Open access • Journal Article • DOI:10.1093/PCP/PCQ196

Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. — Source link

Akhter Most Sharoni, Mohammed Nuruzzaman, Kouji Satoh, Takumi Shimizu ...+5 more authors

Institutions: International Rice Research Institute

Published on: 01 Feb 2011 - Plant and Cell Physiology (Oxford University Press)

Topics: Rice tungro spherical virus, Gene family, Biotic stress, Rice stripe virus and Rice dwarf virus

PCP
PLANT & CELL PHYSIOLOGY

Gene Structures, Classification and Expression Models of the AP2/EREBP Transcription Factor Family in Rice

Akhter Most Sharoni[1], Mohammed Nuruzzaman[1], Kouji Satoh[1,2], Takumi Shimizu[2], Hiroaki Kondoh[1], Takahide Sasaya[2], Il-Ryong Choi[3], Toshihiro Omura[2] and Shoshi Kikuchi[1,*]

[1]Plant Genome Research Unit, Division of Genome and Biodiversity Research, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki, 305-8602 Japan
[2]Research Team for Vector-borne Plant Pathogens, National Agricultural Research Center, Tsukuba, Ibaraki, 305-8666 Japan
[3]Plant Breeding, Genetics, and Biotechnology Division, International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines
*Corresponding author: E-mail, skikuchi@nias.affrc.go.jp; Fax, +81-29-838-7007.
(Received November 17, 2010; Accepted December 7, 2010)

Regular Paper

We identified 163 *AP2/EREBP* (APETALA2/ethylene-responsive element-binding protein) genes in rice. We analyzed gene structures, phylogenies, domain duplication, genome localizations and expression profiles. Conserved amino acid residues and phylogeny construction using the AP2/ERF conserved domain sequence suggest that in rice the *OsAP2/EREBP* gene family can be classified broadly into four subfamilies [AP2, RAV (related to ABI3/VP1), DREB (dehydration-responsive element-binding protein) and ERF (ethylene-responsive factor)]. The chromosomal localizations of the *OsAP2/EREBP* genes indicated 20 segmental duplication events involving 40 genes; 58 redundant *OsAP2/EREBP* genes were involved in tandem duplication events. There were fewer introns after segmental duplication. We investigated expression profiles of this gene family under biotic stresses [infection with rice viruses such as rice stripe virus (RSV), rice tungro spherical virus (RTSV) and rice dwarf virus (RDV, three virus strains S, O and D84)], and various abiotic stresses. Symptoms of virus infection were more severe in RSV infection than in RTSV and RDV infection. Responses to biotic stresses are novel findings and these stresses enhance the ability to identify the best candidate genes for further functional analysis. The genes of subgroup B-5 were not induced under abiotic treatments whereas they were activated by the three RDV strains. None of the genes of subgroups A-3 were differentially expressed by any of the biotic stresses. Our 44K and 22K microarray results suggest that 53 and 52 non-redundant genes in this family were up-regulated in response to biotic and abiotic stresses, respectively. We further examined the stress responsiveness of most genes by reverse transcription–PCR. The study results should be useful in selecting candidate genes from specific subgroups for functional analysis.

Keywords: Abiotic stress • Biotic stress • Microarray • Phylogenetic analysis • Rice.

Abbreviations: AP2, APETALA 2; DAI, days after inoculation; DEG, differentially expressed gene; DREB, dehydration-responsive element-binding protein; ERF, ethylene-responsive factor; EREBP, ethylene-responsive element-binding protein; LEA, late-embryogenesis abundant; MAPK, mitogen-activated protein kinase; RAV, related to ABI3/VP1; RDV, rice dwarf virus; RSV, rice stripe virus; RT–PCR, reverse transcription–PCR; RTSV, rice tungro spherical virus; TF, transcription factor.

Introduction

Under different adverse environmental conditions such as drought, cold, high salinity, flood, submergence and pathogen attack, genes show specific expression patterns in accordance with their biological and physiological functions. Transcription factors (TFs) are important for maintaining expression of functional protein genes in the genome. Proteins enhance or repress the TFs of candidate genes in response to biotic or abiotic stimuli and developmental processes. In the plant kingdom, *AP2/EREBP* (APETALA2/ethylene-responsive element-binding protein) is a large family of TF genes. TFs encoded by *AP2/EREBP* genes contain the highly conserved AP2/ERF DNA-binding domain (Riechmann and Meyerowitz 1998). Jofuku et al. (1994) first reported this conserved DNA-binding domain in the homeotic gene AP2, and Ohme-Takagi and Shinshi (1995) found EREBPs in tobacco. Three proteins (Pti-4, Pti-5 and Pti-6) were identified by Zhou et al. (1997) as interacting with the tomato disease resistance protein Pto in yeast two-hybrid assays. Each of these proteins has a conserved domain that is very similar to the domain identified by Ohme-Takagi and Shinshi (1995); this domain defines a subgroup of the AP2 family that was later designated as the ethylene-responsive factor (ERF, or ethylene-responsive element-binding factor) subfamily. On the basis of the number of AP2/ERF domains encoded and the gene function,

Plant Cell Physiol. 52(2): 344–360 (2011) doi:10.1093/pcp/pcq196, available online at www.pcp.oxfordjournals.org

the *AP2/EREBP* gene family has been divided into four subfamilies: AP2, RAV (related to ABI3/VP1), dehydration-responsive element-binding protein (DREB) and ERF (Sakuma et al. 2002).

Proteins encoded by the *AP2/EREBP* supergene family are defined by the AP2/ERF domain, which consists of 50–60 amino acids, and these proteins are involved in a variety of regulatory mechanisms throughout the plant life cycle. Both the DREB and ERF subfamilies are of particular interest owing to their involvement in plant responses to stresses. TFs encoded by genes in the DREB subfamily play an important role in the resistance of plants to abiotic stresses by recognizing the dehydration-responsive element (DRE), which has a core motif of A/GCCG AC (Liu et al. 1998). ERF subfamily genes encode a large number of ERFs (Fujimoto et al. 2000), which have been shown to participate in the plant response to biotic stresses such as pathogens by recognizing the *cis*-acting element AGCCGCC, known as the GCC box (Hao et al. 1998). TFs encoded by some members of this gene family bind to both GCC and the DRE *cis*-element (Park et al. 2001, Sakuma et al. 2002). ERF and DREB subfamily TFs have been identified in various plant species, including rice (Cao et al. 2006), Arabidopsis (Liu et al. 1998) and cotton (Jin and Liu 2008). The roles of ERF and DREB proteins in the plant response to biotic and abiotic stress have also been extensively documented (Agarwal et al. 2006, Agarwal et al. 2010). In the genomes of Arabidopsis (Sakuma et al. 2002), grapevine (Jaillon et al. 2007) and poplar (Zhuang et al. 2008), 145, 132 and 200 AP2/ERF-related proteins, respectively, are encoded. Genetic and molecular approaches have been used in combination to characterize a series of regulatory genes of the *AP2/EREBP* gene family involved in many different pathways, including genes related to drought (Dubouzet et al. 2003), high salt concentration (Dubouzet et al. 2003), low temperature (Qin et al. 2007), diseases (Gutterson and Reuber 2004, Agarwal et al. 2006) and the control of flowering (Elliott et al. 1996). Overexpression of DREB1A (CBF3) or CBF1 (DREB1B) in transgenic Arabidopsis has been shown to induce strong expression of target stress-inducible genes and result in improved tolerance to drought, high salt and freezing (Jaglo-Ottosen et al. 1998, Liu et al. 1998, Gilmour et al. 2000, Quan et al. 2010, Zhu et al. 2010). Likewise, overexpression of some ERF genes enhances resistance to biotic and abiotic stresses in tobacco (Park et al. 2001). Recently, members of the RAV subfamily have been shown to be involved in the ethylene response (Alonso et al. 2003), the brassinosteroid response (Hu et al. 2004), and biotic and abiotic stress responses in pepper (Sohn et al. 2006). In general, the *AP2/EREBP* gene family has been well studied in Arabidopsis (Sakuma et al. 2002).

Characterization of *AP2/EREBP* family genes in rice can help us understand the molecular mechanisms of resistance to stress, and thus aid in the development of rice varieties, using transgenic technology, with greater tolerance to many adverse environments. Some *AP2/EREBP* family genes have been isolated from rice (Dubouzet et al. 2003), but most of their functions remain to be determined. Completion of the high-quality sequencing of the rice genome (International Rice Genome Sequencing Project 2005) has provided an excellent opportunity for genome-wide analysis of genes belonging to specific gene families. Here, we identified 163 *OsAP2/EREBP* genes in rice by database searches and classified these genes according to their homology with known genes. In this study, we describe subfamilies more specifically, and we present novel information on their role in the plant response to biotic stresses. Some specific subgroups of this gene family were differentially expressed under biotic and abiotic stress conditions. *OsAP2/EREBP* genes play an important role in the cross-talk of signaling pathways of different kinds of stresses. We analyzed the phylogenetic relationships of the *AP2/EREBP* genes in rice, and attempted the complete alignment of the ERF subfamily. In this study, we examined segmental and tandem expression of duplicated genes and exon and intron structures of *OsAP2/EREBP* genes. Furthermore, we studied the expression intensities of *OsAP2/EREBP* genes under different biotic and abiotic stresses using 44K and 22K microarray data. Taken together, our results should be helpful for determining the functions of each *OsAP2/EREBP* gene.

Results

Detection of *OsAP2/EREBP* genes

To identify *OsAP2/EREBP* genes in the rice genome, we searched several different public databases (see Materials and Methods), and used non-redundant sequences of the same genes from the different databases. First, we identified 167 putative *OsAP2/EREBP* genes. We then confirmed all protein sequences of the 167 *OsAP2/EREBP* genes by SMART and Pfam searches for the presence of the AP2/ERF domain. Four of the 167 genes (LOC_Os03g06920, LOC_Os05g28800, LOC_Os12g40960 and LOC_Os12g41040) that had a very small domain were excluded from further analysis, because this domain could not be used to construct an acceptable phylogenesis. Therefore, we used 163 *OsAP2/EREBP* genes for phylogenetic and expression analysis, which had corresponding locus IDs in the databases Michigan State University Rice Genome Annotation (MSU, http://rice.plantbiology.msu.edu/) and Database of Rice Transcription Factors (DRTF, http://drtf.cbi.pku.edu.cn; Gao et al. 2006). Detailed information about the genes and the structures of representative *OsAP2/EREBP* genes is presented in **Supplementary Table S1**. Hereafter, in this paper the 'LOC_' prefix is omitted from the MSU locus IDs.

Phylogenetic analysis and alignments

It was reported that at least 139 ERF members with a conserved AP2/ERF domain exist in the *japonica* rice genome (Nakano et al. 2006, Oh et al. 2009). A total of 163 *AP2/EREBP* genes were identified in rice from the different database searches. These were divided into four subfamilies based on their sequence similarities and the number of AP2/ERF domains in the encoded protein. Among them, (i) 24 genes were predicted

to encode two complete AP2/ERF domains and assigned to the AP2 subfamily; and (ii) five genes predicted to encode a single AP2/ERF domain and a single B3 domain were assigned to the RAV subfamily. Thus, *AP2/EREBP* genes having a double domain (domain a and domain b) were divided into subfamilies AP2 and RAV (**Fig. 1**). The phylogenetic relationships of the genes are the same in each subfamily; for example, in the AP2 subfamily, the domain of all the proteins duplicated before the gene duplication, whereas the RAV subfamily proteins showed domain duplication but no gene duplication. This is not surprising as the *AP2/EREBP* genes are an old family and functionally diverse. This is a common phenomenon in old gene families (Nuruzzaman et al. 2008). (iii) A large number of *AP2/EREBP* genes (134) encoding a single AP2/ERF domain were assigned to the DREB and ERF subfamilies. To clarify the phylogenetic relationships among ERF subfamily genes and to

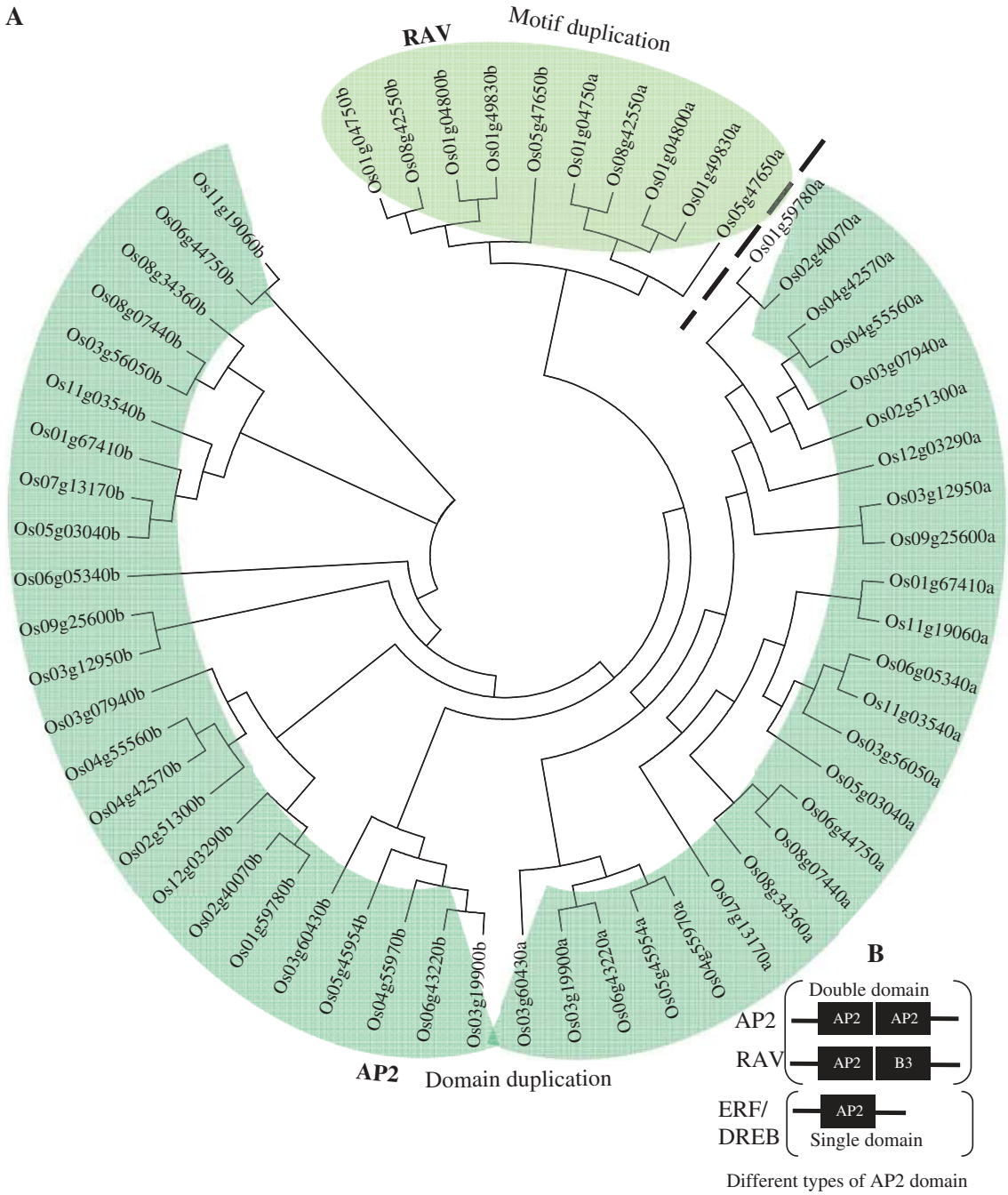

Fig. 1 (A) Evolutionary relationships among rice AP2/EREBP domain sequences. The unrooted tree was generated by the Neighbor–Joining method with the ClustalX program. Domains a and b were found in subfamilies RAV and AP2, e.g. in *Os01g59780a* and *Os01g59780b*, respectively. The RAV subfamily showed motif duplication, and the AP2 subfamily showed domain duplication. (B) Structure of *AP2/EREBP* genes.

infer the evolutionary history of this gene family, a combined phylogenetic tree was constructed by alignment of the 27 published AP2/ERF domain sequences (**Fig. 2**). The 134 genes were further classified into two subfamilies on the basis of similarity of the amino acid sequence of the AP2/ERF domains: 57 of them encode DREB/CBF-like proteins (subfamily DREB; **Fig. 2**, **Supplementary Fig. S1**), and 77 encode ERF-like proteins (subfamily ERF; **Fig. 2**, **Supplementary Fig. S1**). Of these, 127 genes contain a conserved WLG motif of the AP2/ERF domain (excepting seven genes: *Os02g45420*, *Os06g06540*, *Os06g47590*, *Os07g22730*, *Os08g43200*, *Os09g26420* and *Os10g26590*) (**Supplementary Fig. S1**). In all of the proteins encoded by genes of subfamily DREB, position 14 is normally valine (important), whereas position 19 is glutamic acid (not important) in the DREB/CBF DNA-binding domain (Sakuma et al. 2002). Most of the genes of subfamily B have alanine in the A14 position and aspartic acid in the D19 position in the ERF-DNA binding domain. We found that all of the genes of subfamilies DREB and ERF contain alanine in the A37 position; an α-helix in the A37 position has been demonstrated to be essential for binding with the DRE and GCC box *cis*-elements (Liu et al. 2006), and Allen et al. (1998) have described the three-dimensional structure (β-sheets and α-helices) of the ERF domain. The proteins of subfamily DREB/CBF were divided into six subgroups. Among these subgroups A-4 and A-5 contained the greatest numbers of genes. Likewise, the proteins of subfamily ERF were divided into seven subgroups and the largest numbers of genes were assigned to subgroups B-1 and B-3 (**Fig. 2**).

Chromosomal locations of *AP2/EREBP* genes

To examine the genomic distribution of *AP2/EREBP* genes on rice chromosomes, we identified their positions by MSU database search. A total of 163 rice *AP2/EREBP* genes could be

Fig. 2 Relationships among rice AP2/EREBP proteins after alignment with ClustalW. OsAP2/EREBP proteins were allocated to two distinct subfamilies (DREB, subgroups A-1 to A-6; ERF, subgroups B-1 to B-7).

localized on the 12 chromosomes with an obviously uneven distribution. *OsAP2/EREBP* genes were present in all regions on a single chromosome (i.e. at the telomeric ends, near the centromere and in between) and could be distributed individually or in clusters (**Fig. 3**). Chromosome 2 had the largest number (25) of *OsAP2/EREBP* genes, and chromosomes 4 and 6 had 20 each. Only four *OsAP2/EREBP* genes were identified on each of chromosomes 11 and 12. Interestingly, all four *OsAP2/EREBP* genes on chromosome 11 were found on the short arm, whereas three of the four *OsAP2/EREBP* genes on chromosome 12 were on the long arm. Two *OsAP2/EREBP* genes (*Os03g22770* and *Os07g22730*) encoding proteins having only one AP2/ERF domain were positioned around the centromere on chromosome 7 and were arranged as a tandem duplication. Additionally, only one gene (*Os08g45110*) was located near a telomeric region. There were <10 *OsAP2/EREBP* genes on each of chromosomes 10, 11 and 12.

In plants, gene numbers are expanded by segmental and tandem duplication in gene families (Cannon et al. 2004). To elucidate the potential mechanism of evolution of the *OsAP2/EREBP* gene family, we analyzed segmental and tandem duplication events, identifying 20 segmental duplication events in rice (**Fig. 3**, **Supplementary Table S2**). One segmental duplication was found in the RAV subfamily, and 19 segmental duplications were found in the DREB and ERF subfamilies. Neither segment was duplicated in subfamily AP2 in rice. We also found 29 tandemly duplicated genes in rice (**Fig. 3**, **Supplementary Table S3**). In rice, 58 *OsAP2/EREBP* genes were involved in tandem duplications, consisting of 29 pairs or 22 clusters in rice (**Fig. 3**, **Supplementary Table S3**). The number of *OsAP2/EREBP* genes arranged in tandem repeats varied from two to four, and these sets of genes showed the same orientation on the chromosomes, with the exception of seven pairs (*Os02g43970* and *Os02g43820*; *Os03g07830* and *Os03g07940*; *Os04g46220* and *Os04g46240*; *Os04g55520* and *Os04g55560*; *Os06g11860* and *Os06g11940*; *Os08g44960* and *Os08g45110*; and *Os09g39810* and *Os09g39850*). The degree of homology in the protein sequences of these genes is shown in **Supplementary Tables S2** and **S3**.

We analyzed the intron and exon structures of *OsAP2/EREBP* genes belonging to the four subfamilies based on the phylogenetic tree. We found no obvious differences among the different subgroups of the DREB and ERF subfamilies, except for subgroups A-5, B-1 and B-6 (**Supplementary Fig. S2**). Both the locations and the numbers of introns and exons of the *OsAP2/EREBP* gene family varied among genes. The highest

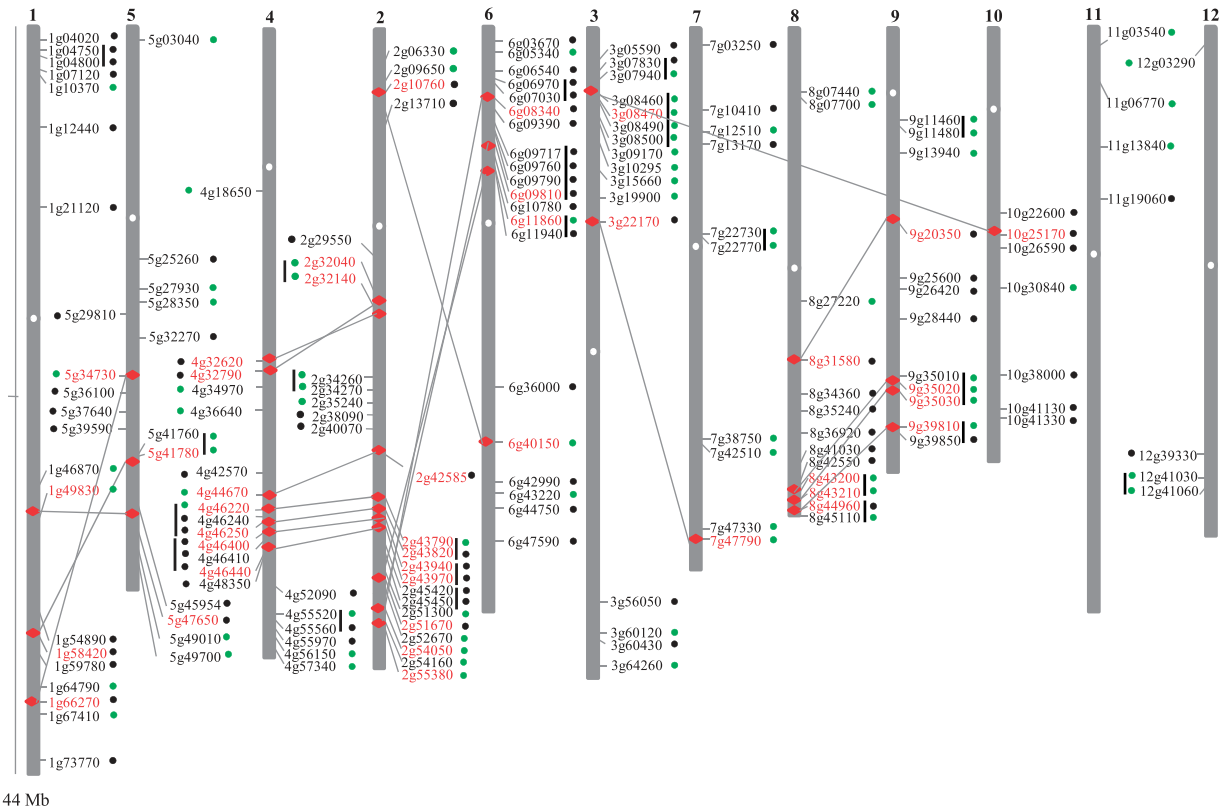

Fig. 3 Locations of *AP2/EREBP* genes on the 12 rice chromosomes. Chromosome numbers are indicated at the top of each bar. The filled white circle on each chromosome (vertical bars) shows the putative position of the centromere. Genes with open reading frames with opposite orientations are marked on the chromosome (filled black circles represent downward and filled green circles indicate upward orientations). Straight lines connect the *OsAP2/EREBP* genes present on duplicated chromosomal segments, and tandemly duplicated gene clusters are marked by vertical black bars. The scale is in megabases (Mb).

numbers of introns and exons were identified in subfamily AP2. In subfamily AP2, most coding sequences of *OsAP2/EREBP* genes were disrupted by introns, with the number of introns varying from five to 11 (**Supplementary Fig. S2**). Most of the genes in subfamilies RAV and ERF had no introns. In rice, some researchers have reported that the rate of intron loss is faster than the rate of intron gain after segmental duplication (Lin et al. 2006, Nuruzzaman et al. 2008). Therefore, we speculate that the genes in subfamilies RAV and ERF may be young members of this gene family; in each subfamily, those genes with more introns may be the original genes in that subfamily, or they may have diversified from subfamily AP2 (**Supplementary Fig. S2**). The mechanisms of intron gain and loss are not yet clear, however (Lin et al. 2006).

Comparison of expression profiles of duplicated *OsAP2/EREBP* genes

In the course of evolution, there are three possible novel functions of gene duplication: non-functionalization, neofunctionalization and subfunctionalization (Lynch and Conery 2000). Divergence of gene expression plays a very important role in the preservation of duplicated genes. In this study, we examined the expression patterns of segmentally and tandemly duplicated genes under different stress conditions. Probes were matched to 11 of the 20 genes located in segmentally duplicated regions. Nine pairs of genes showed highly similar expression patterns under most of the tested stresses, indicating subfunctionalization after the duplication events (**Fig. 4A**). Likewise, we identified 22 clusters of tandemly duplicated *OsAP2/EREBP* genes (**Supplementary Table S3**). Among them, only eight clusters of the gene probe set were found in our 22K microarray data. Five clusters of genes showed highly similar expression intensities, which may indicate subfunctionalization. The expression patterns of four clusters of tandemly duplicated genes (such as *Os02g32040* and *Os02g32140*) were dissimilar, which may indicate neofunctionalization (**Fig. 4B**). Neofunctionalization occurs when a neofunctionalized allele is fixed in one of the duplicated genes and it asserts that after duplication one daughter gene retains the ancestral function while the other acquires new functions. The *Os02g32040* gene showed down-regulation under cold, drought and submergence while the *Os02g32140* gene was up-regulated under all abiotic treatments. On the basis of various roles, we may predict that these duplicated genes have diverse functions under stress conditions.

Gene expression profiles under biotic and abiotic treatments

We used microarray analyses to investigate the response of the *OsAP2/EREBP* genes to both biotic and abiotic stresses. We investigated the expression profiles of *OsAP2/EREBP* genes in rice seedlings infected with rice stripe virus (RSV), rice tungro spherical virus (RTSV) and rice dwarf virus (RDV; three virus strains S, O and D84), and expressed the results as fold changes

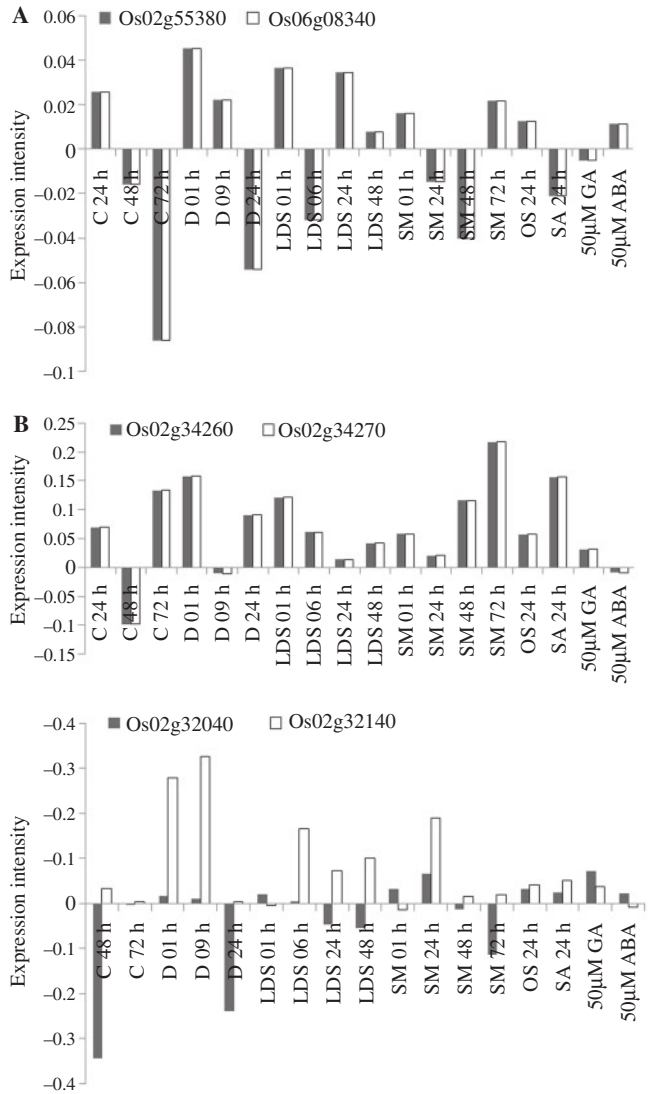

Fig. 4 Examples of expression patterns of *OsAP2/EREBP* genes found in duplicated regions of the rice genome. (A) Expression patterns of two *OsAP2/EREBP* genes found in segmentally duplicated regions. (B) Expression patterns of two pairs of tandemly duplicated *OsAP2/EREBP* genes. The different stresses are shown on the *x*-axis and the expression intensity on the *y*-axis. Gene names are indicated at the top of each graph.

with respect to the controls. Multiplication and motion of viruses might be inhibited by the plant defense system (Kitanaga et al. 2006). We found 37 non-redundant genes differentially expressed under RSV infection at 3, 6, 9 and 12 days after inoculation (DAI) (**Supplementary Table S4**). Among these, 13 were up-regulated and 24 were down-regulated. The number of up-regulated genes was highest at 12 DAI, followed by 3, 6 and 9 DAI (listed in decreasing order) under RSV infection, although many genes involved in gene silencing were activated at 12 DAI (**Fig. 5A**). This may suggest that the gene silencing system in the host was not activated in a timely manner to suppress RSV replication. Recent research has

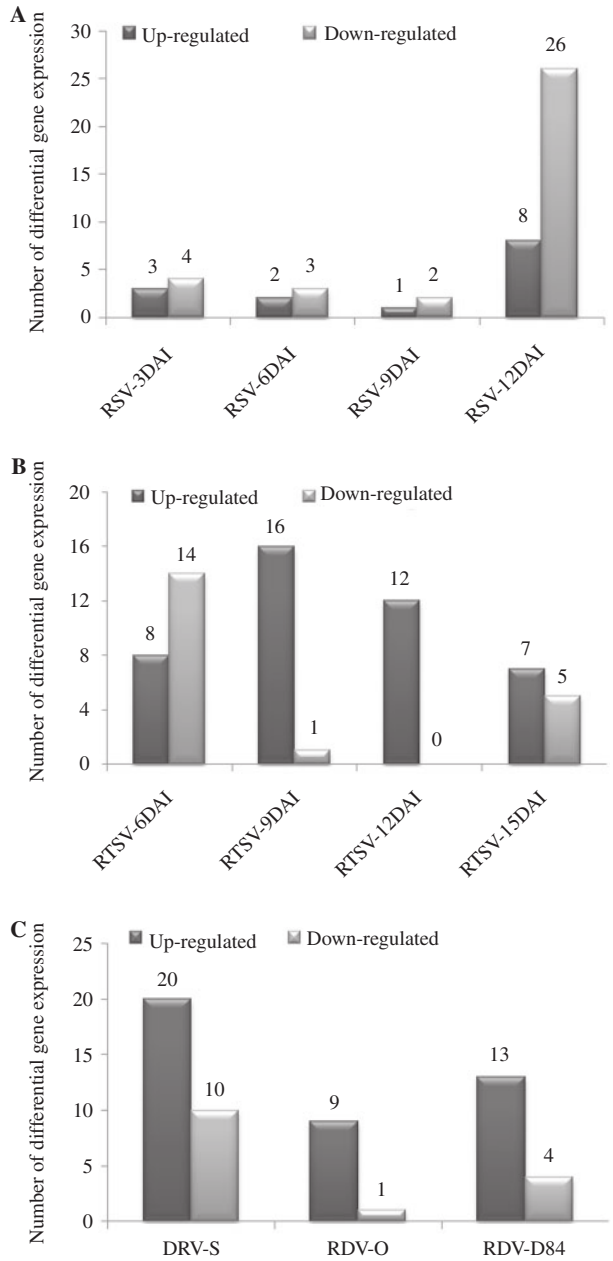

Fig. 5 Number of differentially expressed genes at different time points: (A) under RSV infection at 3, 6, 9 and 12 days after inoculation (DAI); (B) under RTSV infection at 6, 9, 12 and 15 DAI. Different DAI are shown on the *x*-axis, and the number of differentially expressed genes on the *y*-axis. (C) The differentially expressed genes were expressed under RDV infection with three strains (S, O and D84). The number of strains is shown on the *x*-axis, and the differentially expressed genes on the *y*-axis.

shown that the p3 (*NS3*) gene of RSV encodes a gene silencing suppressor that inhibits local and systemic gene silencing (Levy et al. 2008, Xiong et al. 2008, Xiong et al. 2009). We also found 46 non-redundant differentially expressed genes (DEGs) under RTSV infection at 6, 9, 12 and 15 DAI (**Fig. 5B**, **Supplementary Table S5**). Among these, 28 were up-regulated and 18 were down-regulated. Only one gene (*Os09g35030*) was found to be up-regulated at all time points. The number of up-regulated genes was greatest at 9 DAI and smallest at 12, 6 and 15 DAI (in that order) under RTSV infection (**Fig. 5B**). Out of 10 genes of subgroup A-1, eight were differentially expressed under RTSV. Six genes of subgroup B-1 were up- or down-regulated under RTSV infection. Interestingly, only two genes (*Os01g12440* and *Os09g35020*) were very commonly up-regulated under both RSV and RTSV infection. Furthermore we observed 54 non-redundant DEGs under RDV infection with the three strains S, O and D84. The number of genes activated was highest in the RDV S strain and then decreased in the order of strain D84 and strain O (**Fig. 5C**, **Supplementary Table S6**). In the case of RSV, fewer genes were up-regulated than were down-regulated, while in the case of RTSV and RDV, more genes were up-regulated than were down-regulated. Defense systems are activated in response to RTSV and RDV but in the case of RSV the host was nearly dead. In the phylogenetic tree, we found that expression of 46 non-redundant genes of subfamilies DREB and ERF was induced by all biotic stresses, whereas four genes from subfamily RAV were up-regulated under both stresses (RSV and RTSV) and three genes of this subfamily were induced by the RDV S strain. Only two genes from the AP2 subfamily were up-regulated under RSV infection and all of the genes in this subfamily were not differentially expressed by RDV treatment except Os04g42570. Out of 12 genes of subgroup B-4, Os02g32140 was activated by RSV and the RDV S strain, respectively. Most of the genes of subgroups A-2, A-3 and B-6 were not differentially expressed by all biotic treatments. Interestingly, seven of 10 genes in subgroup A-1 were up-regulated under RTSV infection, whereas only one gene of subgroup A-1 showed high expression under RSV infection. Likewise, subgroups B-1 and B-5 of the ERF subfamily genes were preferentially induced under RTSV, RSV and RDV, respectively (**Table 1**). Comparatively speaking, the gene responses of the *OsAP2/EREBP* family were higher under RDV than under RTSV and RSV infections. Some TFs are known to act specifically on transcriptional regulation of genes responding to biotic and abiotic stresses (Shimono et al. 2007). The result of global gene expression analysis for plants infected with these three viruses suggested intercorrelations among the numbers of DEGs, the degrees of gene response, the symptom severity and the accumulation of virus in plants.

The expression levels of 70 non-redundant *OsAP2/EREBP* genes were up- or down-regulated under different abiotic stresses (cold, drought, laid-down (complete) submergence, submergence, osmotic stress, salt and hormone stress) compared with their expression in control seedlings (**Supplementary Table S7**). Under the different stress conditions, more than 60% of the genes were up-regulated. The greatest number of genes (33) were up-regulated under laid-down submergence stress, and the lowest number (1) was up-regulated under osmotic or salt stress. Under the ABA (hormone) treatment, none of the genes was up- or down-regulated. Thirteen genes were up-regulated by at least three abiotic treatments and,

Table 1 Genes preferentially induced in specific subgroups under (A) biotic and (B) abiotic stresses conditions

(A)

Subgroup	No. of genes	RTSV		RSV		RDV					
						S		O		D84	
		Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
A-1	**10**	**7**	**1**	1	4	1	0	2	0	0	0
A-2	**6**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
A-3	**5**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
A-4	12	1	4	1	4	3	2	1	1	1	2
A-5	15	2	2	1	3	3	2	0	0	2	0
A-6	9	1	1	0	1	1	1	0	0	1	0
B-1	**18**	**5**	**1**	**2**	**3**	1	2	1	0	1	0
B-2	12	2	0	1	3	1	0	0	0	1	0
B-3	19	2	2	0	0	2	2	1	0	2	1
B-4	12	2	4	1	1	1	0	0	0	0	0
B-5	**4**	1	0	1	1	**2**	**0**	**3**	**0**	**2**	**0**
B-6	**3**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**
B-7	9	1	1	0	2	1	1	0	0	2	0
AP2	24	0	1	2	3	1	0	1	0	0	1
RAV	5	3	0	1	0	3	0	0	0	1	0

(B)

Subgroup	Induced genes	Treatments	Total	Subgroup	Induced genes	Treatments	Total
A-1	2/10	C, LDS, SM	3	B-1	4/18	C, D, LDS, SM	4
A-2	3/6	LDS, SM	2	B-2	4/12	LDS, SM	2
A-3	**4/5**	**C, D, LDS, SM, GA**	**5**	**B-3**	**7/19**	**C, D, LDS, SM**	**4**
A-4	4/12	C, D, LDS, SM	4	B-4	5/12	C, D, LDS, SM	4
A-5	3/15	C, D, LDS, SM, GA	5	**B-5**	**0/4**	**0**	**0**
A-6	1/9	LDS, SM	2	B-6	2/3	SA	1
AP2	8/24	C, LDS, SM, GA	4	**B-7**	**3/9**	**C, LDS, SM, OS**	**3**
RAV	2/5	LDS, SM	2				

C, cold; D, drought; LDS, laid-down submergence; SM, submergence; GA, gibberellic acid; SA, salt.
Numbers in bold of specific subgroups indicate up- or down-regulation under different stress conditions.

comparatively, subgroups A-3 and A-4 of subfamily DREB were preferentially induced by more than three abiotic treatments (**Table 2**). Five genes showed both up- and down-regulation at different time points of laid-down submergence stress, among which two genes (*Os04g57340* and *Os09g35010*) had identical patterns of up- and down-regulation under both submergence stresses (**Supplementary Table S7**). Altogether, expression of 10 genes of subfamilies AP2 and RAV was influenced by one or more abiotic treatments, and six of the 10 genes were affected by both laid-down submergence and submergence. A small number of genes of the different A and B subgroups were up-regulated under different abiotic stresses, whereas a number of genes belonging to the A-3 and B-3 subgroups were up-regulated similarly in response to various abiotic stresses (**Supplementary Table S7**).

The *Os01g04750* and *Os01g04800* genes of the RAV subfamily, *Os05g45954* of the AP2 subfamily, *Os05g25260* of the A-2 subgroup, *Os04g34970* of the A-4 subgroup, *Os02g43970* of the A-5 subgroup, *Os04g48350* of the B-1 subgroup, *Os09g11460* of the B-2 subgroup and *Os09g28440* of the B-4 subgroup were up-regulated under different abiotic and biotic stresses. All the genes of subgroup B-5 were not induced under abiotic treatments, while only one gene was up-regulated by RSV and RTSV infections, but most of the genes were activated by the three RDV strains. On the other hand, subgroup A-1 of the DREB subfamily comprises several genes that regulate disease resistance pathways, as inferred from the increased resistance to pathogens conferred by their overexpression under the control of a constitutive promoter (Shin et al. 2002). Most of the genes of subgroup A-3 of the DREB subfamily were up-regulated under five different abiotic treatments, while several genes of subgroup B-3 of the ERF subfamily showed up-regulation under four abiotic treatments. Out of the three genes, two showed preferential expression under salt treatment; these are assigned to the B-6 subgroup of the ERF subfamily. On a broad scale, our findings might improve our knowledge of gene responses in research on different biotic and abiotic stresses. Some researchers have reported that

Table 2 The 13 genes expressed (up-regulated) by at least three abiotic treatments

Up-regulated gene	Cold	Drought	Laid-down submergence	Submergence	Osmotic	Salt	50 μM GA	50 μM ABA	Subgroup
Os01g07120	1	1	1	1					**A4**
Os02g09650		1	1	1					**A3**
Os02g43970			1	1			1		A5
Os03g08470		1	1	1					B3
Os04g32620	1	1	1						B1
Os04g34970	1	1	1						**A4**
Os06g07030		1	1	1			1		**A3**
Os06g09760	1	1	1	1					A5
Os06g11940	1	1	1	1					**A4**
Os06g36000			1	1			1		**A3**
Os07g12510		1	1	1					B3
Os09g13940	1		1	1	1				B7
Os11g06770		1	1	1					B4

Genes up-regulated ≥1.5-fold were assigned a value of 1. Numbers in bold for specific subgroups indicate the highest numbers of genes induced by different abiotic stresses.

overexpression of some TF genes from different families such as NAC (Jeong et al. 2010), DREB (Haake et al. 2002), bZIP (Uno et al. 2000) and zinc finger (Mukhopadhyay et al. 2004) improved tolerance to various abiotic stresses. AP2/EREBP transcription factors have a very wide range of functions in specific plant species. Taken together, these results increase our knowledge of the involvement of OsAP2/EREBP TFs in plant resistance and show that certain members or subgroups of this gene family are involved in resistance to abiotic and biotic stresses.

Discussion

Domain/motif duplication and gene expansion

We have presented a detailed, specific phylogenetic analysis of AP2/EREBP proteins based on their DNA-binding domains. The subgroups obtained by our analysis fit well with known AP2/EREBP function classes. Proteins with similar domains may have the same or similar biological functions (Lin et al. 2007). For example, the different subgroups of ERF or DREB (**Fig. 2**) represent five well-known functions of AP2/EREBP proteins: regulation of seed development, cell division and expansion, organ formation or initiation, response to biotic or abiotic stress and cross-talk between different signaling pathways. Our goals in this study were to (i) annotate the *OsAP2/EREBP* gene family; (ii) predict domain and motif duplication of the subfamilies AP2 and RAV; (iii) provide a new avenue to predict the group-based classification in plants; (iv) determine the expression patterns of duplicated genes; (v) assess the number of genes responding to different stresses (responses to biotic stresses, in particular, are a novel contribution); and (vi) select the best candidate genes for further functional analysis. There are some inconsistencies in *OsAP2/EREBP* gene classification between our results and previous findings. *OsAP2/EREBP* genes play vital roles in various developmental processes, including signaling, stress responses and plant defenses. In combination, the processes of gene duplication, nucleotide substitution, domain duplication and intron/exon shuffling can generate a complex set of related genes that may differ substantially in their expression patterns and functions. Gene duplication is one of the major evolutionary mechanisms leading to functional diversification and speciation (Lynch and Conery 2000). Likewise, portions of genes, such as specific exons or those encoding specific domains, may be duplicated within a gene, further complicating the history of gene evolution. Just as we compare gene trees with species trees to co-infer patterns of organismal and genetic evolution, we can compare gene trees with 'domain trees' to investigate patterns of domain evolution.

We found 163 *OsAP2/EREBP* genes in the rice genome that could be classified into different subfamilies distributed on all 12 chromosomes, and Nakano et al. (2006) have reported that 147 *AtAP2/EREBP* genes are distributed on the Arabidopsis chromosomes. Gene duplication is a primary driving force of new gene functions in the evolution of genetic systems and genomes (Moore and Purugganan 2003). Cannon et al. (2004) reported that segmental duplications occur commonly in the slowly evolving *MYB* gene family. Tandem duplications in local genomic clusters with low levels of retention of segmental duplications are common in the large NBS-LRR disease resistance gene family (Cannon et al. 2004). The results of the present study revealed more tandemly duplicated genes than duplicated chromosomal segments in the *AP2/EREBP* gene family in rice. Therefore, we consider the number of *OsAP2/EREBP* family genes to have increased rapidly during the course of evolution, and the tandem duplications of chromosomal regions to have played a key role in the expansion of this family. This phenomenon has also been found in the F-box family of rice genes (Jain et al. 2007). The number of segmental

duplication genes in Arabidopsis was around twice the number in rice, whereas there were about 2.5 times the number of tandemly duplicated genes in rice as in Arabidopsis (Nakano et al. 2006). Gene duplication of this gene family is thus species specific.

Gene function under biotic and abiotic stress conditions

OsAP2/EREBP genes play crucial roles in various developmental processes, including signaling, stress responses and plant defenses. ERF and DREB proteins constitute subfamilies of the AP2/EREBP TFs that are distinctive to plants, and they play significant roles in plant resistance to biotic and abiotic stresses. In the *AP2/EREBP* gene family, 53 non-redundant genes were up-regulated under all tested biotic stresses (RSV, RTSV and RDV); of these 53 genes, the ERF and DREB subfamilies comprised 46 non-redundant genes. Comparing symptoms induced by the virus infections (RSV, RTSV and RDV) and changes of gene expression of *AP2/EREBP* genes suggested that members of this family contribute heavily to the defense mechanism against virus infection. The global gene response and the defense system against RSV infection have been described (Satoh et al. 2010). Significantly, several ERF TFs that confer enhanced disease resistance when overexpressed, such as *ERF1*, *Pti4* and *AtERF1*, are transcriptionally regulated by pathogens and by the plant hormones ethylene and jasmonic acid (Fujimoto et al. 2000, Gutterson and Reuber 2004). Post-transcriptional regulation of *ERF* genes by phosphorylation may be a significant form of regulation. The *OsEREBP1* gene of rice has been shown to be phosphorylated (Fujimoto et al. 2000) by pathogen-induced mitogen-activated protein kinase (MAPK). The basic domain leucine-zipper (bZIP), MYC, MYB, NAC and WRKY-binding (WRKY) ERF TFs are important families of stress-responsive TFs. WRKY proteins, which are unique to plants, contain either one or two WRKY domains and play a key role in regulating pathogen-induced defense responses (Dong et al. 2003). Interestingly, all of the AP2/EREBP sequences that are known to play a role in disease resistance responses are in one AP2/EREBP subfamily. The AP2/EREBP TF family has undergone extensive expansion through gene duplication events. Like that within other TF families (e.g. the MYB) family (Dias et al. 2003), the rate of evolution within the *AP2/EREBP* gene family differs significantly between the conserved domains and other regions of the proteins. Evidence from both the MYB family and the basic helix–loop–helix family (Atchley et al. 1994) suggests that this differential sequence conservation reflects differing structural constraints on the functions of the conserved domain and other domains. We found that the expression levels in our results were very similar to the intensities of the microarray data, and the transcription of many genes encoding TFs was induced at 12 and 9 DAI under the RSV and RTSV biotic stresses, respectively. Similar expression models have been published for genes encoding proteins containing protein kinase, leucine-rich, NB-ARC and EF-hand domains, which might function in signal transduction for defense systems (Tameling and Baulcombe 2007, Li et al. 2009). Expression of the genes tested in this study was induced in the middle and late phases of infection, when plants showed obvious symptoms (**Fig. 5**). Some host defense systems have been associated with genes for TFs in the WRKY family (*OsWRKY45*; Shimono et al. 2007). We studied large numbers of differentially expressed genes by reverse transcription–PCR (RT–PCR) analysis to determine mRNA expression levels from our 44K microarray data (**Figs. 6**, **7**). We found that the expression levels in our results were very similar to the intensities of the microarray data, and the transcription of many genes encoding TFs was affected on different days after inoculation under both biotic stresses. The inactivation of defense systems in rice plants in the early infection stage probably promotes the propagation of RSV and RTSV.

Subfamily DREB has been demonstrated to play a key role in the resistance of plants to abiotic stresses such as cold, high salinity and drought (Shinozaki et al. 2000). Also some DREBs, such as DBF1, DBF2 (Kizis and Pages 2002) and CBF4 (Haake et al. 2002), are responsive to ABA. DREB/ERF proteins share a conserved 50–60 amino acid domain (the AP2/ERF domain) that binds to two *cis*-elements: the GCC box, found in many pathogen-related gene promoters conferring ethylene responsiveness (Gu et al. 2000), and the C-repeat CRT/DRE element motif, involved in the expression of cold- and dehydration-responsive genes. *OsDREB1A–1D* genes are induced under different abiotic stresses in rice (Dubouzet et al. 2003). Similarly, in many studies, overexpression of stress-inducible DREB transcription factors activates the expression of many target genes having DRE elements in their promoters, and the resulting transgenic plants show improved stress tolerance (**Table 3**). The present study showed in detail that expression of rice *AP2/EREBP* genes is induced by abiotic as well as biotic stresses. From the various studies summarized in **Table 3**, it is clear that DREB proteins are important transcription factors regulating abiotic stress-related genes and that they play a critical role in imparting stress endurance to plants. To obtain an overview of the expression pattern changes in the rice *AP2/EREBP* gene family under different stress conditions, we examined responses in 10-day-old seedlings and 30-day-old calli to cold, drought, laid-down submergence, submergence, osmotic, salt and hormone stresses. The expression patterns of the *OsAP2/EREBP* genes may provide clues to determine the function of each gene under both biotic and abiotic conditions. A total of 70 genes were differentially expressed under at least one of these abiotic stress conditions, among which 52 were up-regulated by at least one time point under one or more of the treatments and some subgroups (i.e. A-1, A-3 and B-5) were discriminatorily induced by abiotic or biotic stress. For confirmation, we carried out an RT–PCR analysis to determine mRNA expression levels; all genes had strikingly higher or lower expression profiles in seedlings than in controls at one or more time points (**Fig. 8A–D**). Similarly, MeV cluster analysis based on log 10 ratio values showed that *OsAP2/EREBP* genes had very

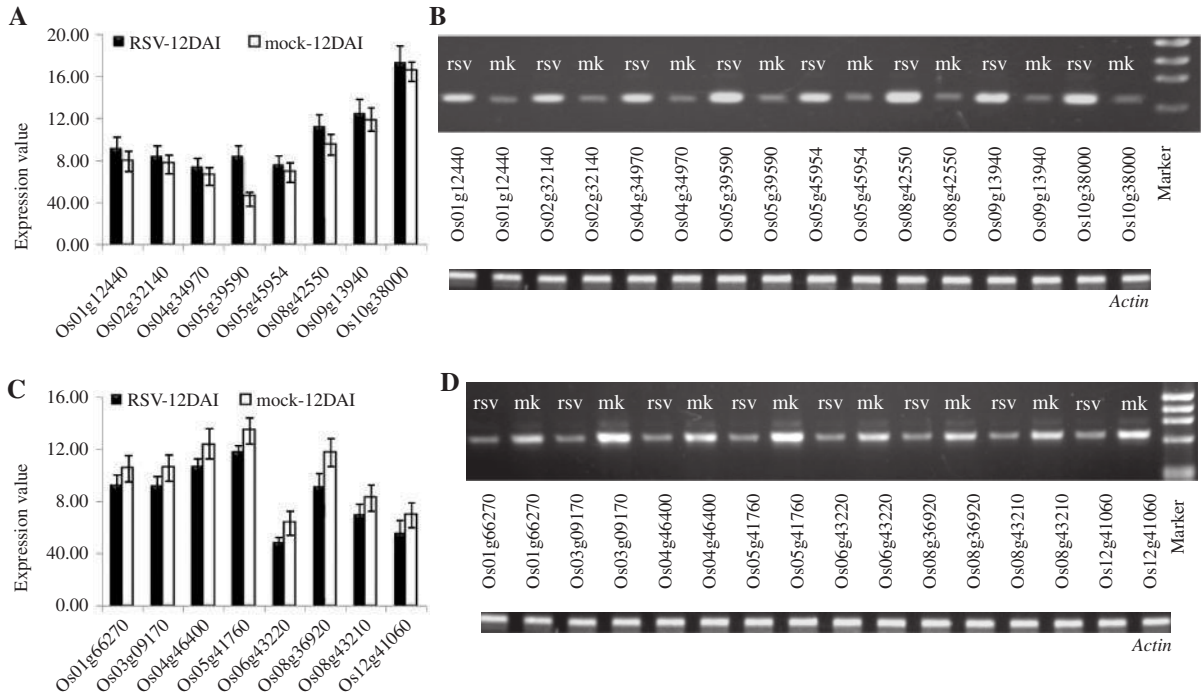

Fig. 6 (A) Histogram of microarray data for eight *OsAP2/EREBP* genes differentially expressed under RSV infection. (B) RT–PCR analysis of eight genes that showed higher expression than the control under RSV infection at 12 DAI. (C) Histogram of microarray data for eight genes down-regulated under RSV infection. (D) RT–PCR analysis of eight genes that showed lower expression than the control under RSV infection at 12 DAI. *Actin* gene expression was used as the internal control; mk, mock (control); rsv, RSV-infected plants sampled 12 DAI. The differentially expressed genes are shown along the *x*-axis of the histograms, and the log2 ratios of expression values are shown on the *y*-axis.

diverse expression patterns (**Fig. 9**, **Supplementary Table S7**). In order to gather more details on the expression profiles under different stresses, we included three biotic and eight abiotic stresses not considered before.

Functional and regulatory proteins are involved in gene expression in response to stress conditions. Functional proteins include membrane proteins that control water movement through membranes, proteins [such as late-embryogenesis abundant (LEA) proteins, osmotin and mRNA-binding proteins]. LEA proteins, for example, improve the drought or high-salinity tolerance of plants. Regulatory proteins (e.g. bZIP, MYC, MYB and DREB), protein kinases (e.g. MAPK, CDP kinase and receptor protein kinase) and proteinases are involved in the regulation of signal transduction and gene expression. TFs, together with *cis*-elements, function in the promoter region of different stress-related genes, and overexpression of these genes may improve tolerance to abiotic or biotic stress. A microarray analysis in Arabidopsis has shown that there are several pathways that independently respond to abiotic stress, and one such important pathway involves the DREB regulon (Fowler and Thomashow 2002). By a transcriptome analysis, genes induced by different stresses can be identified from the functions of their products.

In conclusion, taken together, our data showed gene and domain/motif duplication in subfamilies AP2 and RAV. Segmental and tandem duplications might have been the main contributors to the expansion of this family in rice. Some subgroups showed a high level of expression in abiotic stress and biotic stress, suggesting that they might have functional divergence. Interestingly, two genes were common and up-regulated under two biotic stresses (RSV and RTSV), whereas three genes were induced by RDV infection (all strains). Thirteen genes were activated by at least three abiotic stress conditions. *OsAP2/EREBP* genes showed temporal and spatial patterns of expression under different kinds of stress conditions. Specific subgroups of this gene family provide a new avenue for determining the best candidate genes for functional analysis. Overexpression, knockdown or mutagenesis, and promoter analyses of selected members of this gene family are underway in our laboratory so that we can accurately determine molecular pathways in the *OsAP2/EREBP* gene family.

Materials and Methods

Compilation and classification of *OsAP2/EREBP* gene family members

We searched for AP2/EREBP TF genes in rice by searching various databases: the DRTF database (http://drtf.cbi.pku.cbi.pku.edu.cn; Gao et al. 2006), the MSU database (http://rice.plantbiology.msu.edu/), the Rice Transcription Factor Database (http://ricetfdb.bio.uni-potsdam. de/v3.0; Riano-Pachon et al. 2007),

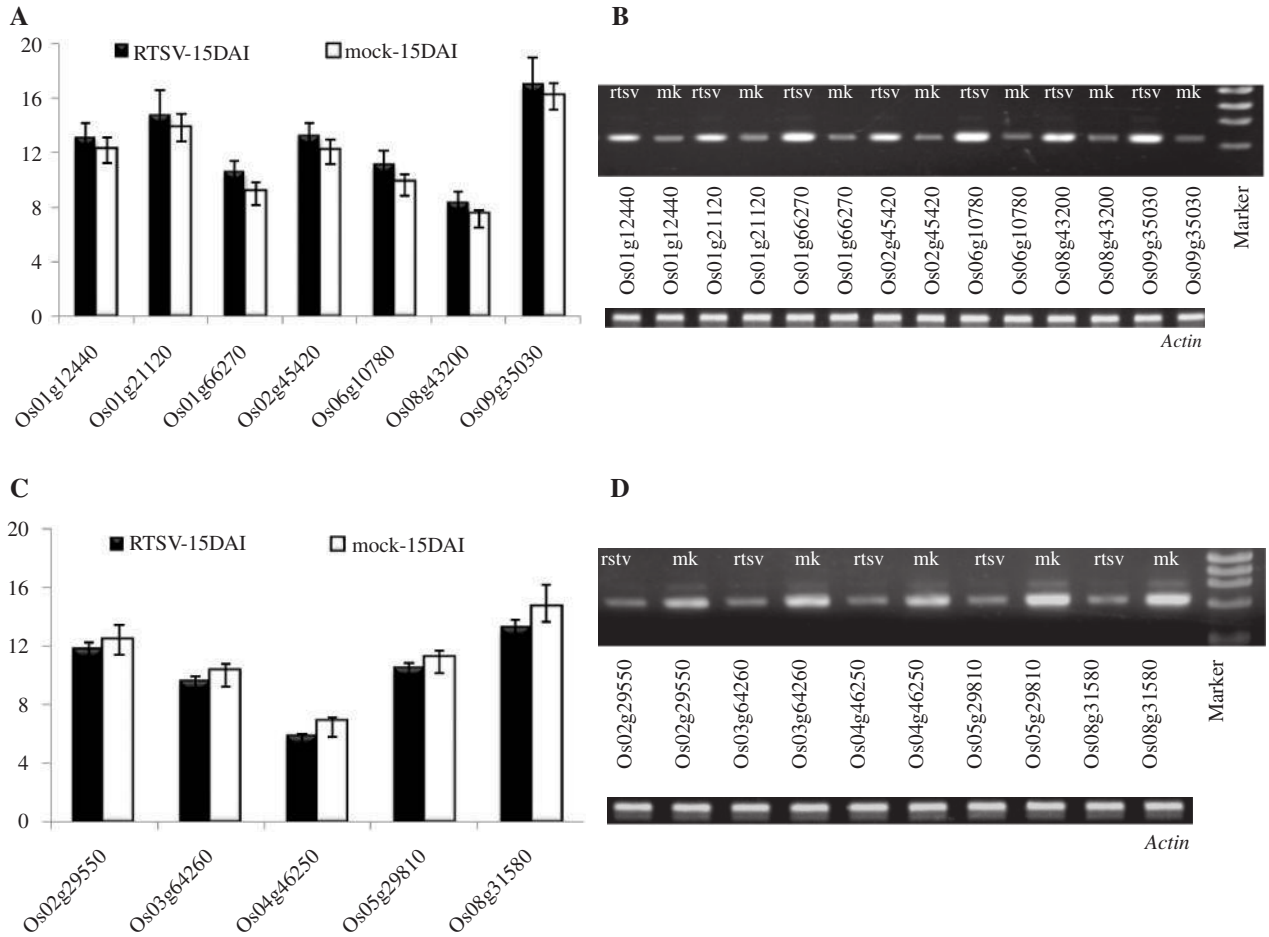

Fig. 7 (A) Histogram of microarray data for seven *OsAP2/EREBP* genes differentially expressed under RTSV infection. (B) RT–PCR analysis of seven genes that showed higher expression than the control under RSV infection at 15 DAI. (C) Histogram of microarray data for five genes down-regulated under RTSV infection. (D) RT–PCR analysis of eight genes that showed lower expression than the control under RSV infection at 15 DAI. *Actin* gene expression was used as the internal control; mk, mock (control); rtsv, RTSV-infected plants sampled 15 DAI. The differentially expressed genes are shown along the *x*-axis of the histograms, and the log2 ratios of expression values are shown on the *y*-axis.

Table 3 Stress response of transgenic plants overexpressing DREBs and ERFs

Gene	Transgenic plants	Performance of transgenic plants	Reference
AtDREB2A	Arabidopsis	Dehydration tolerance	Sakuma et al. (2006)
BNCBF5, BNCBF17	Canola	Freezing tolerance	Savitch et al. (2005)
AtDREB1A	Tobacco	Freezing and dehydration tolerance	Kasuga et al. (2004)
OsDREB1A	Arabidopsis	Freezing, dehydration and salt tolerance	Dubouzet et al. (2003)
AtCBF4	Arabidopsis	Freezing and dehydration tolerance	Haake et al. (2002)
AtCBF1	Potato	Freezing tolerance	Pino et al. (2008)
Tsil1	Tobacco	Biotic, osmotic stress tolerance	Park et al. (2001)
AtCBF3	Arabidopsis	Freezing tolerance	Gilmour et al. (2000)
AtCBF1	Arabidopsis	Freezing tolerance	Jaglo-Ottosen et al. (1998)
AtDREB1A	Arabidopsis	Freezing and dehydration tolerance	Liu et al. (1998), Oh et al. (2005)
TERF1	Rice	Drought and high-salinity tolerance	Gao et al. (2008)
TSRF1	Rice	Drought tolerance	Quan et al. (2010)
PgDREB2A	Tobacco	Salinity and drought tolerance	Agarwal et al. (2010)
RAP2.2	Arabidopsis	Low oxygen response	Hinz et al. (2010)
RAP2.6	Arabidopsis	ABA, salt and osmotic tolerant	Zhu et al. (2010)

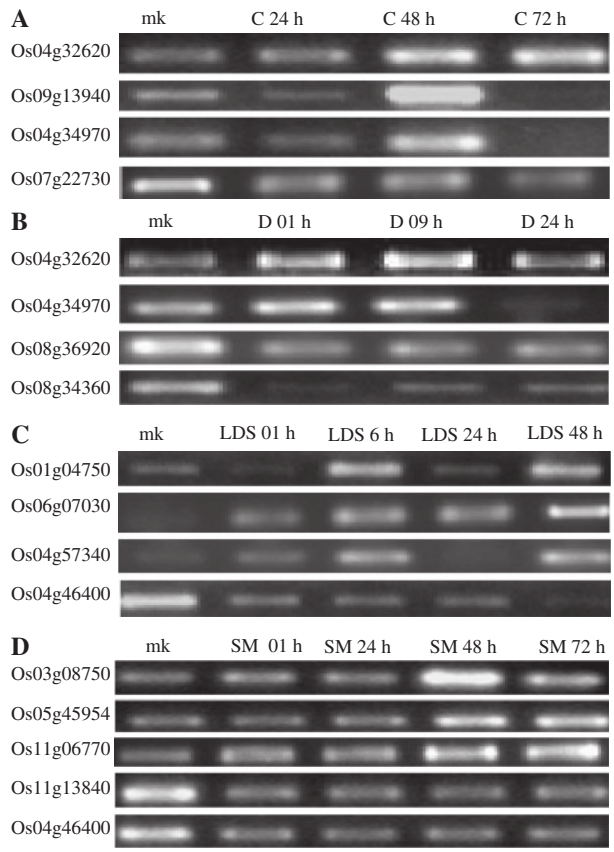

Fig. 8 RT-PCR analysis of differential gene expression under various abiotic treatments. (A) Three genes (*Os04g32620*, *Os09g13940* and *Os04g34970*) were up-regulated and one (*Os07g22730*) was down-regulated under cold stress (mk, mock/control; C 24 h, C 48 h and C 72 h, cold for 24, 48 and 72 h, respectively). (B) Of four genes, three (*Os04g32620*, *Os04g34970* and *Os08g36920*) were up-regulated to high levels and one was down-regulated (*Os08g34360*) to low levels under drought stress (mk, mock/control; D 01 h, D 09 h and D 24 h, drought for 1, 9 and 24 h, respectively). (C) Three genes showed higher expression (*Os01g04750*, *Os06g07030* and *Os04g57340*) and one gene showed lower expression (*Os04g46400*) than the control under laid-down submergence stress (mk, mock/control; LDS 01 h, LDS 06 h, LDS 24 h and LDS 48 h, laid-down submergence for 1, 6, 24 and 48 h, respectively). (D) Among five genes, three genes were up-regulated (*Os03g08470*, *Os05g45954* and *Os11g06770*) and two genes were down-regulated (*Os11g13840* and *Os04g46400*) under submergence stress (mk, mock/control; SM 24 h, SM 48 h and SM 72 h, submergence for 24, 48 and 72 h, respectively).

the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/) and the Knowledge-Based Oryza Molecular Biological Encyclopedia (http://cdna01.dna.affrc.go.jp/cDNA/). We then further conducted BLASTP and TBLASTN searches of the MSU and NCBI databases using two search parameters as follows: maximum number of target sequences, 300; and expected value, <10. SMART (http://smart.embl-heidelberg.de/) and Pfam (http://pfam.sanger.ac.uk/) database searches were used to confirm and classify each predicted *OsAP2/EREBP* gene. Exon and intron structures of this gene family were investigated using the NCPGR (http://gbrowse.ncpgr.cn/cgi-bin/gbrowse/japonica/) database.

Phylogenetic analysis and sequence alignment

A phylogenetic tree was constructed by using OsAP2/EREBP domain sequences, and an unrooted tree was generated by using ClustalX version 1.83 (Thompson et al. 1997) by the Neighbor–Joining method (Saitou and Nei 1987) and bootstrap analysis (1,000 replicates). The tree was displayed by using MEGA software version 4 (Tamura et al. 2007), and sequence alignments were performed with ClustalW (Thompson et al. 1994).

Gene locations on chromosomes and duplications

OsAP2/EREBP genes were located on rice chromosomes according to the positions specified in the MSU rice database. To find large segmental duplications, we identified genome duplications of rice in the MSU database with a maximum permitted distance between collinear gene pairs of 100 or 500 kb (http://rice.plantbiology.msu.edu/segmental_dup/index.shtml). We considered genes to be tandemly duplicated if two *OsAP2/EREBP* genes were separated by three or fewer gene loci according to the Rice Genome Annotation Release 6 of MSU.

Biotic and abiotic treatments

Seedlings (*japonica* cv. Nipponbare) were grown on absorbent tissue paper at 28°C under a 16 h light:8 h dark photoperiod for 10 d and were exposed to the various stress treatments: cold (24, 48 and 72 h, incubation at 10°C); drought (1, 9 and 24 h, with the addition of 25% polyethylene glycol 6000 to the planter box); submergence (24, 48 and 72 h); laid-down submergence (6, 24 and 48 h); osmotic (addition of 260 mM mannitol to the planter box for 24 h); salt (addition of 150 mM sodium chloride to the planter box for 24 h); plant hormones ABA and gibberellic acid (30-day-old calli were treated with 50 μM ABA or 50 μM gibberellic acid for 3 d). Control and treated seedlings were harvested and stored at −80°C until RNA extraction. Methods used to infect rice plants with biotic stresses were published by Satoh et al. (2010).

Data analysis (44K and 22K arrays)

For all biotic stress treatments (RSV, GSE12681; RTSV, GSE16141; RDV, GSE24937), we used 44K microarray data available at NCBI-GEO. Expression patterns of all samples (at least three biological repeats) were transformed into log 2-based numbers and normalized according to the quantile method for standardization of array data. Expression of a gene (up-or down-regulated) was defined as a gene with a log 2-based ratio (RSV, RTSV and RDV-inoculated sample/mock-inoculated sample) higher than 0.585 or lower than −0.585; and a significant difference in gene expression between the treated plants and the control indicated by $P \leq 0.05$ by paired *t*-test [permutations, all possible combinations; false discovery rate (FDR) correction, adjusted Bonferroni method]. Data processing was

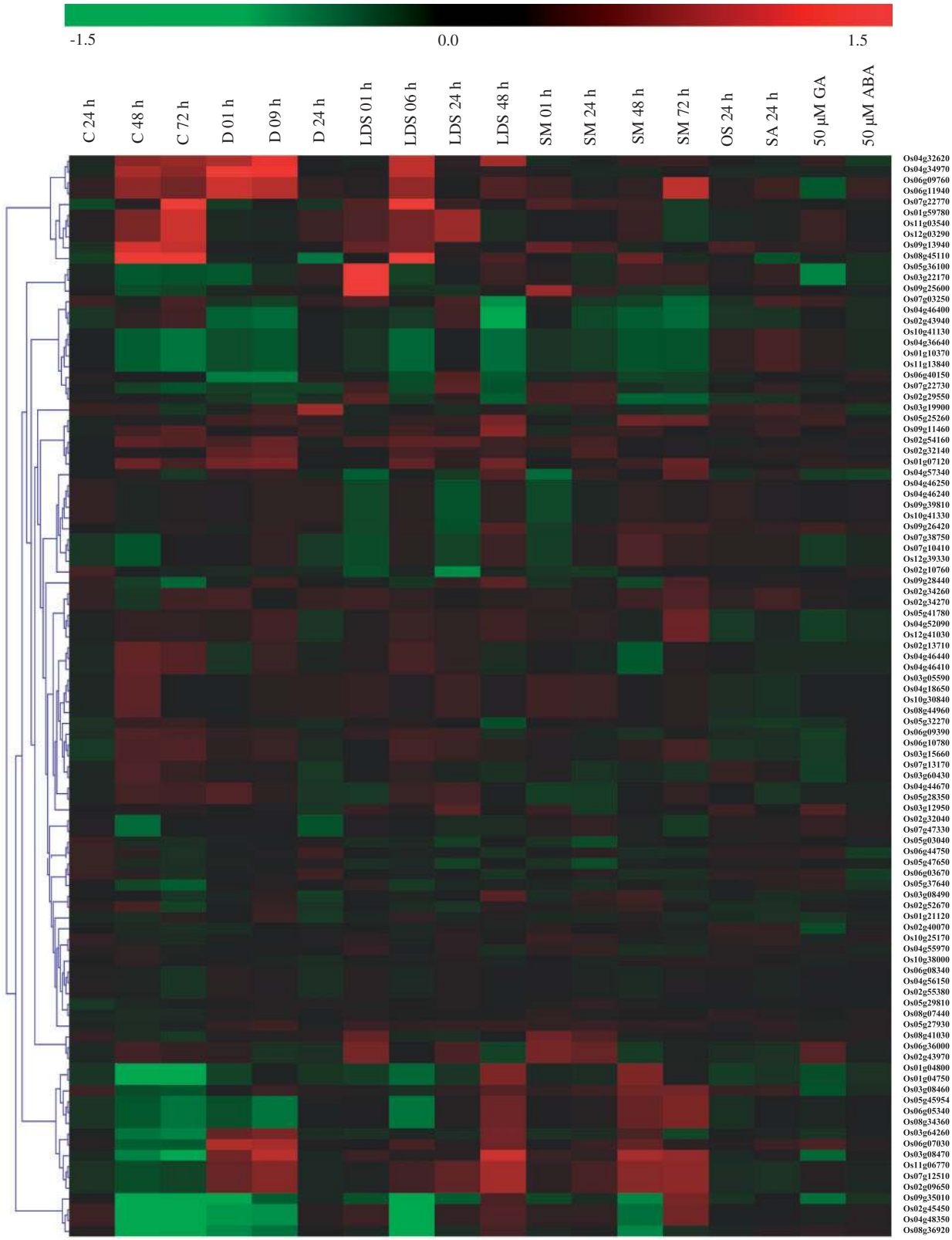

Fig. 9 Differential expression of *OsAP2/EREBP* genes under eight types of abiotic stress (log 10 ratio). The color bar at the top shows levels of expression: red indicates up-regulated genes and green indicates down-regulated genes. Stresses: C 24 h, C 48 h and C 72 h, cold for 24, 48 and 72 h, respectively; D 01 h, D 09 h and D 24 h, drought for 1, 9 and 24 h, respectively; LDS 01 h, LDS 06 h, LDS 24 h and LDS 48 h, laid-down submergence for 1, 6, 24 and 48 h, respectively; SM 24 h, SM 48 h and SM 72 h, submergence for 24, 48 and 72 h, respectively; OS, osmotic stress for 24 h; SA, salt for 24 h; 50 μM GA, 50 μM gibberellic acid.

done with MeV version 4.3. We identified 162 *OsAP2/EREBP* genes from 44K microarray data collected under all biotic stress conditions.

We downloaded 22K microarray data from NCBI-GEO (GSE7532, GSE2415 and GSE661) for the eight abiotic stresses above. In all, 22K microarray gene expression data corresponding to 100 *OsAP2/EREBP* genes were retrieved. Expression intensities in response to all abiotic treatments are log 10 ratio values, and those genes with a threshold value higher than 0.176 or lower than −0.176 in both replications were chosen. The DEGs were either up- or down-regulated in the two repetitions.

Transcript-level analysis

Total RNA samples were extracted from plant materials with an RNA extraction kit (RNeasy Maxi Kit, Qiagen) in accordance with the manufacturer's instructions. Total RNA samples were extensively pre-treated with RNase-free DNase I to eliminate any contaminating genomic DNA. The first-strand cDNA was synthesized from 1 μg of total RNA in a 20 μl reaction volume using Superscript II reverse transcriptase (Invitrogen), and 2 μl of the reaction mixture was subsequently used for RT–PCR runs in a 50 μl reaction volume. RT–PCR was performed using SuperScriptII reverse transcriptase (Invitrogen) in accordance with the manufacturer's instructions to synthesize first-strand cDNA from the DNase I-treated total RNA. The RNA samples for hybridization and RT–PCR were the same. About 1/20 of the first-strand cDNA generated from 1 μg of total RNA was used as a template for PCR in a reaction volume of 50 μl with rTaq DNA polymerase (TAKARA). The RT–PCR runs consisted of 25–38 cycles, depending on the linear range of PCR amplification for expression of each gene. Each PCR was performed (repeated three times) in an ABI 9700 Thermocycler (Applied Biosystems) consisting of incubation at 94°C for 1 min, at 55°C for 50 s and at 72°C for 1 min. The rice *actin* gene was used for RT–PCR as an internal control and the primers are listed in **Supplementary Table S8**.

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) [grants to A.M.S., M.N., K.S., H.K., T.S. and A.H.].

Acknowledgments

We thank Dr. Takayuki Asano for his intensive reading and fruitful comments and Ms. Setsuko Kimura for her technical support in the microarray experiment.

References

Agarwal, P., Agarwal, P.K., Joshi, A.J., Sopory, S.K. and Reddy, M.K. (2010) Overexpression of PgDREB2A transcription factor enhances abiotic stress tolerance and activates downstream stress-responsive genes. *Mol. Biol. Rep.* 37: 1125–1135.

Agarwal, P.K., Agarwal, P., Reddy, M.K. and Sopory, S.K. (2006) Roles of DREB transcription factors in biotic and abiotic stress tolerance in plants. *Plant Cell Rep.* 25: 1263–1274.

Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M. (1998) A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J.* 17: 5484–5496.

Alonso, J.M., Stepanova, A.N. and Leisse, T.J. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.

Atchley, W., Fitch, W. and Bronner-Fraser, M. (1994) Molecular evolution of the MyoD family of transcription factors. *Proc. Natl Acad. Sci. USA* 91: 11522–11526.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D. and May, G. (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4: 10.

Cao, Y., Song, F., Goodman, R.M. and Zheng, Z. (2006) Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. *J. Plant Physiol.* 163: 1167–1178.

Dias, A.P., Braun, E.L., McMullen, M.D. and Grotewold, E. (2003) Recently duplicated maize R2R3 Myb genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. *Plant Physiol.* 131: 610–620.

Dong, J., Chen, C. and Chen, Z. (2003) Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Mol. Biol.* 51: 21–37.

Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, S. et al. (2003) *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought, high-salt- and cold-responsive gene expression. *Plant J.* 33: 751–763.

Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D. et al. (1996) AINTEGUMENTA, an APETALA2-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *Plant Cell* 8: 155–168.

Fowler, S. and Thomashow, F. (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* 14: 1675–1690.

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H. and Ohme-Takagi, M. (2000) *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators and repressors of GCC box-mediated gene expression. *Plant Cell* 12: 393–404.

Gao, G., Zhong, Y., Guo, A., Zhu, Q., Tang, W., Zheng, W. et al. (2006) DRTF: a database of rice transcription factors. *Bioinformatics* 22: 1286–1287.

Gao, S., Zhang, H., Tian, Y., Li, F., Zhang, Z., Lu, X. et al. (2008) Expression of TERF1 in rice regulates expression of stress-responsive genes and enhances tolerance to drought and high-salinity. *Plant Cell Rep.* 27: 1787–1795.

Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D. and Thomashow, M.F. (2000) Overexpression of *Arabidopsis* CBF3

transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol.* 124: 1854–1865.

Gu, Y.Q., Yang, C., Thara, V.K., Zhou, J. and Martin, G.B. (2000) Pti4 is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. *Plant Cell* 12: 771–786.

Gutterson, N. and Reuber, T.L. (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Biol.* 7: 465–471.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F. and Zhang, J.Z. (2002) Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. *Plant Physiol.* 130: 639–648.

Hao, D.Y., Ohme-Takagi, M. and Sarai, A. (1998) Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *J. Biol. Chem.* 273: 26857–26861.

Hinz, M., Wilson, I.W., Yang, J., Buerstenbinder, K., Llewellyn, D., Dennis, E.S. et al. (2010) *Arabidopsis* RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. *Plant Physiol.* 153: 757–772.

Hu, Y.X., Wang, Y.X., Liu, X.F. and Li, J.Y. (2004) *Arabidopsis* RAVI is down-regulated by brassinosteroid and may act as a negative regulator during plant development. *Cell Res.* 14: 8–15.

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

Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O. and Thomashow, M.F. (1998) *Arabidopsis CBF1* overexpression induces *COR* genes and enhances freezing tolerance. *Science* 280: 104–106.

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A. et al. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449: 463–467.

Jain, M., Nijhawan, A., Arora, R., Agarwal, P., Ray, S., Sharma, P. et al. (2007) F-box proteins in rice. Genome wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 143: 1467–1483.

Jeong, J.S., Kim, Y.S., Baek, K.H., Jung, H., Ha, S.H., Do Choi, Y. et al. (2010) Root-specific expression of *OsNAC10* improves drought tolerance and grain yield in rice under field drought conditions. *Plant Physiol.* 153: 185–197.

Jin, L.G. and Liu, J.Y. (2008) Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton. *Plant Physiol. Biochem.* 46: 46–53.

Jofuku, K.D., den Boer, B.G.W., Van Montagu, M. and Okamuro, J.K. (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* 6: 1211–1225.

Kasuga, M., Miura, S., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2004) A combination of the *Arabidopsis* DREB1A gene and stress-inducible *rd29A* promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol.* 45: 346–350.

Kitanaga, Y., Jian, C., Hasegawa, M., Yazaki, J., Kishimoto, N., Kikuchi, S. et al. (2006) Sequential regulation of gibberellin, brassinosteroid, and jasmonic acid biosynthesis occurs in rice coleoptiles to control the transcript levels of anti-microbial thionin genes. *Biosci. Biotechnol. Biochem.* 70: 2410–2419.

Kizis, D. and Pages, M. (2002) Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway. *Plant J.* 30: 679–689.

Levy, A., Dafny-Yelin, M. and Tzfira, T. (2008) Attacking the defenders: plant viruses fight back. *Trends Microbiol.* 16: 194–197.

Li, H., Zhou, S.Y., Zhao, W.S., Su, S.C. and Peng, Y.L. (2009) A novel wall-associated receptor-like protein kinase gene, *OsWAK1*, plays important roles in rice BLAST disease resistance. *Plant Mol. Biol.* 69: 337–346.

Lin, H., Zhu, W., Silva, C.J., Gu, X. and Buell, C.R. (2006) Intron gain and loss in segmentally duplicated genes in rice. *Genome Biol.* 7: 41.

Lin, R., Zhaom, W., Mengm, X., Wang, M. and Peng, Y. (2007) Rice gene *OsNAC19* encodes a novel NAC-domain transcription factor and responds to infection by *Magnaporthe grisea*. *Plant Sci.* 172: 120–130.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. et al. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10: 1391–1406.

Liu, Y., Zhao, T.J., Liu, J.M., Liu, W.Q., Yan, Y.B. et al. (2006) The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. *FEBS Lett.* 580: 1303–1308.

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

Moore, R.C. and Purugganan, M.D. (2003) The early stages of duplicate gene evolution. *Proc. Natl Acad. Sci. USA* 100: 15682–15687.

Mukhopadhyay, A., Vij, S. and Tyagi, A.K. (2004) Overexpression of a zinc finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. *Proc. Natl Acad. Sci. USA* 101: 6309–6314.

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

Nuruzzaman, M., Gupta, M., Zhang, C., Wang, L., Xie, W., Xiong, L. et al. (2008) Sequence and expression analysis of the thioredoxin protein gene family in rice. *Mol. Genet. Genomics* 280: 139–151.

Oh, S.J., Kim, S.Y., Kwon, C.W., Park, H.K., Jeong, J.S. and Kim, J.K. (2009) Overexpression of the transcription factor AP37 in rice improves grain yield under drought conditions. *Plant Physiol.* 150: 1368–1379.

Oh, S.J., Song, S.I., Kim, Y.S., Jang, H.J., Kim, S.Y., Kim, M. et al. (2005) *Arabidopsis* CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. *Plant Physiol.* 138: 341–351.

Ohme-Takagi, M. and Shinshi, H. (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* 7: 173–182.

Park, J.M., Park, C.J., Lee, S.B., Ham, B.K., Shin, R. and Paek, K.H. (2001) Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell* 13: 1035–1046.

Pino, M.T., Skinner, J.S., Jeknic, Z., Hayes, P.M., Soeldner, A.H., Thomashow, M.F. et al. (2008) Ectopic AtCBF1 over-expression enhances freezing tolerance and induces cold acclimation-associated physiological modifications in potato. *Plant Cell Environment*, 31: 393–406.

Qin, Q.L., Liu, J.G., Zhang, Z., Peng, R.H., Xiong, A.S., Yao, Q.H. et al. (2007) Isolation, optimization, and functional analysis of the cDNA encoding transcription factor RdreB1 in *Oryza sativa* L. *Mol. Breed.* 19: 329–340.

Quan, R., Hu, S., Zhang, Z., Zhang, H., Zhang, Z. and Huang, R. (2010) Overexpression of an ERF transcription factor TSRF1 improves rice drought tolerance. *Plant Biotechnol. J.* 8: 476–488.

Riano-Pachon, D.M., Ruzicic, S., Dreyer, I. and Mueller-Roeber, B. (2007) PlnTFDB: an integrative plant transcription factor database. *BMC Bioinform.* 8: 42.

Riechmann, J.L. and Meyerowitz, E.M. (1998) The AP2/EREBP family of plant transcription factors. *Biol. Chem.* 379: 633–646.

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

Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2002) DNA-binding specificity of the AP2/ERF domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Commun.* 290: 998–1009.

Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K. et al. (2006) Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell* 18: 1292–1309.

Satoh, K., Kondoh, H., Sasaya, T., Shimizu, T., Choi, I.R., Omura, T. et al. (2010) Selective modification of rice (*Oryza sativa*) gene expression by rice stripe virus infection. *J. Gen. Virol.* 91: 294–305.

Savitch, L.V., Allard, G., Seki, M., Robert, L.S., Tinker, N.A., Huner, N.P.A. et al. (2005) The effect of overexpression of two *Brassica CBF/DREB1*-like transcription factors on photosynthetic capacity and freezing tolerance in *Brassica napus*. *Plant Cell Physiol.* 46: 1525–1539.

Shimono, M., Sugano, S., Nakayama, A., Jiang, C.J., Ono, K., Toki, S. et al. (2007) Rice WRKY45 plays a crucial role in benzothiadiazole-inducible BLAST resistance. *Plant Cell* 19: 2064–2076.

Shin, R., Park, J., An, J. and Paek, K. (2002) Ectopic expression of Tsi1 in transgenic hot pepper plants enhances host resistance to viral, bacterial, and oomycete pathogens. *Mol. Plant-Microbe Interact.* 15: 983–989.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr. Opin. Plant Biol.* 3: 217–223.

Sohn, K.H., Lee, S.C., Jung, H.W., Hong, J.K. and Hwang, B.K. (2006) Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. *Plant Mol. Biol.* 61: 897–915.

Tameling, W.I. and Baulcombe, D.C. (2007) Physical association of the NB-LRR resistance protein Rx with a Ran GTPase-activating protein is required for extreme resistance to potato virus X. *Plant Cell* 19: 1682–1694.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596–1599.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and 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.

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000) *Arabidopsis* basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc. Natl Acad. Sci. USA* 97: 11632–11637.

Xiong, R., Wu, J., Zhou, Y. and Zhou, X. (2008) Identification of a movement protein of the tenuivirus rice stripe virus. *J. Virol.* 82: 12304–12311.

Xiong, R., Wu, J., Zhou, Y. and Zhou, X. (2009) Characterization and subcellular localization of an RNA silencing suppressor encoded by rice stripe tenuivirus. *Virology* 387: 29–40.

Zhou, J., Tang, X. and Martin, G.B. (1997) The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a *cis*-element of pathogenesis-related genes. *EMBO J.* 16: 3207–3218.

Zhu, Q., Zhang, J., Gao, X., Tong, J., Xiao, L., Li, W. et al. (2010) The *Arabidopsis* AP2/ERF transcription factor RAP2.6 participates in ABA, salt and osmotic stress responses. *Gene* 457: 1–12.

Zhuang, J., Cai, B., Peng, R.H., Zhu, B., Jin, X.F., Xue, Y. et al. (2008) Genome-wide analysis of the *AP2/ERF* gene family in *Populus trichocarpa*. *Biochem. Biophys. Res. Commun.* 371: 468–474.