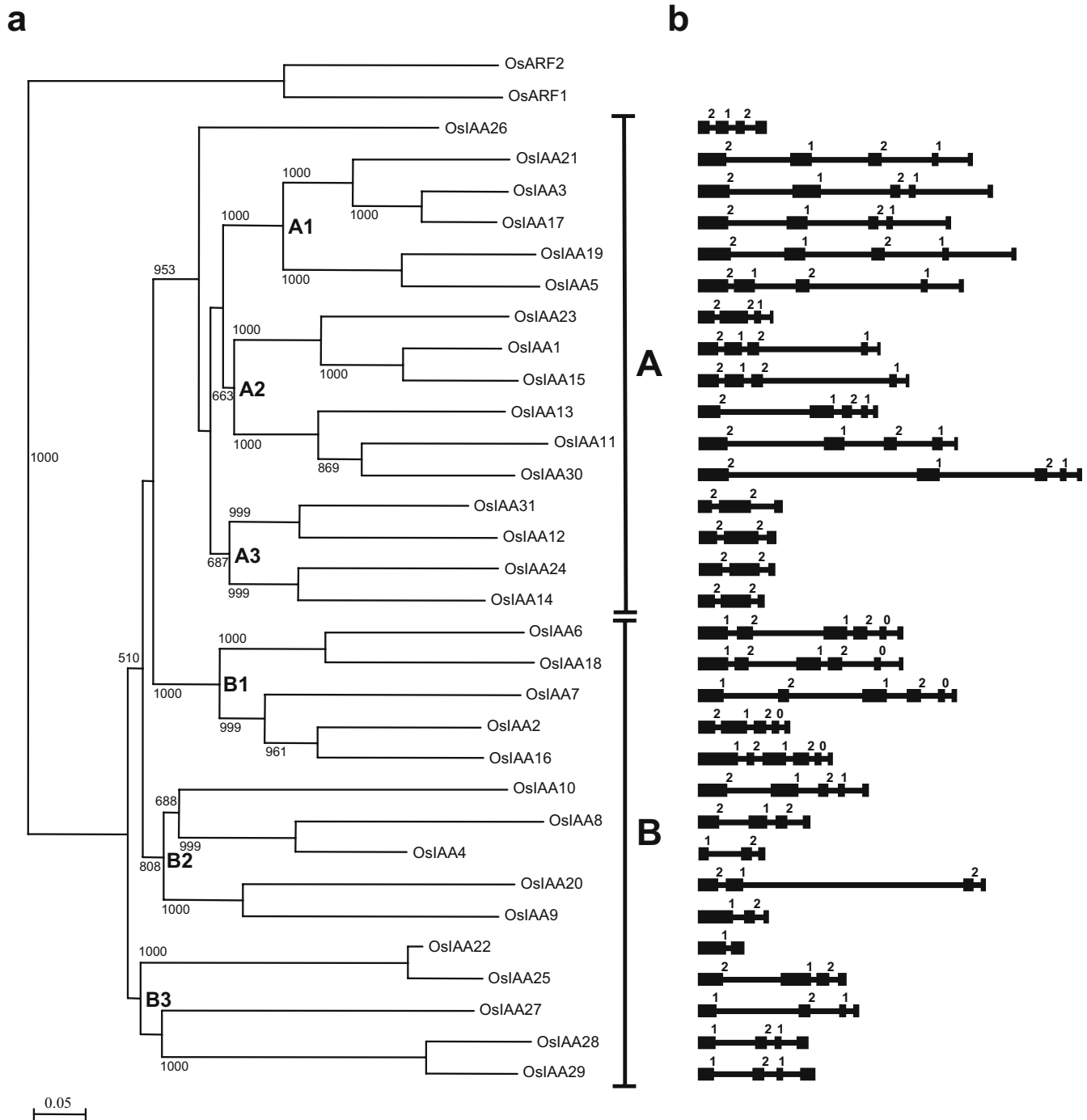
Fig. 3 a Phylogenetic relationship among the rice Aux/IAA proteins. The unrooted tree was generated using ClustalX program by neighbor-joining method. Bootstrap values (above 50%) from 1,000 replicates are indicated at each node. Two rice ARF protein sequences (OsARF1, AJ306306 and OsARF2, AB071293) were used as outgroup. **b** Exon–intron organization of corresponding *Aux/IAA* genes. The exons and introns are represented by *black boxes* and *lines*, respectively. The numbers 0, 1, and 2 represent phase 0, 1, and 2 introns, respectively

rich region and is least conserved among the family members (absent in OsIAA4, 22, 27, 28, and 29; Fig. 2b). The proline-rich domain II is comparatively more conserved (absent in OsIAA4, 8, 27, 28, and 29; Fig. 2b). The classification of *OsIAA4, 27, 28,* and *29* (harboring only domains III and IV) as *Aux/IAA* family members was confirmed by construction of phylogenetic tree based upon domain III and IV amino acid sequences of 31 OsIAA and two representative OsARF proteins. The Aux/IAA and ARF sequences formed two separate clusters and OsIAA4, 27, 28, and 29 clustered distinctly with other OsIAA protein sequences (data not shown). The absence of domain II is really perplexing because it mediates protein degradation by interaction with TIR1 of SCF^{TIR1} complex (Worley et al. 2000; Gray et al. 2001; Kepinski and Leyser 2004), which is presumed to be a prerequisite for a normal auxin response to ensue. Thus, it will be interesting to elucidate whether an SCF^{TIR1}-independent pathway is responsible for degradation of such proteins, or some other as yet undiscovered component ensures their proteolysis. Domains III and IV are invariantly conserved in all the 31 OsIAA proteins (Fig. 2b). An amphipathic βαα fold present within domain III and adjacent hydrophobic amino acid residues is also essentially conserved among all the rice Aux/IAA proteins (Fig. 2b). This fold is similar to the β ribbon of DNA recognition motif of prokaryotic repressors such as Arc and MetJ (Phillips 1994). However, there is no evidence so far that the Aux/IAA proteins bind to the DNA directly. Domains III and IV are involved in homo- and hetero-dimerization among the Aux/IAA proteins and ARFs (Kim et al. 1997; Ulmasov et al. 1997; Ouellet et al. 2001). The Aux/IAA proteins have been shown to be localized to the nucleus (Abel et al. 1994). Two types of putative nuclear localization signals were detected in most of the Aux/IAA proteins in rice. First, a bipartite structure comprised a conserved basic doublet KR between domains I and II and basic amino acids in domain II (Fig. 2b). Second, SV40-type NLS located in domain III (Fig. 2b). These putative NLSs may direct rice Aux/IAA proteins to the nucleus. However, functional validation of these putative NLSs will

be required to ascertain their in vivo role in subcellular localization of rice Aux/IAA proteins. In an investigation on the nuclear localization of at least one member of this family, OsiIAA1 (Thakur et al. 2001), redesignated here as OsIAA13, both the NLS sequences were found to be effective in translocating the GUS fusion protein to the nucleus in the onion epidermal cells (Thakur et al. 2005).

Gene structure and phylogenetic analysis of rice *Aux/IAA* genes

A comparison of the full-length cDNA sequences with the corresponding genomic DNA sequences showed that the coding sequence of the majority of the *Aux/IAA* genes (13 among 31) are indeed disrupted by four introns (Table 1; Fig. 3b) at perfectly conserved positions with respect to their amino acid sequence, suggesting a common ancestral gene with a classical pattern of five exons and four introns. The highly conserved intron phasing (Fig. 3b) also supported the idea of evolution of *Aux/IAA* genes from the common ancestral gene by exon shuffling (Kolkman and Stemmer 2001). However, variations in this basic gene structure were observed for other members, implicating mainly loss of one or more introns. Seven genes have lost one of the four putative ancestral introns, whereas six genes have lost two introns, and one gene showed the loss of three introns. Gain of one additional intron was observed for four genes. To examine the phylogenetic relationship among rice Aux/IAA proteins, an unrooted tree was constructed from alignments of the full-length Aux/IAA protein sequences including two representative OsARF protein sequences as outgroup (Fig. 3a). The Aux/IAA and ARF sequences formed separate clusters. All the rice Aux/IAA protein sequences were grouped into two major groups (group A and B) with well-supported bootstrap value similar to *Arabidopsis* (Remington et al. 2004). Sixteen and 15 OsIAA proteins were included in group A and B, respectively. Group A and B could be further subdivided into three subgroups each (A1–A3 and B1–B3) with varying degree of bootstrap support. Most of the *Aux/IAA* genes grouped together showed conserved gene structure (Fig. 3), in terms of exon/intron organization and intron phasing. Thirty one of the OsIAA proteins formed 12 sister pairs, 10 of which had very strong bootstrap support (99.9–100%). In *Arabidopsis* also, the *Aux/IAA* genes formed ten sister pairs, and all of them were found to be located on homologous duplicated chromosomal segments (Remington et al. 2004). Interestingly, 9 of the 12 sister pairs of rice *Aux/IAA* genes are also located on the duplicated chromosomal blocks described by Paterson et al. (2004). The sister pair OsIAA28 and 29 represents a local duplication event. The two other sister pairs may also be present on unidentified duplicated chromosomal blocks. Thus, it is remarkable that the duplication of the sister pairs of *Aux/IAA* genes is associated with chromosomal block duplications in both rice and *Arabidopsis*. The preferential retention of the duplicated *Aux/IAA* genes in rice might be either to maintain

a

b
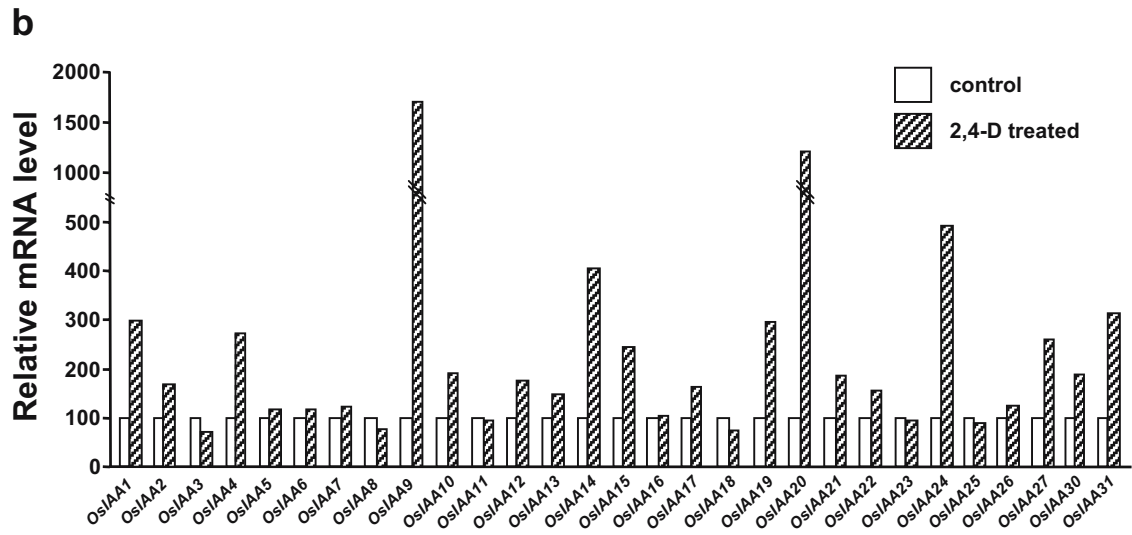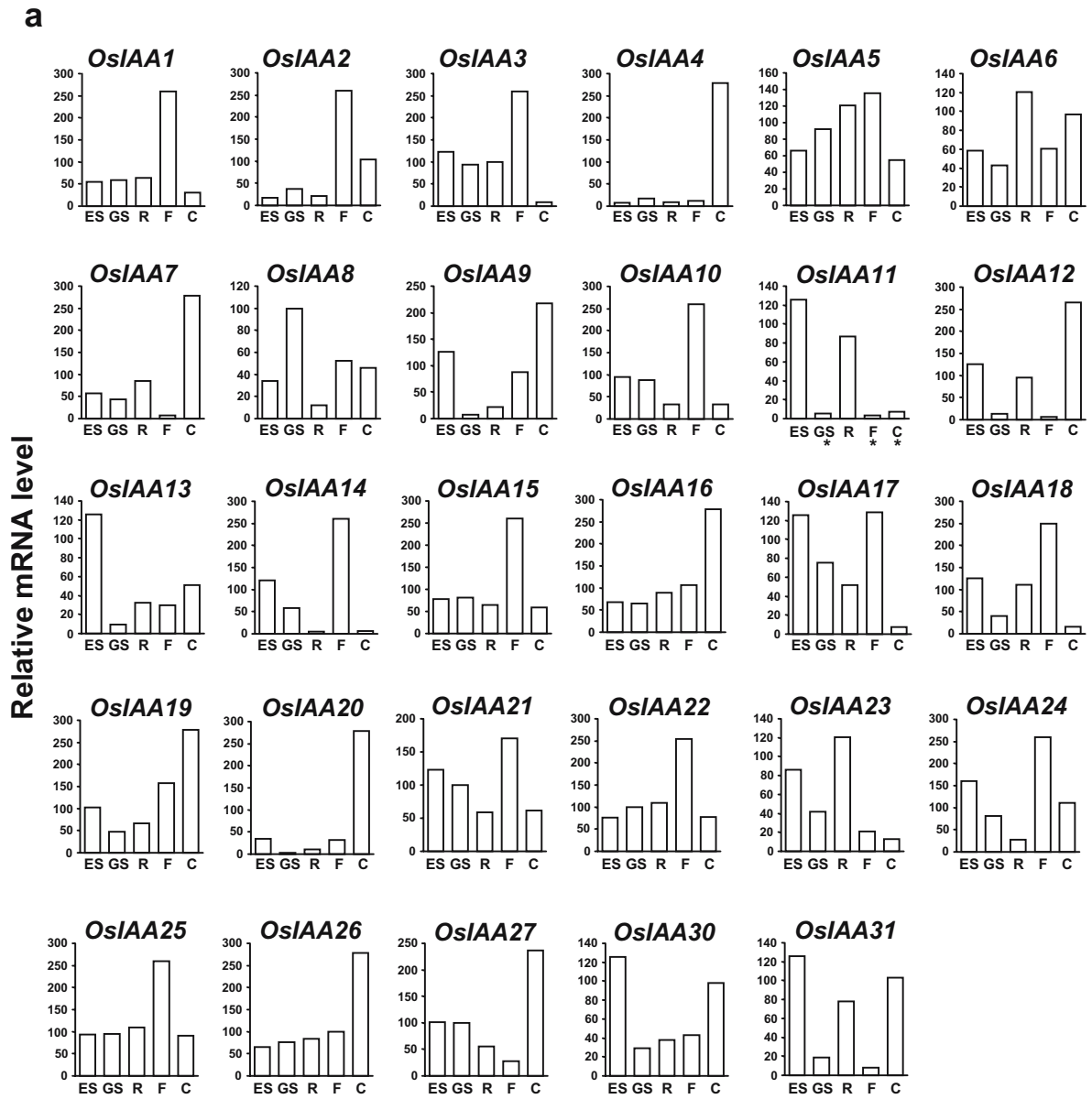
Fig. 5 Real-time PCR expression profiles of individual *OsIAA* genes. **a** The relative mRNA levels of individual *OsIAA* genes normalized with respect to housekeeping gene, *UBQ5*, in different tissues (*ES* etiolated shoots, *GS* green shoots, *R* roots, *F* flowers, *C* callus). **b** The relative mRNA levels of individual *OsIAA* genes in control (16 h auxin depleted) and 2,4-D (30 μm) treated coleoptile segments of 3-day-old etiolated rice seedlings. *Asterisks* indicate that the expression was close to the detection limit. *OsIAA28* could not be amplified at all, whereas *OsIAA29* transcription was detected only in roots and that too at very low level. The range of differences between the duplicates was 8–25%

the proper dosage relationships with interacting proteins such as ARFs or due to the duplication of large chromosomal regions retaining their *cis*-regulatory elements, which is essentially similar to *Arabidopsis* (Remington et al. 2004). The retention of the duplicated *Aux/IAA* genes also supports the idea that the genes involved in transcription and signal transduction have been preferentially retained in *Arabidopsis* (Blanc and Wolfe 2004). The fact that the *Aux/IAA* gene family in both rice and *Arabidopsis* showed high degree of duplicated gene retention is particularly interesting as both species experienced similar evolutionary mechanisms, i.e., polyploidization followed by diploidization (Bowers et al. 2003; Paterson et al. 2004; Wang et al. 2005). Taken together, these observations throw some light on the evolutionary steps encountered during the diversification of *Aux/IAA* genes. It appears that the evolutionary history of the *Aux/IAA* gene family reflects a succession of genomic rearrangements and expansion due to extensive duplication and diversification.

To examine the phylogenetic relationships of rice and *Arabidopsis* Aux/IAA proteins, an unrooted tree was constructed from alignments of their full-length Aux/IAA protein sequences (Supplementary Fig. S1) including two representative OsARF protein sequences as outgroup (Fig. 4). Based upon the sequence homology, the Aux/IAA proteins clustered distinctly into two groups (group A and B) consistent with analysis by Remington et al. (2004). The events of expansion of *Aux/IAA* gene family were found in rice vis-a-vis *Arabidopsis* (Fig. 4), which indicate that the speciation of rice *Aux/IAA* genes preceded their duplications (Guyot and Keller 2004). None of the rice Aux/IAA protein grouped with AtIAA29, 30, and 31. Genes from groups A and B are present both in *Arabidopsis* and rice, indicating that these appeared before the divergence between monocots and dicots.

Differential expression of *OsIAA* genes

The frequency of ESTs or cDNAs available in different databases has been considered as a useful tool for preliminary analysis of gene expression (Adams et al. 1995). A megablast search in EST database available at NCBI resulted in identification of ESTs for most of the *OsIAA* genes, but the frequency of ESTs for individual genes varies greatly (Table 1). For example, 37 ESTs are deposited for

OsIAA6 and more than 50 for *OsIAA3*, *13*, and *17*, whereas no ESTs could be identified for *OsIAA22* and *28* (Table 1). The full-length cDNA clones have been identified for 25 of the 31 members. The transcription of *OsIAA22* and *28* genes, however, remained unclear, as no corresponding full-length cDNA or EST could be identified from the databases. The expression of all the 31 *OsIAA* genes was thus verified experimentally and quantitated by real-time PCR analysis (Fig. 5) in the present study.

Many of the *Aux/IAA* genes in soybean, pea, *Arabidopsis*, and rice were found to be differentially expressed in different tissues or in response to exogenous auxin and light stimuli (Theologis et al. 1985; Ainley et al. 1988; Yamamoto et al. 1992; Abel et al. 1995; Thakur et al. 2001). To determine the organ-specific expression pattern of each *OsIAA* gene, real-time PCR was performed with total RNA isolated from etiolated shoots, green shoots, roots, flowers, and callus. *OsIAA* genes showed complexity of specific and overlapping expression patterns in various tissues/organs analyzed (Fig. 5a), indicating they might perform specific functions or act redundantly. The expression of *OsIAA28* could not be detected in any RNA samples analyzed. In addition, the transcripts for *OsIAA29* were detected in roots only and at a very low level (data not shown). Significant differences were found in the transcript abundance of *OsIAA* genes in etiolated and green shoots (Fig. 5a), indicating their light regulation. In an earlier study from our group, the transcript levels of *OsiIAA1* (renamed as *OsIAA13* in this study) were found to be downregulated when etiolated seedlings were irradiated with white light (Thakur et al. 2001). Many of the auxin-responsive mutants of *Arabidopsis* also exhibit altered light responses mediated by red/far-red or blue light, suggesting a cross-talk between auxin and light signaling (Reed et al. 1998; Stowe-Evans et al. 1998; Soh et al. 1999; Nagpal et al. 2000). Interestingly, six sister pair *OsIAA* genes (*OsIAA1* and *15*, *OsIAA3* and *17*, *OsIAA9* and *20*, *OsIAA12* and *31*, *OsIAA14* and *24*, *OsIAA22* and *25*) showed essentially similar expression profiles in various tissues, suggesting that they may perform redundant functions. Other sister pair genes may have undergone functional divergence after duplication and may have acquired mutually exclusive developmental roles.

The transcript levels of most of the *OsIAA* genes were upregulated by auxin treatment, although to varying degree as is evident from the data presented in Fig. 5b. The effect is more pronounced on *OsIAA9*, *14*, *19*, *20*, *24*, and *31*. Members of *Arabidopsis* gene family have also been shown to respond to exogenous IAA in a highly differential fashion with respect to dosage and time (Abel et al. 1994, 1995). The difference in kinetics between individual *Aux/IAA* genes is likely due to a variety of factors such as tissue-specific auxin reception, cell-type dependence and differential regulation of free auxin concentrations, or different modes of auxin-dependent transcriptional and posttranscriptional regulation.

Functions of rice *Aux/IAA* genes?

The *Aux/IAA* genes have been shown to play a critical role in plant growth and development, photomorphogenesis, light, and auxin signaling by biochemical and molecular genetic analysis of several gain-of-function and loss-of-function mutants in *Arabidopsis* (Liscum and Reed 2002). The results of structural analyses of rice Aux/IAA proteins will pave the way for their functional analysis. Our analysis showed that Aux/IAA proteins from *Arabidopsis* and rice are divided into two major groups. The Aux/IAA proteins classified in the same groups may have similar functions in events common to both monocot and dicot plants. The organ-specific differential expression suggests diverse and overlapping roles of these proteins during plant growth and development. The effect of light and auxin treatment on the transcript levels of *OsIAA* genes reflects their role in light and auxin signal transduction.

To further elucidate the functions of these genes, we investigated the phenotypes of rice *Tos17* retrotransposon insertion mutants (Miyao et al. 2003) of these genes with the aid of *Tos17* mutant panel database (http://www.tos.nias.affrc.go.jp/) using the BLAST program. We could enlist 14 insertion mutants corresponding to only two of the *OsIAA* genes (Supplementary Table S2). The phenotypes of the insertion mutants of these genes showed dwarfism, sterility, weakness, and altered yield. From these phenotypes, it can be speculated that these genes may play a role in different metabolic pathways and cellular processes influenced by light and auxin. A detailed analysis of these insertional mutants already available and RNAi strategy for the remaining *OsIAA* genes will greatly help in elucidation of the role of these genes and their functional validation.

References

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

Abel S, Oeller PW, Theologis A (1994) Early auxin-induced genes encode short-lived nuclear proteins. Proc Natl Acad Sci U S A 91:326–330

Abel S, Nguyen MD, Theologis A (1995) The *PS-IAA4/5*-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. J Mol Biol 251:533–549

Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O et al (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature 377:3–17

Ainley WM, Walker JC, Nagao RT, Key JL (1988) Sequence and characterization of two auxin-regulated genes from soybean. J Biol Chem 263:10658–10666

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

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

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–438

Conner TW, Goekjian VH, LaFayette PR, Key JL (1990) Structure and expression of two auxin-inducible genes from *Arabidopsis*. Plant Mol Biol 15:623–632

Dargeviciute A, Roux C, Decreux A, Sitbon F, Perrot-Rechenmann C (1998) Molecular cloning and expression of the early auxin-responsive *Aux/IAA* gene family in *Nicotiana tabacum*. Plant Cell Physiol 39:993–1002

Dharmasiri N, Dharmasiri S, Estelle M (2005a) The F-box protein TIR1 is an auxin receptor. Nature 435:441–445

Dharmasiri N, Dharmasiri S, Weijers D, Lechner E, Yamada M, Hobbie L, Ehrismann JS, Jurgens G, Estelle M (2005b) Plant development is regulated by a family of auxin receptor F box proteins. Dev Cell 9:109–119

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100

Gray WM, Estelle M (2000) Functions of the ubiquitin proteasome pathway in auxin response. Trends Biochem Sci 25:133–138

Gray WM, Kepinski S, Rouse D, Leyser O, Estelle O (2001) Auxin regulates SCF^{TIR1}-dependent degradation of Aux/IAA proteins. Nature 414:271–276

Guilfoyle TJ (1999) Auxin-regulated genes and promoters. In: Hooykaas PJJ, Hall MA, Libbenga KR (eds) Biochemistry and molecular biology of plant hormones. Elsevier, Amsterdam, The Netherlands

Guyot R, Keller B (2004) Ancestral genome duplication in rice. Genome 47:610–614

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

Jain M, Tyagi SB, Thakur JK, Tyagi AK, Khurana, JP (2004) Molecular characterization of a light-responsive gene, breast basic conserved protein 1 (*OsiBBC1*), encoding nuclear-localized protein homologous to ribosomal protein L13 from *Oryza sativa indica*. Biochim Biophys Acta 1676:182–192

Jain M, Kaur N, Tyagi AK, Khurana JP (2005) The auxin-responsive *GH3* gene family in rice (*Oryza sativa*). Funct Integr Genomics (in press)

Kepinski S, Leyser O (2004) Auxin-induced SCF^{TIR1}–Aux/IAA interaction involves stable modification of the SCF^{TIR1} complex. Proc Natl Acad Sci U S A 101:12381–12386

Kepinski S, Leyser O (2005) The *Arabidopsis* F-box protein TIR1 is an auxin receptor. Nature 435:446–451

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, Yamada H, Ooka H 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

Kolkman JA, Stemmer WPC (2001) Directed evolution of proteins by exon shuffling. Nat Biotechnol 19:423–428

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

Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka Y, Onosato K, Hirochika H (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, Walker L, Young J, Sonawala A, Timpte C, Estelle M, Reed JW (2000) *AXR2* encodes a member of the Aux/IAA protein family. Plant Physiol 123:563–573

Nebenfuhr A, White TJ, Lomax TL (2000) The *diageotropica* mutation alters auxin induction of a subset of the *Aux/IAA* gene family in tomato. Plant Mol Biol 44:73–84

Oeller PW, Keller JA, Parks JE, Silbert JE, Theologis A (1993) Structural characterization of the early indoleacetic acid-inducible genes *PS-IAA4/5* and *PS-IAA6* of pea (*Pisum sativum* L). J Mol Biol 233:789–798

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

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A 101:9903–9908

Phillips SEV (1994) The β-ribbon DNA recognition motif. Annu Rev Biophys Biomol Struct 23:671–701

Ramos JA, Zenser N, Leyser O, Callis J (2001) Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent. Plant Cell 13:2349–2360

Reed JW (2001) Roles and activities of Aux/IAA proteins in *Arabidopsis*. Trends Plant Sci 6:420–425

Reed JW, Elumalai RP, Chory J (1998) Suppressors of an *Arabidopsis thaliana phyB* mutation identify genes that control light signaling and hypocotyl elongation. Genetics 148:1295–1310

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

Rogg LE, Lasswell J, Bartel B (2001) A gain-of-function mutation in *IAA28* suppresses lateral root development. Plant Cell 13:465–480

Rouse D, Mackay P, Stirnberg P, Estelle M, Leyser O (1998) Changes in auxin response from mutations in an *Aux/IAA* gene. Science 279:1371–1373

Soh MS, Hong SH, Kim BC, Vizir I, Park H, Choi G, Hong MY, Chung Y-Y, Furuya M, Nam HG (1999) Regulation of both light- and auxin-mediated development by the *Arabidopsis IAA3/SHY2* gene. J Plant Biol 42:239–246

Stowe-Evans EL, Harper RM, Motchoulski AV, Liscum E (1998) NPH4, a conditional modulator of auxin-dependent differential growth responses in *Arabidopsis*. Plant Physiol 118:1265–1275

Thakur JK, Tyagi AK, Khurana JP (2001) *OsIAA1*, an *Aux/IAA* cDNA from rice, and changes in its expression as influenced by auxin and light. DNA Res 8:193–203

Thakur JK, Jain M, Tyagi AK, Khurana JP (2005) Exogenous auxin enhances the degradation of a light down-regulated and nuclear-localized OsiIAA1, an Aux/IAA protein from rice, via proteasome. Biochim Biophys Acta 1730:196–205

Theologis A, Huynh TV, Davis RW (1985) Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. J Mol Biol 183:53–68

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. Nucleic Acids Res 25:4876–4882

Tian Q, Reed J (1999) Control of auxin-regulated root development by *Arabidopsis thaliana SHY/IAA3* gene. Development 126: 711–721

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

Tiwari SB, Hagen G, Guilfoyle TJ (2004) Aux/IAA proteins contain a potent transcriptional repression domain. Plant Cell 16:533–543

Ulmasov T, Liu ZB, Hagen G, Guilfoyle TJ (1995) Composite structure of auxin response elements. Plant Cell 7:1611–1623

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

Walker JC, Key JL (1982) Isolation of cloned cDNAs to auxin-responsive poly(A) RNAs of elongating soybean hypocotyls. Proc Natl Acad Sci U S A 79:7185–7189

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

Worley CK, Zenser N, Ramos J, Rouse D, Leyser O, Theologis A, Callis J (2000) Degradation of Aux/IAA protein is essential for normal auxin signaling. Plant J 21:553–562

Yamamoto KT, Mori H, Imaseki H (1992) cDNA cloning of indole-3-acetic acid regulated genes: *Aux22* and *SAUR* from mung bean (*Vigna radiata*) hypocotyls tissue. Plant Cell Physiol 33:93–97

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92