Gene 465 (2010) 30–44

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Methods paper

Genome-wide analysis of NAC transcription factor family in rice

Mohammed Nuruzzaman [a], Ramaswamy Manimekalai [a], Akhter Most Sharoni [a], Kouji Satoh [a], Hiroaki Kondoh [a], Hisako Ooka [b], Shoshi Kikuchi [a,*]

[a] Plant Genome Research Unit, Division of Genome and Biodiversity Research, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki 305-8602, Japan
[b] Kurume National College of Technology, Komorino 1-1-1, Kurume, Fukuoka 830-8555, Japan

ARTICLE INFO

Article history:
Accepted 16 June 2010
Available online 26 June 2010

Received by C. Feschotte

Keywords:
Rice
Phylogenetic analysis
Microarray
Stress

ABSTRACT

We investigated 151 non-redundant *NAC* genes in rice and 117 in Arabidopsis. A complete overview of this gene family in rice is presented, including gene structures, phylogenies, genome localizations, and expression profiles. We also performed a comparative analysis of these genes in rice and Arabidopsis. Conserved amino acid residues and phylogeny construction using the NAC conserved domain sequence suggest that *OsNAC* gene family was classified broadly into two major groups (A and B) and sixteen subgroups in rice. We presented more specific phylogenetic analysis of OsNAC proteins based on the DNA-binding domain and known gene function, respectively. Loss of introns was observed in the segmental duplication. Homologous, paralogous, and orthologous searches of rice and Arabidopsis revealed that the major functional diversification within the *NAC* gene family predated the divergence of monocots and dicots. The chromosomal localizations of *OsNAC* genes indicated nine segmental duplication events involving 18 genes; 32 non-redundant *OsNAC* genes were involved in tandem duplications. Expression levels of this gene family were checked under various abiotic stresses (cold, drought, submergence, laid-down submergence, osmotic, salinity and hormone) and biotic stresses [infection with rice viruses such as RSV (rice stripe virus) and RTSV (rice tungro spherical virus)]. Biotic stresses are novel work and increase the possibilities for finding the best candidate genes. A preliminary search based on our microarray (22K and 44K) data suggested that more than 45 and 26 non-redundant genes in this family were upregulated in response to abiotic and biotic stresses, respectively. All of the genes were further investigated for their stress responsiveness by RT-PCR analysis. Six genes showed preferential expression under both biotic RSV and RTSV stress. Eleven genes were upregulated by at least three abiotic treatments. Our study provides a very useful reference for cloning and functional analysis of members of this gene family in rice.

1. Introduction

A number of transcription factors are involved to regulate the gene expression in living organism. As part of their regulatory function, transcription factors interact with plant-specific DNA sequences. Knowledge of the structure of the DNA-binding domains of transcription factors is essential for understanding their function and mechanism. The *NAC* gene family name was derived from the names of three transcription factors: (i) NAM (no apical meristem, Petunia), (ii) ATAF1–2, and (iii) CUC2 (cup-shaped cotyledon, Arabidopsis), all of which have the same DNA-binding domain (Souer et al., 1996; Aida et al., 1997). *NAC* genes encode plant-specific transcriptional regulators that constitute a large transcription factor family in plants (Olsen et al., 2005); this family was first identified by mutations (Souer et al., 1996; Aida et al., 1997). Transcription factor families are expressed at much higher levels in plants than in animals. In the plant kingdom, more than 50 families of different transcription factors have been identified by sequence analyses of model species such as rice (Xiong et al., 2005), and numerous reports suggest that transcription factors and *cis*-acting elements are involved in almost all aspects of cellular activity as part of their related roles in promoting gene expression (Xiong et al., 2005).

NAC protein family members are highly conserved at the N-terminal NAC binding domain and have a highly variable C-terminal domain that plays a major role in the regulation of transcription (Olsen et al., 2005). This unstable C-terminal domain of NAC proteins generally operates as a functional domain and acts as a transcriptional activator or repressor (Tran et al., 2004; Hu et al., 2006; Kim et al., 2007a). This variable C-terminal domain is very large and also has protein binding activity. Kim et al. (2007a) reported that the C-terminal domain of Arabidopsis calmodulin-binding (CB) NAC can bind with calmodulin proteins.

There are 105 redundant putative *NAC* genes in Arabidopsis (Ooka et al., 2003), and 140 putative *NAC* or *NAC*-like genes in rice (Fang

---

Abbreviations: NAC, NAMATAF1/2 and CUC transcription factor; RSV, Rice stripe virus; RTSV, Rice tungro spherical virus; RT-PCR, Reverse transcription-polymerase chain reaction; BAC, Bacterial artificial chromosome; Os, *Oryza sativa*; At, *Arabidopsis thaliana*; GA, Gibberellic acid; ABA, Abscisic acid.

* Corresponding author. Tel./fax: +81 29 838 7007.
E-mail address: skikuchi@nias.affrc.go.jp (S. Kikuchi).

0378-1119/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2010.06.008

et al., 2008). The NAC transcription factors appeared to control of biochemical and molecular pathways that can save plants under different stress conditions. The NAC transcription factors are multifunctional proteins with various roles in the plant life cycle, such as maintenance of the shoot apical meristem (Souer et al., 1996; Kim et al., 2007b), cotyledon development (Aida et al., 1997), lateral root development (He et al., 2005), flower formation (Sablowski and Meyerowitz, 1998), hormone signaling (Greve et al., 2003), response to pathogen infection (Xie et al., 1999; Ren et al., 2000; Olsen et al., 2005; Nakashima et al., 2007), plant organ senescence (Liu et al., 2009), embryo development (Duval et al., 2002), response to different abiotic stresses (Tran et al., 2004; He et al., 2005; Hu et al., 2006; Bhatnagar-Mathur et al., 2007; Nakashima et al., 2007; Yoo et al., 2007), formation of secondary walls (Zhong et al., 2007), cell division (Kim et al., 2006), fiber development (Ko et al., 2007), seed development (Sperotto et al., 2009), and senescence (Uauy et al., 2006). In addition, expression of a large array of genes is induced by a number of stress conditions, meaning that numerous proteins are produced to join the pathways that lead to synergistic improvement of stress tolerance (Seki et al., 2003).

NAC transcription factor production is induced by drought, as was first shown in Arabidopsis. Overexpression of three *Arabidopsis thaliana* (*At*) *NAC* genes (*ANAC019*, *ANAC055*, and *ANAC072*) in transgenic plants increased the stress tolerance of the plants and altered the expression of drought, salinity, and low-temperature-stress-inducible genes (Tran et al., 2004). Furthermore, some researchers have reported important functions of these genes in abiotic and biotic stress resistance during the plant life cycle. One field study showed that introduction of *SNAC1* into rice increased drought tolerance in the transgenic plants and gave 22% to 34% greater seed set than in the negative control population upon exposure to severe drought during flowering (Hu et al., 2006). *SNAC1* is also involved in cold and salt tolerance (Hu et al., 2006). *Oryza sativa* (*Os*) *NAC6*/*SNAC2* and *OsNAC10* are drought-tolerance gene that improves the expression of several protein-encoding genes under stress conditions (Nakashima et al., 2007; Jeong et al., 2010) and belongs to the ATAF subfamily (Kikuchi et al., 2000; Ooka et al., 2003). In Arabidopsis, *AtNAC2* expression is induced by salt and abscisic acid and amends salt tolerance and root development (He et al., 2005). Hegedus et al. (2003) have identified nine *NAC* genes that are upregulated in *Brassica napus* (rapeseed) by cold temperature. In wheat (*Triticum aestivum*) *NAC* gene, (*NAM-B1*) played an important role to move nutrient remobilization from leaves to developing grains (Uauy et al., 2006). *TaNAC4* gene functions as a transcriptional activator involved in wheat response to abiotic and biotic stresses (Xia et al., 2010). Lin et al. (2007) found that *OsNAC19* transcript was elevated by the infection of *Magnaporthe grisea*, suggesting that *OsNAC19* involved in rice defense response to *M. grisea* infection. NAC proteins are also involved in response to viral infection during vegetative development of plants (Xie et al., 1999; Ren et al., 2000).

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 all the genes belonging to specific gene families. In rice, only a few *NAC* genes have been characterized, and the functions of most of them remain to be determined. Here, we identified 151 *OsNAC* genes in rice by database searches and classified these genes according to reported genes. The subgroups are more specific than previous work and biotic stresses are new work in this present study. *OsNAC* genes play an important role in the crosstalk of different kinds of stresses signaling. We analyzed the phylogenetic relationships of the *NAC* genes in rice and Arabidopsis, as well as the segmental and tandem duplications and exon and intron structures of *OsNAC* genes. We also studied the expression intensities of *OsNAC* genes under different abiotic and biotic stresses from our 22K and 44K microarray data. The data generated should be very helpful in studies of the biological functions of each *OsNAC* gene.

2. Materials and methods

2.1. Collection and classification of OsNAC gene family members

To identify members of the rice (*Oryza sativa* L. subsp. *japonica* cv.) NAC transcription factor gene family, multiple database searches were performed. The Database of Rice Transcription Factors (DRTF, http://drtf.cbi.pku.edu.cn; Gao et al., 2006), rice genome annotation (Machigan State University, MSU; http://rice.plantbiology.msu.edu/), Rice Transcription Factor Database (Rice-TFDB, http://ricetfdb.bio.uni-potsdam.de/v2.1/; Riano-Pachon et al., 2007), National Centre for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/), and the Knowledge-Based Oryza Molecular Biological Encyclopedia (KOME; http://www.cdna01.dna.affrc.go.jp/cDNA) were used to search for members of the *NAC* gene family. We then further explored the MSU database to gather a more complete collection of putative *NAC* genes in rice by using keywords (NAC, NAM) and a domain (PF02365) search. The BLASTP and TBLASTN search parameters of the three databases (MSU, NCBI, and KOME) were set as follows: maximum target sequences 350, and expected value less than 10. SMART (http://smart.embl-heidelberg.de/) and Pfam database (http://pfam.sanger.ac.uk/) searches were used to confirm and classify each predicted *OsNAC* gene. Information about gene structure, transcripts, length, chromosomal localization, full-length cDNA, BAC accessions for each gene, and the characteristics of the corresponding proteins were procured from MSU, KOME, and GRAMENE (Liang et al., 2008). Exon and intron structures were investigated by using the National Center of Plant Gene Research (NCPGR, http://gbrowse.ncpgr.cn/cgi-bin/gbrowse/japonica/) database.

2.2. NAC genes from Arabidopsis and other species

Various database searches were performed to collect data on all members of the Arabidopsis *NAC* gene family. We used the BLAST programs (TBLASTN and BLASTP) available in The Arabidopsis Information Resource (TAIR), MSU Arabidopsis databases and the NCBI Arabidopsis genome database. As a query sequence, we first used the amino acid sequence of the NAM domain from the rice NAC. To increase the extent of the database search results, we also performed a position-specific iterated BLAST search (Altschul et al., 1997) against the Arabidopsis database on the NCBI website. To confirm completion of the collection we also performed database searches using the amino acid sequences of the NAM domain of some members of the Arabidopsis NAC family as query sequences. We indentified non-redundant 117 *AtNAC* genes from database searches. We used the Plant Transcription Factor Data Base (PTFDB, http://plntfdb.bio.uni-potsdam.de/v3.0/fam_mem.php?family_id=NAC) for other species (sorghum, maize, and poplar).

2.3. Phylogenetic analysis and sequence alignment

Phylogenetic trees were constructed by using OsNAC domain sequences, and an unrooted tree was generated by ClustalX version 1.83 (Thompson et al., 1997) by the neighbor-joining method (Saitou and Nei, 1987) and bootstrap analysis (1000 replicates). The tree was analyzed and displayed by using MEGA software version 4 (Tamura et al., 2007). Another two unrooted trees were constructed by the same method to align OsNAC and AtNAC protein sequences and to align the NAC protein sequences of five monocot and dicot species (rice, sorghum, maize, poplar, and Arabidopsis). We defined two proteins with 100% support in the same species as homologous proteins and two with 100% support in different species as orthologous proteins. Multiple sequence alignments were performed with ClustalX version 1.83 (Thompson et al., 1997).

2.4. Gene positions on chromosomes, and duplications

OsNAC genes were located on rice chromosomes according to their positions given in the MSU rice database. The positions of *OsNAC* genes on the rice chromosome maps were drawn and modified manually, with annotation. 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 kb or 500 kb (http://rice.plantbiology.msu.edu/segmental_dup/index.shtml). We selected tandemly duplicated genes if two *OsNAC* genes were separated by four or fewer gene loci according to the Rice Genome Annotation Release 6 of MSU.

2.5. Plant growth and abiotic and biotic treatments

Rice seeds (*Oryza sativa* L. ssp. *japonica* cv. 'Nipponbare') were supplied by Drs. M. Yano and T. Matsumoto of the National Institute of Agrobiological Sciences (NIAS), Japan. Seedlings were grown on absorbent tissue paper at 28 °C under a 16:8 light:dark cycle for 10 days. Seedlings were then exposed to the following conditions: cold stress (24 h, 48 h, or 72 h, incubation at 10 °C); mild drought stress (01 h, 09 h, or 24 h, addition of 25% polyethylene glycol 6000 to the planter box for mild drought stress); submergence (24 h, 48 h, or 72 h); laid-down submergence (06 h, 24 h, and 48 h); severe osmotic stress (addition of 260 mM mannitol to the planter box for 24 h); salt stress (addition of 150 mM sodium chloride to the planter box for 24 h); and plant hormone, abscisic acid (ABA) and gibberellic acid (GA) (30-day-old calli were treated by 50 μM ABA and 50 μM GA for three days). Controls for all microarray experiments were 10-day-old seedlings before each stress treatment. Sample seedlings were harvested after the stress treatment and stored at −80 °C. After 10 days of growth, control seedlings were harvested and stored at −80 °C.

Taichung Native 1 (TN1) is an *indica* rice cultivar susceptible to rice tungro spherical virus (RTSV) (Encabo et al., 2009). RTSV strain A (Cabauatan et al., 1995) maintained in TN1 was used as the source of inoculum. Insect inoculation of plants with viruses was done by the tube method as described by Cabauatan et al. (1995). Green leafhoppers (*Nephotettix virescens* Distant) were given acquisition access for 3 days to RTSV-infected TN1 and were allowed a 24-h inoculation access period to 9-day-old plants at three insects per plant. Inoculated plants were maintained in a temperature-controlled greenhouse (28 ± 3 °C, natural sunlight). At 3, 6, 9, 12, 15, 18, and 21 days after inoculation (DAI), the shoots of the inoculated plants (excluding the meristem) were cut 3 to 5 cm above the soil surface. After measurement of fresh weight, the plant samples were frozen in liquid nitrogen and stored at −80 °C until extraction of RNA. Rice plants infected with virus (RSV) methods were published by Satoh et al. (2010).

2.6. Estimation of virus accumulation

Relative amounts of RTSV accumulated in TN1 were estimated as described by Shibata et al. (2007) using the absorbance values obtained by enzyme-linked immunosorbent assay (ELISA) for RTSV coat protein (CP) in plants. Nine-day-old plants were inoculated with RTSV as described above. Shoots of 10 individual plants were collected at 0, 3, 6, 9, 12, 15, 18, and 21 DAI. Shoot extracts from the individual plants were diluted 1 in 10 (w/v) in phosphate-buffered saline and subjected to ELISA. The average absorbance value of extracts from plants of the same treatment was taken as the relative level of RTSV accumulation in the plant at each time point. The experiments were repeated three times.

2.7. Rice oligo microarray analysis (22K and 44K)

Target complementary RNA (cRNA) was amplified and labeled with cyanine-3 (Cy3) and cyanine-5 (Cy5) from 200 to 500 ng mRNA by using a Fluorescent Linear Amplification Kit (Agilent Technologies). The quality of the targets was determined with a model 2100 Bioanalyzer (Agilent Technologies) and the concentration with a spectrophotometer (Nanodrop ND-1000; Nanodrop Technologies). We hybridized 1 μg fluorescent linear-amplified Cy3- and Cy5-labeled cRNA (500 ng of each cRNA) to a customized rice oligo microarray. Labeled cRNAs from each of the six treatment groups and the control were mixed and fragmented to an average size of 100 to 200 bases with an in situ Hybridization Kit Plus (Agilent Technologies) by incubation at 60 °C for 30 min. Fragmented cRNAs were added to hybridization buffer, applied to the microarray, and hybridized for 17 h at 60 °C. The slides (except in the case of cold stress) were washed with SSC (10 min in 6× SSC and 0.005% Triton X-102 at room temperature, then 5 min in 0.1× SSC and 0.005% Triton X-102 at 4 °C). Cold-stress slides were washed by the saline–sodium phosphate–EDTA (SSPE) method. Slides were dried and scanned on an Agilent Technologies G2565BA Microarray Scanner System. Scanned microarray images were analyzed by using Feature Extraction 6.1.1 software (Agilent Technologies), and dye-normalized, background-subtracted intensity and ratio data were exported to a text file. This software flags corrupted spots and detects lack of differences between sample spots and the background. The availability of spots on the microarray was greater than 85%.

Individual total RNA samples were extracted from shoot tissues pooled from five plants of the same treatment by using an RNeasy Maxi kit (Qiagen) in accordance with the manufacturer's instructions. Cy3- or Cy5-labeled cRNA samples were synthesized from 850 ng of the total RNA by using a low-input RNA Labeling Kit (Agilent Technologies). Hybridization solution was prepared with 825 ng each of Cy3- and Cy5-labeled cRNA preparations using an in situ Hybridization Kit Plus (Agilent Technologies). A custom 60-mer oligoarray printed with probes for about 43,500 rice genes was used for comparison of gene expression levels between TN1 inoculated with RTSV (Cy3-labeled samples) and those inoculated with virus-free green leafhoppers (Cy5-labeled samples) harvested on the same DAI (Satoh et al., unpublished). Hybridization and washing of microarray slides were performed in accordance with the manufacturer's protocols. The slide images were scanned by a DNA microarray scanner (G2505B; Agilent Technologies). Information on probe arrangement on the array (GPL7252) is available at NCBI-Gene Expression Omnibus GEO.

2.8. Data analysis (22K and 44K arrays)

Abiotic gene expression data corresponding to the 84 *OsNAC* genes were retrieved from the 22K microarray data. The microarray data download from NCBI-GEO (GSE7532, GSE2415, and GSE661) for eight abiotic stresses (cold, drought, submergence, laid-down submergence, osmotic, salt, 50 μM ABA, and 50 μM GA). Expression intensities of all treatments are log10 ratio value and each of treatments is two biological repeats. We selected the common genes in both replications with a threshold higher than 0.176 or lower than −0.176. The log10 ratio values of the common genes are averaged and selected the up- or down regulated genes in the different abiotic treatments (Supplementary Table S1).

Image files of the slides were processed by Feature Extraction version 9.5 (Agilent Technologies). Cy3 and Cy5 signal intensities were normalized by rank-consistency filtering and the LOWESS method. The microarray data used (RSV; GSE12681 and RTSV; GSE16141) are available at NCBI-GEO. Signal intensities of all samples were transformed into log2-based numbers and normalized according to the quantile method for standardization among the array data. A gene was declared “expressed” if its average signal intensity was higher than six under at least one condition; otherwise, the gene was considered “not expressed”. A significantly differentially expressed gene was defined as an expressed gene with (i) a log2-based ratio (RSV and RTSV-inoculated sample/mock-inoculated sample) higher than 0.585 or lower than −0.585; and (ii) significance in changes in gene expression between the two plants of $P \leq 0.05$ by paired *t*-test (permutations: all possible combinations; FDR correction: adjusted Bonferroni method). Data processing was done using MeV version 4.3. We also found 130 *OsNAC* genes from our 44K microarray data collected under RSV or RTSV stress conditions.

2.9. RT-PCR analysis to confirm microarray data

RT-PCR was performed to confirm the differential expression of representative *OsNAC* genes identified by microarray data analysis using gene-specific primers. Total RNA was extracted from 10-day-old seedlings by using an RNA extraction kit (RNeasy Maxi Kit, Qiagen) in accordance with the manufacturer's instructions. The concentration and quality of RNA samples were examined by spectrophotometer (Nanodrop ND-1000; Nanodrop Technologies) and a model 2100 Bioanalyzer (Agilent Technologies). Before the RT reaction, total RNA was treated with amplification-grade DNase I (Invitrogen) for 15 min to degrade possibly contaminated residual genomic DNA. 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. About 1/20 of the first-strand cDNA generated from 1 μg total RNA was used as a template for PCR in a reaction volume of 50 μL with rTaq DNA polymerase (Takara). PCR was performed in an ABI 9700 Thermocycler (Applied Biosystems) with the following cycling profile: 94 °C for 3 min; 25 to 35 cycles (depending on the expression levels of the different genes) at 94 °C for 50 s, 55 °C for 50 s, and 72 °C for 1 min. Fifteen microliters of the PCR product was separated in a 1% agarose gel and stained with ethidium bromide for visualization. We used a pair of primers specific to the rice *actin* gene (LOC_Os05g36290) for RT-PCR as an internal control. We checked the transcript levels of the *actin* gene and there was no significant variation among different RNA samples. This gene has also been used as the internal control in other RT-PCR experiments in our laboratory (Satoh et al., 2010). A pair of primers for each *OsNAC* gene (Supplementary Table S2) with a 500- to 600-bp amplicon was used for RT-PCR. All RT-PCRs were repeated three times with independently reverse-transcribed templates.

3. Results

3.1. Identification of OsNAC genes

To identify *OsNAC* genes in the rice genome we used four public databases: MSU, NCBI, DRTF, and KOME. By removing redundant sequences from the four databases and different transcripts of the same gene, preliminarily we identified 155 putative *OsNAC* genes, including 16 (AK068153, AK107407, AK108080, AK071020, AK062952, AK109860, AK063943, AK102902, AK069733, AK069423, AK072275, AK068393, AK064178, AK288271, AK287632, and AK119495) found only in the KOME and MSU DNA searches (rice full-length searches). The protein sequences encoded by all 155 putative *OsNAC* genes were confirmed by SMART and Pfam searches for the presence of the NAM domain. For convenience, the “LOC_” prefix has been omitted from the MSU locus IDs in the rest of the manuscript. Four genes (Os01g23710, Os10g25630, Os10g26250 and Os04g45340) with no NAM domain and one gene (Os09g24670) with only a very small domain were excluded from further analysis, as they cannot be used to construct acceptable phylogenesis. Therefore, there were at least 151 *OsNAC* genes in the rice genome, of which 135 had corresponding locus IDs in the MSU and DRTF databases. The number of NAC proteins in rice (151 members) is greater than that in Arabidopsis (117 members). Detailed information and the structures of representative *OsNAC* genes can be found in Supplementary Table S3.

3.2. Phylogenetic analysis and multiple sequence alignments

To clarify the phylogenetic relationships among the *OsNAC* genes and infer the evolutionary history of this gene family, a combined phylogenetic tree was constructed with the aligned OsNAC domain and sixteen NAC published domain sequences of different species (Fig. 1). Those of published NAC domain proteins bound specifically to CATGTG motif of promoter region (Tran et al., 2004) or act as a functional motif and an activation domain (Oh et al., 2005). The conserved NAC domain consisting of approximately 160 residues is located in N-terminal region and NAC domain proteins are highly involved in stress responses than C-terminal. Examination of the tree revealed that the OsNAC family fell into two major groups (group A and group B). Group A and group B were further divided into 7, and 9 subgroups, respectively. The subgroups are named by the presence of known NAC proteins in the respective subgroup. In group A, the subgroups were ONAC4, ONAC5, ONAC2, ONAC3, ONAC7, ONAC1, and ONAC6. Likewise, the SNAC (Stress-associated NAC), ANAC34, NEO, SND, NAC22, NAC1, NAM/CUC3, OMNAC (*Oryza* membrane NAC), and TIP (*Turnip crinkle* virus interacting protein) subgroups were assigned to group B. Group A consisted of 65 *OsNAC* genes; 14 belonged to ONAC4, three to ONAC5, sixteen to ONAC2, seventeen to ONAC3, five to ONAC7, eight to ONAC1 and two to ONAC6. Group A does not include any analyzed members of NAC and its members have less homology with the Arabidopsis NACs. The members showed diversity in the NAC domain structure. Group B consisted of 86 *OsNAC* genes; 14 members of SNAC, thirteen of ANAC34, two of NEO, fourteen of SND, five of NAC22, six of NAC1, seventeen of NAM/CUC3 one of OMNAC, and fourteen of TIP. The NAC members of group B showed high homology with the Arabidopsis and others NACs.

With the help of comparative genomics, it is possible to analyze proteins corresponding to the same gene family among different species. To evaluate the phylogenetic relationship among the NAC proteins in rice and Arabidopsis, we constructed another unrooted tree from alignments of all 268 NAC protein sequences of the two species as described in the Materials and methods (Supplementary Fig. 1). This phylogenetic analysis suggested that most OsNAC and AtNAC proteins were clustered in many plant-specific clades with very high bootstrap support. We found three pairs of orthologous proteins with 1000 replicates: Os06g15690 and AT1G25580, Os11g45950 and AT3g04410, and Os09g32260 and AT2G24430. We identified 34 pairs of homologous/paralogous proteins in rice and 29 pairs in Arabidopsis (e.g. Os01g48130 and Os05g48850; AT3G01600 and AT5G14490). Paralogous genes were derived in both species by gene duplication. This result suggested that some members of the *OsNAC* and *AtNAC* gene families originated from the same ancestral genes before divergence of monocots and dicots.

We then compared the NAC proteins in rice with those in maize, sorghum, poplar, and Arabidopsis. We accumulated 747 putative non-redundant NAC protein sequences, comprising 135 sequences from rice, 216 from maize, 112 from sorghum, 167 from poplar, and 117 from Arabidopsis. We constructed a phylogenetic tree based on the protein sequences of the *NAC* gene family. Examination of the phylogenetic tree revealed that the plant-specific *NAC* gene family became highly diversified during evolution. Fourteen *ONAC2* genes out of 16 (blue dots in Supplementary Fig. 2) and 12 *ONAC3* genes out of 17 (pink dots in Supplementary Fig. 2) were plant-specific. The NAC domain composition of ONAC2 and ONAC3 was tightly conserved (data not shown). *NAC* gene family analysis of tobacco has been reported by Rushton et al. (2008) and showed that tobacco *NAC* genes (TNAC) are Solanaceae specific.

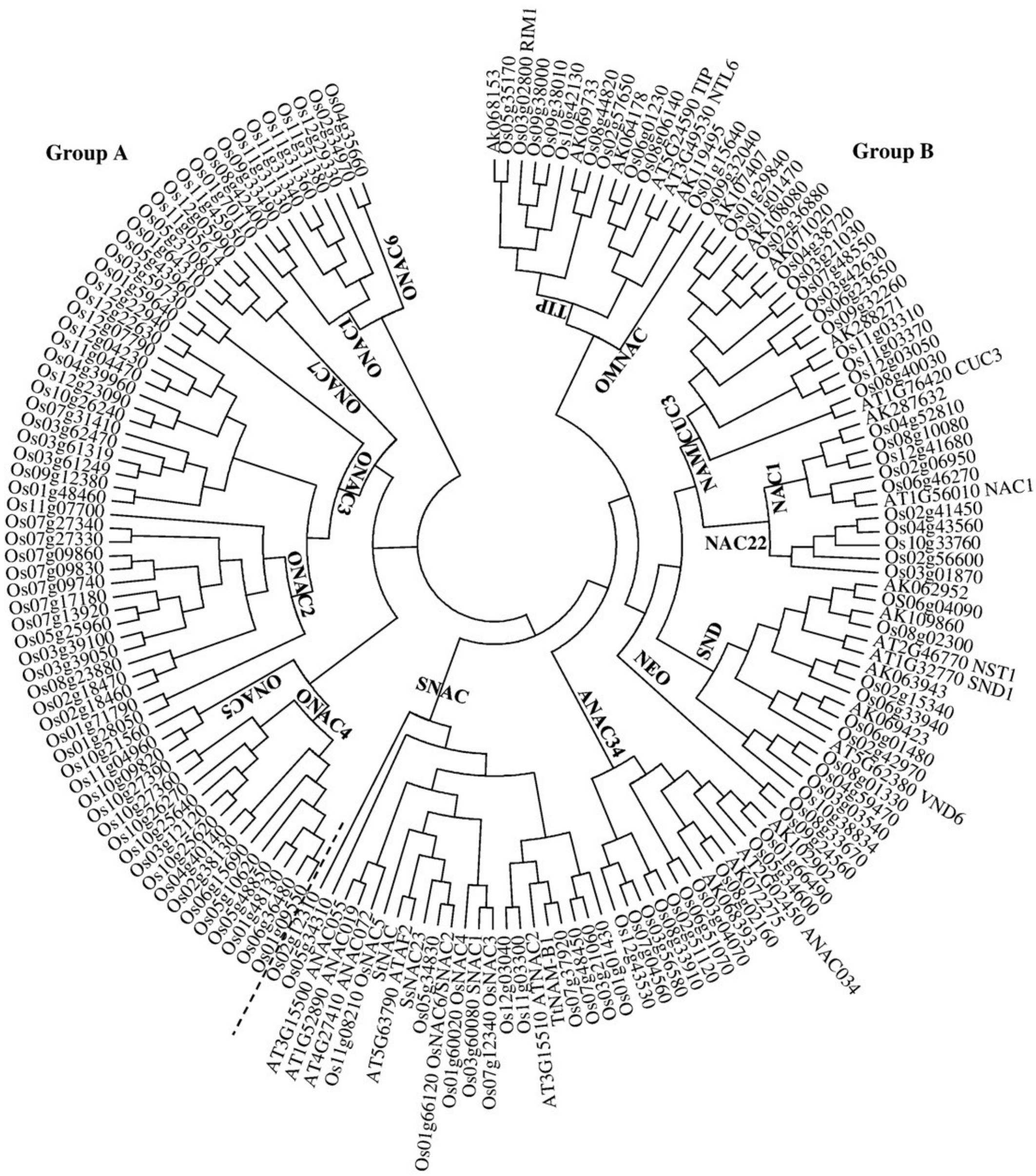

Fig. 1. Evolutionary relationship among the rice OsNAC domain sequences. The unrooted tree was generated by using the ClustalX program with the neighbor-joining method. Bootstrap values from 1000 replicates are indicated at each node. OsNAC proteins were allocated to two distinct groups (A and B).

3.3. Chromosomal locations of OsNAC genes

To examine the genomic distribution of *OsNAC* genes on the chromosomes, we identified the position based on MSU database search. The position and length of each gene are shown in Supplementary Table S3. In total, 135 of the 151 *OsNAC* genes could be localized on the 12 chromosomes, with obviously uneven distribution. *OsNAC* genes are present 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. 2) Chromosome 3 had the largest number (17) of *OsNAC* genes, followed by 15 on chromosome 1. In contrast, only seven *OsNAC* genes were found on each of chromosomes 4 and 9. All seven *OsNAC* genes on chromosome 4 were found on the long arm. Eight out of the 13 *OsNAC* genes on chromosome 11 were located on the short arm, and six *OsNAC* genes out of the seven located on chromosome 9 were placed on the long arm. Fewer than 10 *OsNAC* genes were found on chromosomes 4, 5, 6, and 9. Two *OsNAC* genes (Os03g39100 and Os03g39050) encoding proteins having only the NAC domain were positioned around the centromere on chromosome 3 and showed tandem duplication. Additionally, nine *OsNAC* genes (Os01g01430, Os01g01470, Os03g03540, Os06g01230, Os08g01330, Os11g03300, Os11g03310, Os12g03040, and Os12g03050) were located near the telomeric regions. Six of these genes, on chromosomes 1, 11, and 12, showed tandem duplication in the telomeric regions.

During evolution, both segmental duplication and tandem duplication have contributed to the large number of gene families in plants (Cannon et al., 2004). To elucidate the potential mechanism of

Fig. 2. Distribution of *OsNAC* genes on the 12 rice chromosomes. The scale is in megabases (Mb). Chromosome numbers are indicated at the top of each bar. White dot on each chromosome (vertical bars) shows the rough position of the centromere. Genes with open reading frames in opposite orientations are marked on the chromosome (the black triangles represent upward and the blue triangles indicate downward). Straight lines connect the *OsNAC* genes presented on duplicated chromosomal segments, and tandemly duplicated gene clusters are marked by green bars.

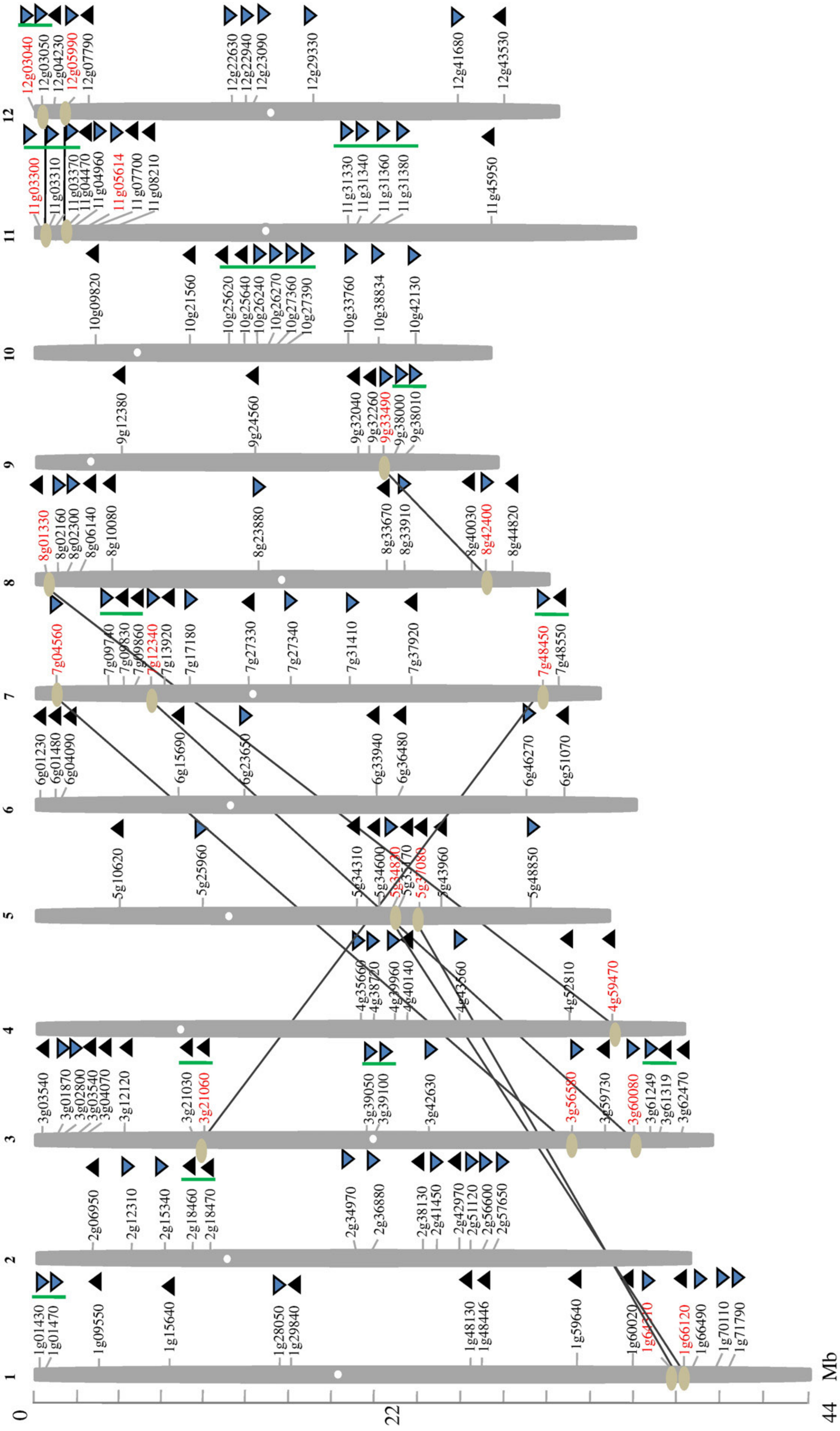
0
22
44
Mb
1
1g01430
1g01470
1g09550
1g15640
1g28050
1g29840
1g48130
1g48446
1g59640
1g60020
1g64310
1g66120
1g66490
1g70110
1g71790
2
2g06950
2g12310
2g15340
2g18460
2g18470
2g34970
2g36880
2g38130
2g41450
2g42970
2g51120
2g56600
2g57650
3
3g03540
3g01870
3g02800
3g03540
3g04070
3g12120
3g21030
3g21060
3g39050
3g39100
3g42630
3g56580
3g59730
3g60080
3g61249
3g61319
3g62470
4
4g35660
4g38720
4g39960
4g40140
4g43560
4g52810
4g59470
5
5g10620
5g25960
5g34310
5g34600
5g34870
5g35170
5g37080
5g43960
5g48850
6
6g01230
6g01480
6g04090
6g15690
6g23650
6g33940
6g36480
6g46270
6g51070
7
7g04560
7g09740
7g09830
7g09860
7g12340
7g13920
7g17180
7g27330
7g27340
7g31410
7g37920
7g48450
7g48550
8
8g01330
8g02160
8g02300
8g06140
8g10080
8g23880
8g33670
8g33910
8g40030
8g42400
8g44820
9
9g12380
9g24560
9g32040
9g32260
9g33490
9g38000
9g38010
10
10g09820
10g21560
10g25620
10g25640
10g26240
10g26270
10g27360
10g27390
10g33760
10g38834
10g42130
11
11g03300
11g03310
11g03370
11g04470
11g04960
11g05614
11g07700
11g08210
11g31330
11g31340
11g31360
11g31380
11g45950
12
12g03040
12g03050
12g04230
12g05990
12g07790
12g22630
12g22940
12g23090
12g29330
12g41680
12g43530

Table 1
The 11 genes expressed (upregulated) by at least three abiotic treatments.

Upregulated gene	Cold	Drought	Submergence	Laid-down submergence	Osmotic	Salt	50 μM ABA	50 μM GA
Os12g41680	1	1	1	1				
Os02g34970		1	1	1				
Os07g48450		1	1	1			1	
Os01g01430		1	1	1			1	
Os07g48550	1	1	1	1				
Os11g03300	1	1	1	1				
Os12g03040	1	1	1	1				
Os01g66120	1	1	1	1				
Os05g34830	1	1	1	1				
Os03g21030	1	1	1					
Os01g48460		1	1	1				

Genes upregulated more than 1.5 fold were assigned a value of 1.

evolution of the *OsNAC* gene family, we analyzed segmental and tandem duplication events. We found nine segmental duplication events (Fig. 2, Supplementary Table S4). Although the ONAC3 subgroup contained the greatest members of *OsNAC* genes, none of the *OsNAC* genes belonging to this subgroup was on duplicated chromosomal segments on the chromosomes. Similarly, in the second-largest subgroup, ONAC2 no pairs of segmentally duplicated genes were found. No segmental duplication events occurred among the three chromosomes 2, 6, and 10. The highest numbers of segmentally duplicated genes were found in subgroups SNAC and ONAC7, whereas only one pair of *OsNAC* genes belonging to subgroups ANAC34, SND, and ONAC1 was found on duplicated chromosomal segments.

On the basis of the phylogenetic tree results, we analyzed the intron and exon structures of *OsNAC* genes belonging to the above subgroups. There were no obvious differences among the different subgroups, except in the case of subgroups ONAC2 and ONAC3 (Supplementary Fig. 3). Both the location and numbers of introns and exons of the *OsNAC* gene family were diversified on each gene. The highest number of introns and exons was found in subgroups ONAC4 and TIP (Supplementary Fig. 3). The coding sequences of all the *OsNAC* genes were disrupted by introns: the number of introns varied from 2 to 16 in subgroup ONAC4 and from 2 to 6 in subgroup TIP (Supplementary Fig. 3). No introns were observed in subgroups ONAC2 and ONAC3, except in five genes (Os03g39050, Os10g26240, Os12g23090, Os11g04470, Os01g59640, Os03g59730 and Os05g43960). Most of the genes in subgroups NAM/CUC3, NAC1, NAC22, SND, ANAC34, and SNAC had two introns. In rice, researchers have shown 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; Roy and Penny, 2006). Therefore, it seems that the genes in subgroups SND, ANAC34, SNAC, ONAC7, and ONAC1 may represent young members of this gene family; in each subgroup the genes with more introns may be the original genes from that subgroup or may have diversified from subgroup ONAC4 or subgroup TIP (Supplementary Fig. 3). However, the mechanisms of intron gain and loss are not yet clear (Lin et al., 2006).

Tandem duplication is a common source for the expansion of gene families. We found 36 *OsNAC* genes involved in tandem duplications, consisting of 18 pairs (Fig. 2). The number of *OsNAC* genes arranged in the tandem repeats varied from two to four, and most of these genes were assigned in the same orientation on the chromosomes, with the exception of two pairs (Os07g09740 and Os07g09830, and Os07g48450 and Os07g48550). The identity of the protein sequences of these genes ranged from 42% to 86% (data not shown). Information on the duplicated genes is given in Supplementary Table S5.

3.4. Expression profiles with abiotic treatments

During their life cycles, plants frequently encounter external stress conditions that adversely affect growth and development. To investigate the response of the *OsNAC* genes to both abiotic and biotic stresses, we performed a microarray analysis. We identified 46 non-redundant *OsNAC* genes that were expressed up- and downregulation under one or more of the abiotic stresses (cold, drought, submergence, laid-down submergence, osmotic, salt and hormone) compared with expression in the control seedlings (Supplementary Table S6). More than 50% of these genes were upregulated under different stresses. The greatest number of upregulated genes (21) was found under laid-down submergence stress. The lowest number of upregulated gene (1) was observed under osmotic stress and under salt stress. None of the genes was up- and down regulated under ABA treatment. The greatest number of downregulated genes (14) was found under laid-down stress, whereas the lowest number (3) was found under submergence stress. We identified 11 genes upregulated by at least three abiotic treatments (Table 1). Interestingly, under the different abiotic treatments, changes in the expression of eight genes were very common; these genes showed upregulation under four different abiotic treatments (Table 1). Five genes (Os01g01430, Os03g04070, Os04g43560, Os05g37080, and Os07g48450) were common and showed up- and downregulation under different abiotic stresses. Likewise, six genes (such as Os01g66120 and Os05g34830) up-and down regulated in the different time points of laid-down submergence stress (Supplementary Table S6). Two genes were downregulated by at least three abiotic treatments (Table 2). Seven genes of the ONAC4 subgroup were downregulated under the three abiotic stresses, whereas upregulation of a number of genes belonging to the TIP, NAM/CUC3, NAC1, and SNAC subgroups was common in response to various abiotic stresses (Supplementary Table S6). In the SNAC subgroup, expression of the Os01g01430, Os01g66120, Os05g34830, Os07g48450, Os11g03300, and Os12g03040 genes was induced by four abiotic stresses (Supplementary Table S6). Likewise, Os07g48550 and Os12g41680 genes in subgroup NAM/CUC3, and subgroup NAC1 showed upregulation under four abiotic stresses (Supplementary Table S6). Only a very few upregulated genes were observed in subgroups ONAC1, NAC22, and ONAC3. Os01g66120, an ortholog of the sugarcane *NAC* gene *SsNAC23*, was upregulated under cold, drought, submergence, and laid-down submergence stresses. The search for *cis*-elements (http://hpc.irri.cgiar.org/tool/nias/ces; Doi et al., 2008) identified ABRE and CBE motifs in Os01g66120. Nakashima et al. (2007) have reported that these motifs are induced by drought, cold, and salt. The above genes (such as Os07g48550 and Os12g41680) contained ABRE and CBE motifs in its upstream region and its expression might be induced by stress. These genes were upregulated under four different abiotic stresses in our study. A search for *cis*-elements within the promoter region of SNAC subgroup members revealed the binding motifs for leucine zipper (bZIP) factors and helix–loop–helix (bHLH) factors.

Table 2
The 2 genes expressed (downregulated) by at least three abiotic treatments.

Downregulated gene	Cold	Drought	Submergence	Laid-down submergence	Osmotic	Salt	50 μM ABA	50 μM GA
Os05g48850	−1	−1	−1	−1				
Os01g64310	−1	−1	−1					

Those downregulated less than 1.5 fold were assigned a value of −1.

3.5. Expression profiles under biotic treatments

To investigate the responses of *OsNAC* genes to biotic stresses, we investigated the expression profiles of *OsNAC* genes in seedlings treated with RSV or RTSV and expressed the results as fold changes with respect to the controls. We identified 34 non-redundant genes differentially expressed under RSV stress at 3, 6, 9, and 12 DAI (Supplementary Table S7). Nineteen of these genes were upregulated, whereas 15 were downregulated. Five genes (Os04g40140, Os05g10620, Os11g05614, Os12g03050, and Os12g05990) were expressed at high levels (five-fold compared with the controls); among them, Os12g05990 was expressed at the highest level (nine-fold). Surprisingly, under RSV stress, the number of genes upregulated was highest at 12 DAI and then decreased in the order of 9 DAI, 6 DAI, and 3 DAI (Fig. 3A). We also identified 24 non-redundant genes differentially expressed under RTSV stress at 6 DAI, 9 DAI, 12 DAI, and 15 DAI (Supplementary Table S8). Among them, 13 genes were upregulated and 11 were downregulated. Interestingly, only one gene (Os05g37080) was upregulated at all DAI. Under RTSV treatment, the number of genes upregulated was greater at 15 DAI but was smaller at 9, 6, and 12 DAI (Fig. 3B). Six genes (Os02g34970, Os02g38130, Os11g03310, Os11g03370, Os11g05614, and Os12g03050) were very common; these genes were upregulated under both RSV stress and RTSV stress. From the phylogenetic tree, we found that expression of most of the genes of subgroup NAM/CUC3, and subgroup SNAC, was induced by both abiotic and biotic stresses, whereas five genes from subgroup TIP showed upregulation under RSV stress. Os07g48450 gene was induced by four abiotic stresses (dorought, submergence, laid-down submergence, and 50 μM GA), and one biotic RTSV stress.

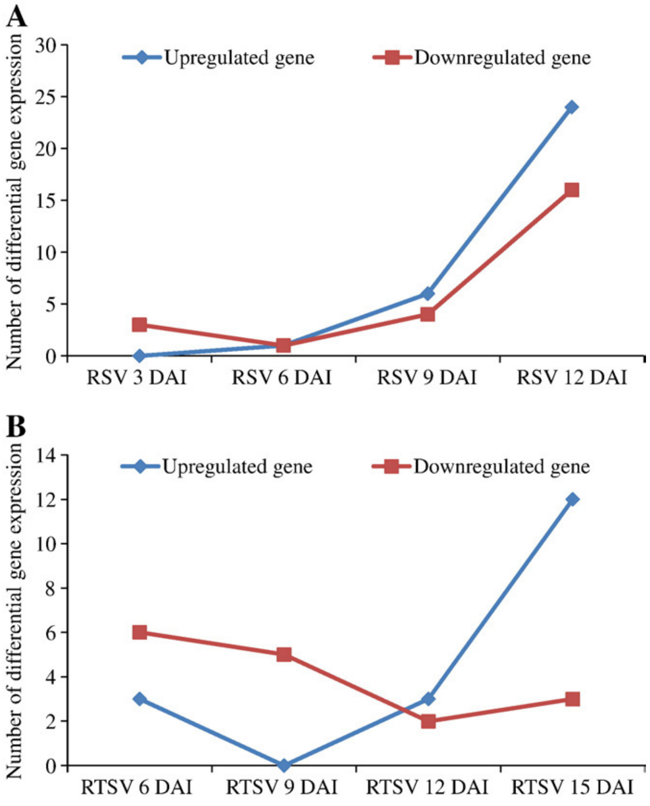

Fig. 3. Number of genes differentially expressed (A) under RSV stress at different days after inoculation (DAI), i.e., 3 DAI, 6 DAI, 9 DAI, and 12 DAI; and, (B) under RTSV stress at 6 DAI, 9 DAI, 12 DAI, and 15 DAI. X-axis represents the different days of inoculation and Y-axis indicates the number of differential gene expression.

3.6. Comparison of expression intensities of duplicated OsNAC genes

Gene expression divergence may play an important role in the preservation of duplicated genes. Lynch and Conery (2000) have found that during the course of evolution some of the alternative outcomes of duplicated genes have included (i) nonfunctionalization, (ii) neofunctionalization, and (iii) subfunctionalization. We examined the expression patterns of segmentally and tandemly duplicated genes under the different stress conditions. Probe numbers were available for six of the nine pairs of genes located in the segmentally duplicated regions. Out of six pairs of genes, five showed similar, high expression patterns under most of the tested stresses, indicating subfunctionalization after the duplication events (Fig. 4A). We identified 14 clusters of tandemly duplicated *OsNAC* genes (Supplementary Table S5). Among them only four clusters of the gene probe set were found in our 22K microarray data. The expression patterns of two tandemly duplicated genes (Os07g48550 and Os07g48450) were dissimilar, which may indicate neofunctionalization. Three clusters of genes showed similar expression, which may indicate subfunctionalization (Fig. 4B).

4. Discussion

4.1. Function of different subgroups

The objectives of this study were (i) to perform an annotation of the *OsNAC* gene family, (ii) to provide a new tool to predict the group based classification in plants, (iii) to predict number of genes response to different stresses, especially biotic stresses are novel study in this manuscript, and (iv) to select the best candidate genes for functional analysis. We have presented more specific phylogenetic analysis of NAC proteins, based on the DNA-binding domains. Group A, an unknown and group B, known are two major groups in phylogenetic analysis (Fig. 1). These are the inconsistencies in our *NAC* gene classification than previous study. Using this approach, the subgroups fit well with known NAC function classes. Proteins with similar domains may have the same or similar biological functions (Lin et al., 2007). For example, the different subgroups of group B in Fig. 1, represent four well known functions of the NAC proteins (i) regulation of plant cell wall development (ii) cell division and expansion (iii) organ formation or initiation, and (iv) response to abiotic or biotic stress. Therefore, we speculate that there is functional redundancy among the *NAC* genes. *OsNAC* genes play crucial roles in various developmental processes, including signaling, stress responses and plant defenses. Membrane-associated NAC members are clustered in two subgroups (TIP and OMNAC). Subgroup TIP has known members of TIP and NTL. TIP NACs were implicated stress

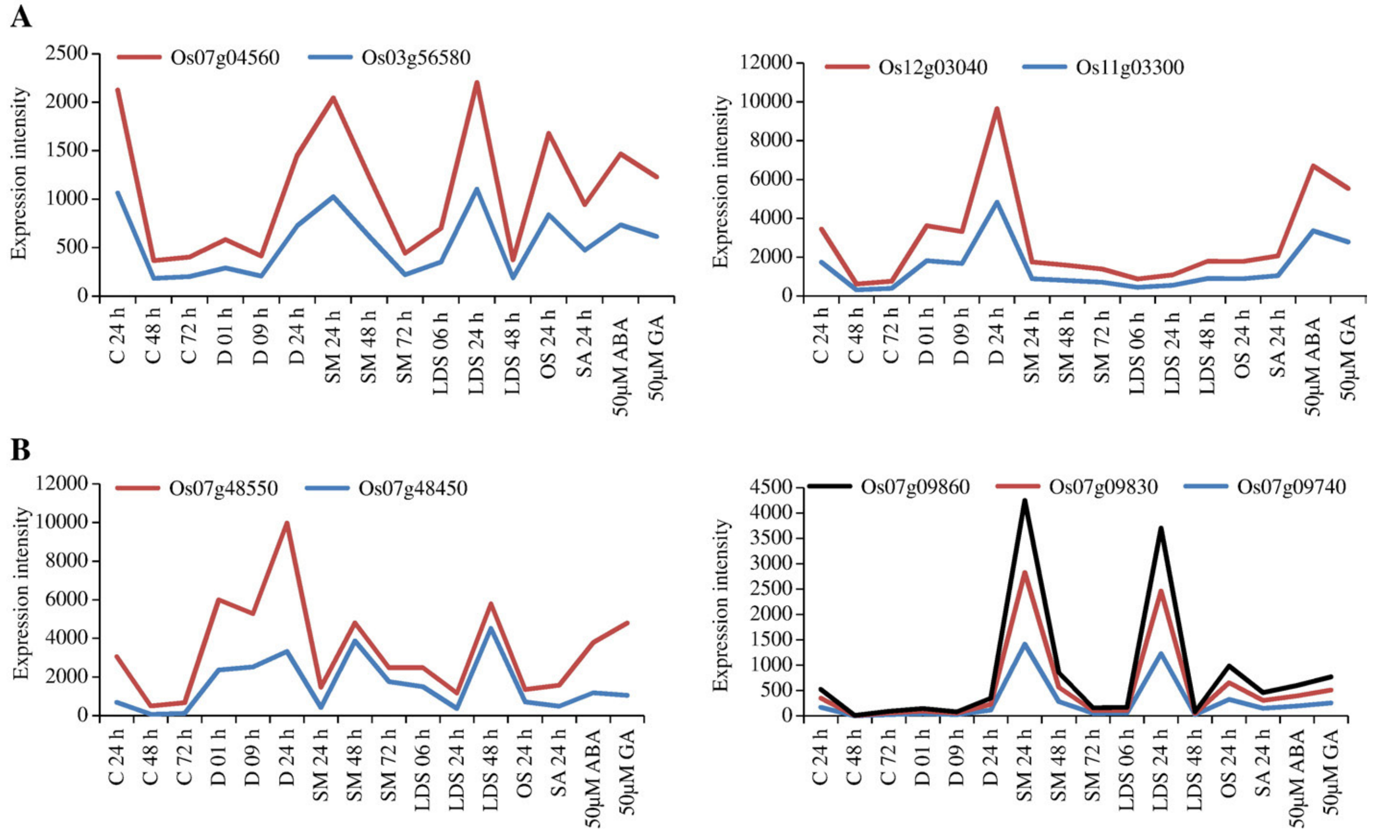

Fig. 4. Examples of expression patterns of *OsNAC* genes found in duplicated regions of the rice genome. (A) Expression patterns of some *OsNAC* genes found in segmentally duplicated regions. (B) Expression patterns of some tandemly duplicated *OsNAC* genes. X-axis represents the different stresses and Y-axis indicates the expression intensity. Gene names are listed at the top of each graph.

response (NTL6) and play various roles in a variety of plant growth and development processes (Kim et al., 2007a,b), and TIP acts as a positive regulator of the host basal defense responses (Yoshii et al., 2009). In our study, eight genes of this subgroup showed upregulation in both abiotic and biotic stresses. Some genes (Os01g15640, Os06g01230, Os08g06140, and Os10g42130) of this subgroup share high similarity with NTL6 and RIM1. It is assumed that members of this subgroup maybe involved in morphogenesis such as stress, plant growth and development processes. Very recently, Kim et al. (2010) reported that subgroup TIP was induced by different abiotic stresses. Arabidopsis TIP is involved in viral resistance (Ren et al., 2000). The subgroup NAM/CUC3 includes known NAC protein CUC3 and other members in that subgroup also have high homology with Arabidopsis NAM/CUC3 proteins. NAM/CUC3 is involved in shoot apical meristem formation and development (Kim et al., 2007b). Eight genes of this subgroup are induced by abiotic and biotic stresses. The subgroup NAC1 (a known member from Arabidopsis ANAC021) members are similar to ANAC021/22. The subgroup NAC22 members are similar to NAC22 of Arabidopsis. The members in the subgroup NAC22 also share a conserved N-terminal region with NAC1 members. It is speculated that members in subgroups NAC1 and NAC22 may be involved in morphogenesis such as root development. Subgroup SND includes characterized by NAC members, NST1 and SND1, but none of the genes is induced by different stresses. NST1 and SND1 are involved in secondary cell wall synthesis (Zhong et al., 2007). Most of the members of this subgroup share high similarity with NST1 and SND1, which may be involved in cell wall related functions in rice. The SNAC subgroup include known NAC members from rice, sugar cane (*SsNAC23*), Arabidopsis (*ATAF2*, *ANAC55*, *NAC2*, *ANAC19*, and *ANAC72*) and wheat (*TtNAM-B1*) that are involved in abiotic stress responses (Fang et al., 2008), and four *SNAC* genes were upregulated under either RSV stress or RTSV stress in this study (Fig. 1), which may indicate that *SNAC* genes induced by biotic stresses also. From our array analysis we present more information that rice *NAC* genes induced by abiotic and biotic stresses. Some reports suggested that *CaNAC1*, *BnNACs*, and *OsNAC6* are members of subgroup SNAC share common function in the plant induction response to biotic and abiotic stresses (Hegedus et al., 2003; Oh et al., 2005; Nakashima et al., 2007). Our microarray experiments showed that nine genes of subgroup SNAC were induced by the infection of RBSDV (Rice black-streaked dwarf fijivirus), RGSV (rice grassy stunt tenuivirus), and RRSV (rice ragged stunt oryzavirus) while, all the members of subgroup TIP were induced by the infection of RGSV and RDV (rice dwarf phytoreovirus; Nuruzzaman et al., unpublished results). These results suggested that *OsNAC* gene might play an important role in the cross talk of different kinds of stresses. Transcriptome analyses have generated considerable data, which show extensive overlapping on gene expression between biotic and abiotic stresses. Many of these overlapped genes encode signaling components, including transcription factors and protein kinases (Schenk et al., 2000; Seki et al., 2002). Transcirption factors are promising candidates of common player involved in crosstalk between abiotic and biotic stress signaling. There are high homology with known genes and tight clustering of members of each subgroup. The members of above subgroup are involved in stress response which is confirmed by the expression analyses at mRNA level in this manuscript. Through our phylogenetic analysis, we have predicted many candidate genes that are function in controlling stress responses. Although phylogenetic analysis provides important bioinformatics support for candidate gene selection, we are alert that it alone cannot clearly indicate function. For this reason, our tissue expression analysis at the level of mRNA transcription and the conserved group-specific residues of the NAC domain defined in this study are good candidates for future investigation to further understand the structure–function relationships of NAC transcription factors.

4.2. Gene multiplication

We identified 151 *OsNAC* genes in the rice genome that could be classified into different subgroups and were distributed on all 12 chromosomes. Gene duplication is believed to be a major mechanism for the establishment of new gene functions and is one of the preliminary drivers of the evolution of genetic systems and genomes (Moore and Purugganan, 2003). Segmental duplications occur commonly in the slowly evolving in the *MYB* gene family (Cannon et al., 2004). In the large and quickly evolving NBS-LRR disease-resistance gene family, tandem duplications in local genomic clusters with low levels of retention of segmental duplications are common (Cannon et al., 2004). Our analysis suggested that there were more tandemly duplicated genes than duplicated chromosomal segments in the *OsNAC* gene family. Therefore, we consider that the number of *OsNAC* gene members increased rapidly during the course of evolution, and the tandem duplications of chromosomal regions played a key role in the expansion of this family. This phenomenon has been found in the F-box family of rice genes (Jain et al., 2007). Interestingly, most of the segmentally duplicated genes

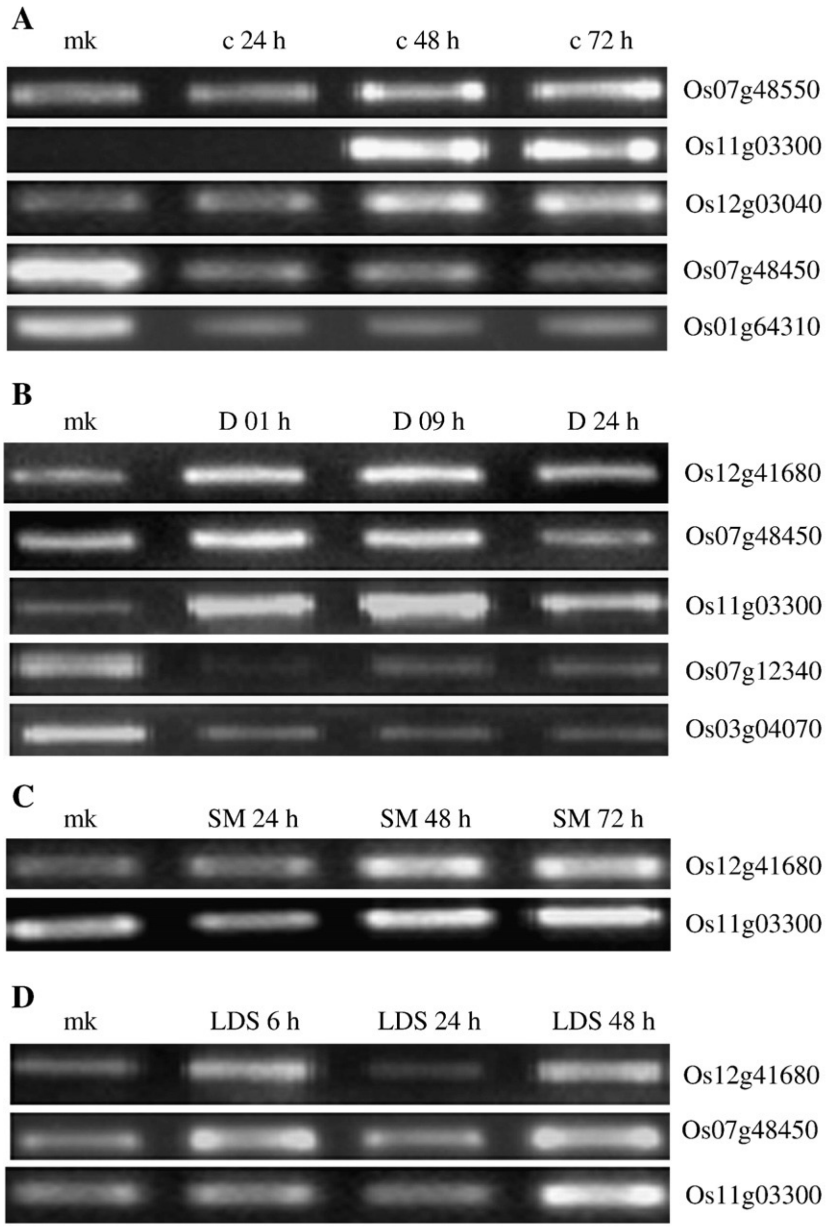

Fig. 5. RT-PCR analysis of differential gene expression under abiotic treatments. (A) three genes (Os07g48550, Os11g03300, and Os12g03040) were upregulated and two (Os07g48450 and Os01g64310) were downregulated under cold stress (mk = mock/control; C-24 h, C-48 h, and C-72 h = cold for 24, 48, and 72 h, respectively); (B) of five genes, three were upregulated (Os12g41680, Os07g48450, and Os11g03300) to high levels and two were downregulated (Os07g12340 and Os03g04070) to low levels under drought stress (mk = mock/control; D-01 h, D-09 h, and D-24 h = drought for 01, 09, and 24 h, respectively); (C) Os12g41680 and Os11g03300 genes were expressed at high levels under submergence stress (mk = mock/control; SM-24 h, SM-48 h, and SM-72 h = submergence for 24, 48, and 72 h, respectively); (D) three genes (Os12g41680, Os07g48450, and Os11g03300) were upregulated under laid-down submergence stress (mk = mock/control; LDS-06 h, LDS-24 h, and LDS-48 h = laid-down submergence for 06, 24, and 48 h, respectively).

belonged to the SNAC subgroup and ONAC7 subgroup. The finding that the number of members of *SNAC* gene (14) in rice is less than that in Arabidopsis (17) suggests that *SNAC* gene members in Arabidopsis have undergone multiple duplication events to gain more members during the process of evolution (Supplementary Fig. 1).

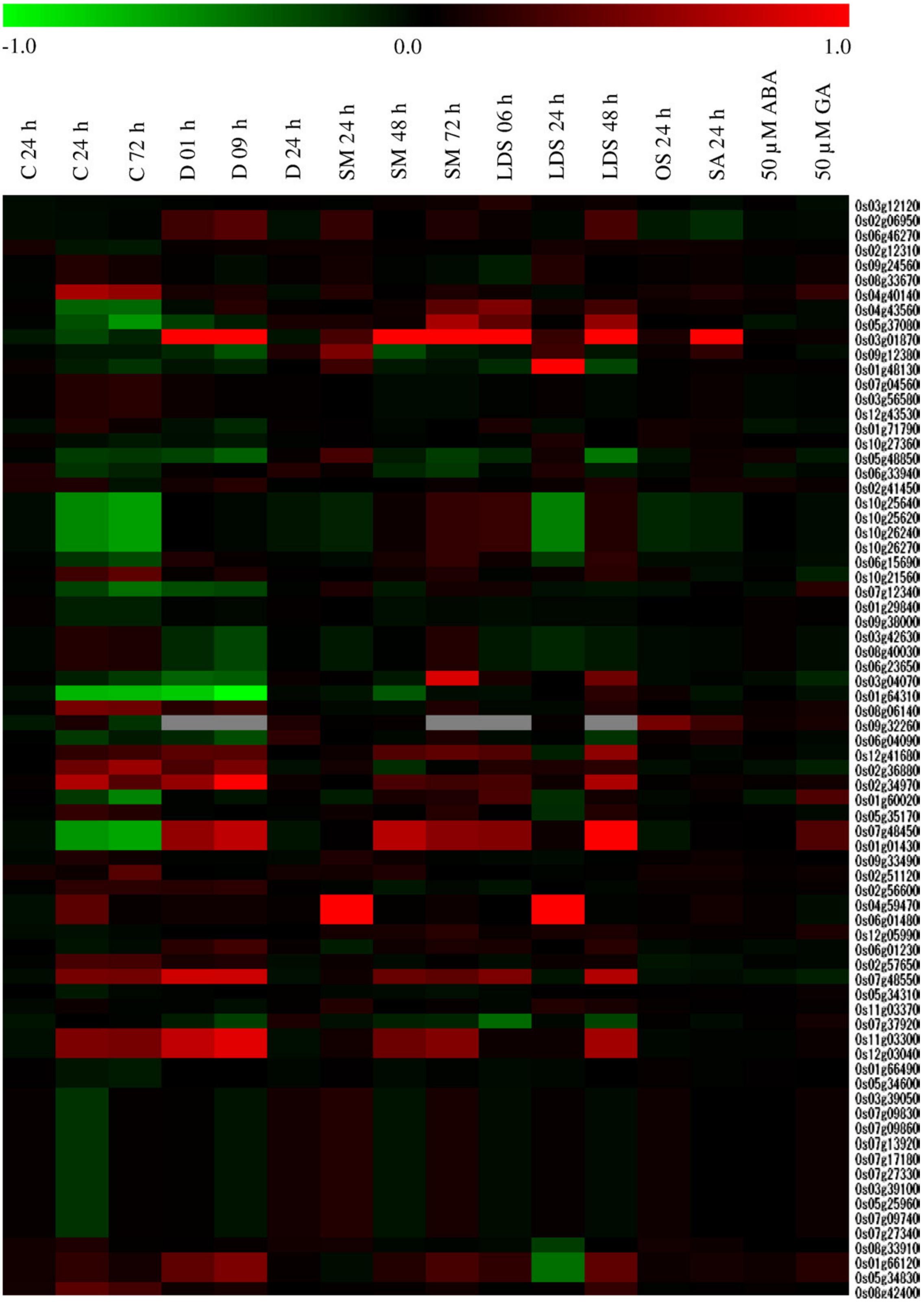

Fig. 6. Differential expression of *OsNAC* genes under eight types of abiotic stress (log10 ratio); color bar at top shows level of expression. Red indicates expressed genes and green indicates unexpressed ones; C-24 h, C-48 h, and C-72 h = cold for 24, 48, and 72 h, respectively; D-01 h, D-09 h, and D-24 h = drought for 01, 09, and 24 h, respectively; SM-24 h, SM-48 h, SM-72 h = submergence for 24, 48, and 72 h, respectively; LDS-06 h, LDS-24 h, and LDS-48 h = laid-down submergence for 06, 24, and 48 h, respectively; OS = osmotic 24 h; SA = salt 24 h; ABA = 50 µM abscisic acid; GA = 50 µM gibberellic acid.

4.3. OsNAC genes play an important role under different stresses

In this paper, we show that in detailed expression of rice NAC is induced by abiotic stress as well as biotic stresses. To obtain an overview of the expression-level changes in the rice NAC family under stress conditions, we extracted the stress signal ratios of the *OsNAC* genes from our Nipponbare microarray data for cold, drought, submergence, laid-down submergence, osmotic, salt, and hormone stresses in 10-day-old seedlings, and in 30-day-old calli. Determining the expression profiles of the *OsNAC* genes may provide clues to the function of each gene under both abiotic and biotic conditions. Our analyses suggested that a total of 46 genes were expressed (up or down) under at least one of the seven abiotic stress conditions. For confirmation we performed an RT-PCR analysis to determine the mRNA expression levels; all genes had strikingly higher or lower expression profiles in seedlings than in the controls (mock) (Fig. 5A–D). Similarly, MeV cluster analysis based on log10 ratio values showed that the *OsNAC* genes had very diverse expression patterns (Fig. 6, Supplementary Table S1). In order to gather more details of expression profiles under large abiotic and biotic stresses, we included submergence, laid-down submergence, osmotic, hormone, RSV, and RTSV stresses, which were not considered by Fang et al. (2008). In addition, we found that the Os10g42130 genes of the TIP subgroup; Os06g46270 and Os12g41680 genes of the NAC1 subgroup; Os07g48450 and Os11g03300 genes of the SNAC subgroup were upregulated under different abiotic and biotic stresses (most of the genes of subgroup TIP and SNAC are induced by the infected of virus, unpublished results). Thus, our findings might help to increase our understanding of gene responses in research on abiotic and biotic stress over a broad scale. Expression of some transcription factor genes from different families such as DREB (Liu et al., 1998; Kasuga et al., 1999; Haake et al., 2002; Yamaguchi-Shinozaki and Shinozaki, 2001), MYB (Abe et al., 1997), bZIP (Uno et al., 2000), and zinc finger (Mukhopadhyay et al., 2004) has been reported to improve tolerance in various abiotic stresses. Increasing evidence suggests that some members of the

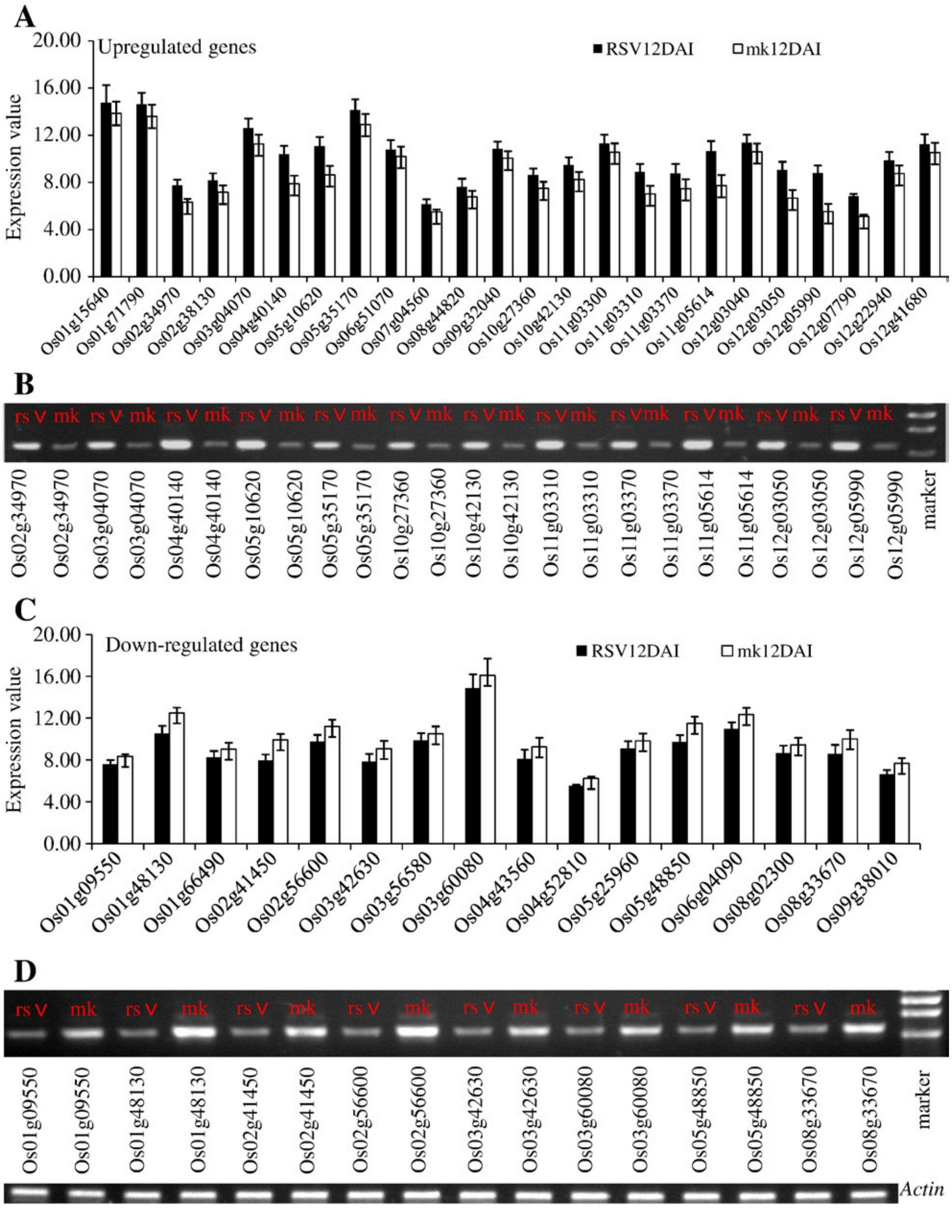

Fig. 7. (A) Histogram of 24 *OsNAC* genes differentially expressed under RSV stress; (B) RT-PCR analysis of 12 genes that showed high expression than control under RSV stress 12 days after inoculation (DAI); (C) Histogram of 16 genes downregulated under RSV stress; (D) RT-PCR analysis of eight genes that showed differential expression in response to RSV stress. *Actin* gene expression was used as an internal control; mk = mock (control). X-axis represents the differentially expressed genes and Y-axis indicates the log10 ratio expression values.

NAC gene family also contribute to abiotic stress tolerance in Arabidopsis and soybean (Tran et al., 2004; Yoo et al., 2007; Tran et al., 2009). Very recently, Jeong et al. (2010) reported that *OsNAC10* enhance abiotic tolerance in rice, which is belong to SNAC subgroup. In our array analysis, *OsNAC10* gene is induced by the infection of RDV, RBSDV, RGSV, and RRSV virus (unpublished results). To date, only six genes of the OsNAC family in rice have been functionally characterized for their roles in abiotic stress tolerance; a large number of the gene functions of this family are still unknown in plants. One of these six genes is the drought stress-responsive gene *OsNAC10*, mentioned above (Jeong et al., 2010). Expression of this gene is induced specifically in guard cells by drought stress, and overexpression in rice can promote stomatal closure and substantially improve drought resistance under field conditions. The other three genes are *OsNAC3*, *OsNAC4*, *OsNAC5*, and *OsNAC6*/*SNAC2*, the expression of three genes are also induced by various stresses; transgenic rice plants overexpressing these genes showed significantly improved tolerance to dehydration stress (Nakashima et al., 2007; Sperotto et al., 2009; Zheng et al., 2009). Plant-specific NAC (NAM, ATAF1/2, CUC2/3) transcription factor play regulatory roles in diverse developmental process and stress responses.

Pathogens attack plants under environmental conditions favoring such pathogens, and plants defend themselves through different signaling pathways that regulate numerous biochemical, metabolic, and molecular mechanisms to increase tolerance to adverse conditions. Previously, NAC proteins have been identified because they interact with other proteins of biological importance. The wheat

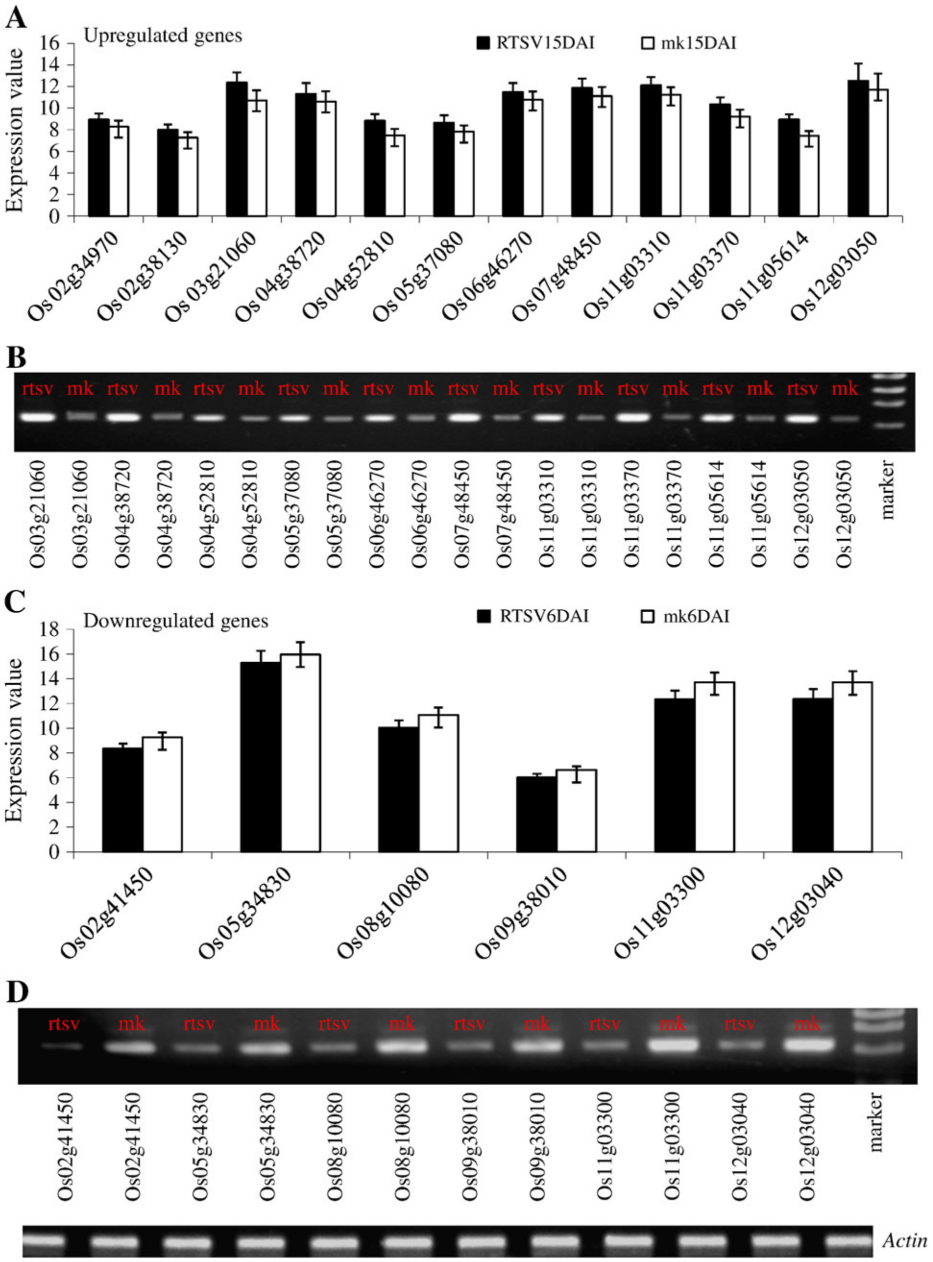

Fig. 8. (A) Histogram of 12 *OsNAC* genes differentially expressed under RTSV treatment; (B) RT-PCR analysis of 10 genes that showed differential mRNA expression under RTSV stress 15 days after inoculation (DAI); (C) histogram of six genes downregulated under RTSV stress; (D) RT-PCR analysis of six genes that showed differential mRNA expression at 6 DAI. *Actin* gene expression was used as an internal control; mk = mock (control). X-axis represents the differentially expressed genes and Y-axis indicates the log10 ratio expression values.

(*Triticum* sp.) geminivirus RepA-binding (GRAB) proteins GRAB1 and GRAB2 were identified because of their ability to interact with the wheat dwarf geminivirus RepA protein (Xie et al., 1999), and the Arabidopsis TIP was identified because of its binding to the *Turnip crinkle* virus capsid protein (Ren et al., 2000; Yoshii et al., 2009). A number of transcription factors, such as WRKY, NAC, and AP2, are known to act specifically on transcriptional regulation of genes responsive to biotic stresses (Gutterson and Reuber, 2004; Ryu et al., 2006; Nakashima et al., 2007; Shimono et al., 2007; Chen and Guo, 2008). We identified a total of 54 genes that were differentially expressed under either of the two biotic (RSV and RTSV) treatments. Among these, 31 genes were upregulated by both treatments. More than half of these genes were studied by RT-PCR analysis to determine mRNA expression levels from our 44K microarray data (Figs. 7, 8). 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 transcription factors was affected at 12 DAI and 15 DAI under both biotic stresses. Data on such expression patterns 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 (Lee and Rudd, 2002; Tameling and Baulcombe, 2007; Li et al., 2009). Host defense systems have been associated with genes for transcription factors in the WRKY family (OsWRKY45; Shimono et al., 2007). Expression of these genes was induced in the late phase of infection, when plants showed obvious symptoms (Fig. 3). The inactivation of defense systems in rice plants in the early infection stage probably stimulates the propagation of RSV and RTSV. Together, above results indicate that some given members of the *OsNAC* gene family is involved in abiotic and biotic stresses condition and increase our knowledge about the involvement of the NAC transcription factor in plant resistance.

5. Conclusion

Taken together, our data suggest that the *OsNAC* gene family has greatly expanded and diversified in monocot and dicot species. Segmental and tandem duplications might have contributed to the expansion of this family. Many subgroups have evolved, resulting in a high level of functional divergence. Most of these subgroups are present in both rice and Arabidopsis, indicating that the appearance of many of the genes in these species preceded the divergence of monocots and dicots. Likewise, some subgroups are present in only one species, suggesting that they have evolved or have been lost in one or more species after this divergence. Interestingly, eight genes were upregulated by at least four abiotic treatments. Six genes were upregulated under both RSV stress and RTSV stress. *OsNAC* genes showed temporal and spatial patterns of expression under different stresses. The preferential expression of *OsNAC* genes provides a new avenue for functional analyses in rice. Further research using, for example, an RNAi strategy or insertion mutagenesis is required to reveal the exact molecular mechanisms or pathways of the *OsNAC* gene family.

Acknowledgments

This study was supported by grants from the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) to M. N, A. M. S, and S. K. It was partly supported by a grant from the Ministry of Agriculture, Forestry, and Fisheries of Japan (Integrated Research Project for Plant, Insect and Animal using Genome Technology, QT1003). R. M. gratefully acknowledges receipt of the Niche Areas of Biotechnology Fellowship to Young Scientists, sponsored by the Department of Biotechnology, Government of India.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2010.06.008.

References

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D., Shinozaki, K., 1997. Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. Plant Cell 9, 1859–1868.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., Tasaka, M., 1997. Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. Plant Cell 9, 841–857.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bhatnagar-Mathur, P., Devi, M.J., Reddy, D.S., Lavanya, M., Vadez, V., et al., 2007. Stress inducible expression of At DREB1A in transgenic peanut (*Arachis hypogaea* L.) increases transpiration efficiency under water-limiting conditions. Plant Cell Rep. 26, 2071–2082.

Cabauatan, P.Q., Cabunagan, R.C., Koganezawa, H., 1995. Biological variants of rice tungro viruses in the Philippines. Phytopathology 85, 77–81.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., 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.

Chen, X., Guo, Z., 2008. Tobacco OPBP1 enhances salt tolerance and disease resistance of transgenic rice. Int. J. Mol. Sci. 9, 2601–2613.

Doi, K., et al., 2008. Development of a novel data mining tool to find *cis*-elements in rice gene. BMC Plant Biol. 8, 20.

Duval, M., Hsieh, T.F., Kim, S.Y., Thomas, T.L., 2002. Molecular characterization of AtNAM: a member of the Arabidopsis NAC domain superfamily. Plant Mol. Biol. 50, 237–248.

Encabo, J.R., et al., 2009. Suppression of two tungro viruses in rice by separable traits originating from cultivar Utri Merah. Mol. Plant-Microbe Interact. 22, 1268–1281.

Fang, Y., You, J., Xie, K., Xie, W., Xiong, L., 2008. Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. Mol. Genet. Genomics 280, 547–563.

Gao, G., et al., 2006. DRTF: a database of rice transcription factors. Bioinformatics 22, 1286–1287.

Greve, K., La Cour, T., Jensen, M.K., Poulsen, F.M., Skriver, K., 2003. Interactions between plant RING-H2 and plant specific NAC (NAM/ATAF1/2/CUC2) proteins: RING-H2 molecular specificity and cellular localization. Biochem. J. 371, 97–108.

Gutterson, N., 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., Zhang, J.Z., 2002. Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. Plant Physiol. 130, 639–648.

He, X.J., Mu, R.L., Cao, W.H., Zhang, Z.G., Zhang, J.S., Chen, S.Y., 2005. AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J. 44, 903–916.

Hegedus, D., et al., 2003. Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Mol. Biol. 53, 383–397.

Hu, H., et al., 2006. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. Proc. Nat. Acad. Sci. U.S.A. 103, 12987–12992.

Jain, M., 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., 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.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K., Shinozaki, K., 1999. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat. Biotechnol. 17, 287–291.

Kikuchi, K., Ueguchi-Tanaka, M., Yoshida, K.T., Nagato, Y., Matsusoka, M., Hirano, H.Y., 2000. Molecular analysis of the NAC gene family in rice. Mol. Genet. Genomics 262, 1047–1051.

Kim, S.G., Lee, S., Seo, P.J., Kim, S.K., Kim, J.K., Park, C.M., 2010. Genome-scale screening and molecular characterization of membrane-bound transcription factors in Arabidopsis and rice. Genomics 95, 56–65.

Kim, Y.S., et al., 2006. A membrane-bound NAC transcription factor regulates cell division in Arabidopsis. Plant Cell 18, 3132–3144.

Kim, S.Y., et al., 2007a. Exploring membrane-associated NAC transcription factors in Arabidopsis: implications for membrane biology in genome regulation. Nucleic Acids Res. 35, 203–213.

Kim, S.G., Kim, S.Y., Park, C.M., 2007b. A membrane-associated NAC transcription factor regulates salt-responsive flowering via FLOWERING LOCUS T in Arabidopsis. Planta 226, 647–654.

Ko, J.H., Yang, S.H., Park, A.H., Lerouxel, O., Han, K.H., 2007. ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. Plant J. 50, 1035–1048.

Lee, J., Rudd, J.J., 2002. Calcium-dependent protein kinases: versatile plant signalling components necessary for pathogen defence. Trends Plant Sci. 7, 97–98.

Li, H., Zhou, S.Y., Zhao, W.S., Su, S.C., 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.

Liang, C., Jaiswal, P., Hebbard, C., Avraham, S., Buckler, E.S., et al., 2008. Gramene: a growing plant comparative genomics resource. Nucleic Acids Res. 36, D947–D953.

Lin, H., Zhu, W., Silva, C.J., Gu, X., 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., 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., 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.Z., Baig, M.N.R., Fan, R., Ye, J.L., Cao, C.Y., Deng, X.X., 2009. Identification and expression pattern of a novel NAM, ATAF, and CUC-like gene from *Citrus sinensis* Osbeck. Plant Mol. Biol. Rep. 27, 292–297.

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

Moore, R.C., Purugganan, M.D., 2003. The early stages of duplicate gene evolution. Proc. Nat. Acad. Sci. U.S.A. 100, 15682–15687.

Mukhopadhyay, A., Vij, S., 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. Nat. Acad. Sci. U.S.A. 101, 6309–6314.

Nakashima, K., Tran, L.S., VanNguyen, D., Fujita, M., Maruyama, K., Todaka, D., Ito, Y., Hayashi, N., Shinozaki, K., Yamaguchi-Shinozaki, K., 2007. Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J. 51, 617–630.

Nuruzzaman, M., et al., 2008. Sequence and expression analysis of the thioredoxin protein gene family in rice. Mol. Genet. Genomics 280, 139–151.

Oh, S.K., Lee, S., Yu, S.H., Choi, D., 2005. Expression of a novel NAC domain-containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. Planta 222, 876–887.

Olsen, A.N., Ernst, H.A., Lo Leggio, L., Skriver, K., 2005. NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci. 10, 1360–1385.

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., et al., 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Res. 10, 239–247.

Ren, T., Qu, F., Morris, T.J., 2000. HRT gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to *turnip crinkle* virus. Plant Cell 12, 1917–1925.

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

Roy, S.W., Penny, D., 2006. Patterns of intron loss and gain in plants: intron loss-dominated evolution and genome-wide comparison of *O. sativa* and *A. thaliana*. Mol. Biol. Evol. 24, 171–181.

Rushton, P., et al., 2008. Tobacco transcription factors: novel insights into transcriptional regulation in the Solanaceae. Plant Physiol. 147, 280–295.

Ryu, H.S., et al., 2006. A comprehensive expression analysis of the WRKY gene superfamily in rice plants during defense response. Plant Cell Rep. 25, 836–847.

Sablowski, R.W., Meyerowitz, E.M., 1998. A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. Cell 92, 93–103.

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

Satoh, K., et al., 2010. Selective modification of rice (*Oryza sativa*) gene expression by rice stripe virus infection. J. Gen. Virol. 91, 294–305.

Schenk, P.M., et al., 2000. Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. Proc. Nat. Acad. Sci. U.S.A. 97, 11655–11660.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., et al., 2002. Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31, 279–292.

Seki, M., Kamei, A., Yamaguchi-Shinozaki, K., Shinozaki, K., 2003. Molecular responses to drought, salinity and frost: common and different paths for plant protection. Curr. Opin. Biotechnol. 14, 194–199.

Shibata, Y., Cabunagan, R.C., Cabauatan, P.Q., Choi, I.R., 2007. Characterization of *Oryza rufipogon*-derived resistance to tungro disease in rice. Plant Dis. 91, 1386–1391.

Shimono, M., et al., 2007. Rice WRKY45 plays a crucial role in benzothiadiazole-inducible BLAST resistance. Plant Cell 19, 2064–2076.

Souer, E., van Houwelingen, A., Kloos, D., Mol, J., Koes, R., 1996. The No Apical Meristem gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. Cell 85, 159–170.

Sperotto, R.A., et al., 2009. Identification of up-regulated genes in flag leaves during rice grain filling and characterization of *OsNAC5*, a new ABA-dependent transcription factor. Planta 230, 985–1002.

Tameling, W.I., 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., et al., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

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

Tran, L.S., et al., 2004. Isolation and functional analysis of Arabidopsis stress inducible NAC transcription factors that bind to a drought responsive *cis*-element in the early responsive to dehydration stress 1 promoter. Plant Cell 16, 2481–2498.

Tran, L.S., et al., 2009. Molecular characterization of stress-inducible *GmNAC* genes in soybean. Mol. Genet. Genomics 281, 647–664.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., Dubcovsky, J., 2006. A *NAC* gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314, 1298–1301.

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K., 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. Nat. Acad. Sci. U.S.A. 97, 11632–11637.

Xia, N., et al., 2010. Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses. Mol. Biol. Re. doi:10.1007/s11330-010-0023-4

Xie, Q., Sanz-Burgos, A.P., Guo, H., Garcia, J.A., Gutierrez, C., 1999. GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a Geminivirus protein. Plant Mol. Biol. 39, 647–656.

Xiong, Y., Liu, T., Tian, C., Sun, S., Li, J., Chen, M., 2005. Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots. Plant Mol. Biol. 59, 191–203.

Yamaguchi-Shinozaki, K., Shinozaki, K., 2001. Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Novartis Found. Symp. 236, 176–186 discussion 186–189.

Yoo, S.Y., Kim, Y., Kim, S.Y., Lee, J.S., Ahn, J.H., 2007. Control of flowering time and cold response by a NAC-domain protein in Arabidopsis. PLoS ONE 2, e642.

Yoshii, M., et al., 2009. Disruption of a novel gene for a NAC-domain protein in rice confers resistance to rice dwarf virus. Plant J. 57, 615–625.

Zheng, X., Chen, B., Lu, G., Han, B., 2009. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. Biochem. Biophys. Res. Commun. 379, 985–989.

Zhong, R., Richardson, E.A., Ye, Z.H., 2007. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell 19, 2776–2792.