Plant Mol Biol
DOI 10.1007/s11103-009-9518-6

The ankyrin repeat gene family in rice: genome-wide identification, classification and expression profiling

Jianyan Huang · Xiaobo Zhao · Huihui Yu · Yidan Ouyang · Lei Wang · Qifa Zhang

Received: 15 March 2009 / Accepted: 12 June 2009

Abstract Ankyrin repeat (ANK) containing proteins comprise a large protein family. Although many members of this family have been implicated in plant growth, development and signal transduction, only a few ANK genes have been reported in rice. In this study, we analyzed the structures, phylogenetic relationship, genome localizations and expression profiles of 175 ankyrin repeat genes identified in rice (*OsANK*). Domain composition analysis suggested OsANK proteins can be classified into ten subfamilies. Chromosomal localizations of *OsANK* genes indicated nine segmental duplication events involving 17 genes and 65 *OsANK* genes were involved in tandem duplications. The expression profiles of 158 *OsANK* genes were analyzed in 24 tissues covering the whole life cycle of two rice genotypes, Minghui 63 and Zhenshan 97. Sixteen genes showed preferential expression in given tissues compared to all the other tissues in Minghui 63 and Zhenshan 97. Nine genes were preferentially expressed in stamen of 1 day before flowering, suggesting that these genes may play important roles in pollination and fertilization. Expression data of *OsANK* genes were also obtained with tissues of seedlings subjected to three phytohormone (NAA, GA3 and KT) and light/dark treatments. Eighteen genes showed differential expression with at least one phytohormone treatment while under light/dark treatments, 13 *OsANK* genes showed differential expression. Our data provided a very useful reference for cloning and functional analysis of members of this gene family in rice.

Keywords Rice · Ankyrin repeat · *OsANK* gene family · Microarray · Expression profiles

Abbreviations

ANK	Ankyrin repeat
Os	*Oryza sativa*
At	*Arabidopsis thaliana*
ANK-M	Only contain ANK domain
TM	Transmembrane
TPR	Tetratricopeptide repeat
RF	Ring finger
BTB	Broad-complex, tramtrack and bric a brac domains
ZnF	Zinc-finger
BPA	BAR, PH and ArfGap domains
IQ	Calmodulin-binding domain
PK	S_TKc domain or Pkinase-tyr domain

Introduction

Ankyrin repeat (ANK) is one of the most common protein domain widely distributed in organisms ranging from virus to human (Sedgwick and Smerdon 1999). This domain, composed of 33 amino acids, was initially found in two yeast cell-cycle regulators Swi6/Cdc10 and the *Drosophila* signaling protein Notch (Breeden and Nasmyth 1987), and was named after the discovery of 24 copies of this sequence in the cytoskeletal protein ankyrin (Lux et al.

Electronic supplementary material The online version of this article (doi:10.1007/s11103-009-9518-6) contains supplementary material, which is available to authorized users.

J. Huang (✉) · X. Zhao · H. Yu · Y. Ouyang · L. Wang · Q. Zhang
National Key Laboratory of Crop Genetic Improvement, National Centre of Plant Gene Research (Wuhan), Huazhong Agricultural University, 430070 Wuhan, China
e-mail: hjy020@webmail.hzau.edu.cn

Published online: 16 July 2009

1990). The ANK, known to mediate protein–protein interactions (Michaely and Bennett 1992; Bork 1993), was found in numerous proteins with diverse functions (Mosavi et al. 2004). In animals and yeast, some ANK proteins played important roles in cell-cycle control, transcriptional regulation, cytoskeleton integrity and signal transduction (Sedgwick and Smerdon 1999).

In plants, ANK proteins have been demonstrated to be involved in a number of important processes. AKR was the first reported ANK protein in higher plant *Arabidopsis* (Zhang et al. 1992). It was regulated by light and played a regulatory role in cell differentiation and development. Overexpressing antisense or sense RNA of the *AKR* gene exhibited a chlorotic phenotype. *AKR2A* was found to play a role in the biogenesis of the chloroplast outer envelope membrane (OEM) proteins (Bae et al. 2008). *BOP1*, an ANK gene from *Arabidopsis* was required for leaf morphogenesis (Ha et al. 2004). EMB506 containing five ANK repeats organized in tandem within the C terminal moiety was essential for embryogenesis in *Arabidopsis* (Albert et al. 1999). Further study revealed that EMB506 interacted with AKRP encoded by *AKR* through their ANK domains. Both EMB506 and AKRP were essential for the plastid differentiation linked to cell differentiation, morphogenesis and organogenesis during the plant life cycle and development (Garcion et al. 2006). *TIP1*, also an ANK gene in *Arabidopsis* encoding an *S*-acyl transferase, affected cell growth throughout the whole life of plant (Hemsley et al. 2005). It was shown that overexpression of *TIP1* led to longer root hairs suggesting it functions in root hair formation. Another ANK gene *XBAT32* also affected the lateral root initiation and its expression was induced by auxin (Nodzon et al. 2004). *LIANK*, an ANK gene in Lily, was essential for pollen germination and pollen tube growth. Overexpression of *LIANK* caused abnormal pollen tube growth while knockout of this gene decreased the growth of the polarized tip of the pollen tube (Huang et al. 2006).

Ankyrin repeat proteins were also found to play important roles in responses to biotic and abiotic stresses in plant. The *Arabidopsis* ANK protein AKR2 might be involved in the regulation of antioxidation metabolism in both disease resistance and stress responses (Yan et al. 2002). In pepper, *CaKR1* was found to play roles in both biotic and abiotic stress responses (Seong et al. 2007b). Transgenic tomato plants expressing *CaKR1* showed enhanced resistance to *Phytophthora infestans* (Seong et al. 2007a). ACD6, acting as a plasma membrane localized positive regulator of salicylic acid signaling, controlled defense responses against virulent bacteria (Lu et al. 2003, 2005). *ZFAR1* was a gene encoding a putative zinc-finger protein with ANK domains, and the *ZFARI* mutants showed increased local susceptibility to *Botrytis* and sensitivity to germination in the presence of abscisic acid (AbuQamar et al. 2006). *ITN1* encoding an ANK-transmembrane protein was implicated in diverse cellular processes such as signal transduction. The *itn1* mutation partially impaired ABA signaling pathways (Sakamoto et al. 2008). NPR1, a positive regulator of acquired resistance responses, was a central activator of SA-regulated gene expression (Cao et al. 1997).

Ankyrin repeat proteins were also found to perform other functions. Examples included *ACBP2* that may play a role in mediation of AtEBP movement between cells (Li and Chye 2004), *MjXB3*, which was involved in the coordination of the senescence program (Xu et al. 2007), *AKT3*, which was a functional transport protein (Ketchum and Slayman 1996), and *IGN1*, which was shown to be required for the maintenance of nitrogen-fixing symbiosis in root nodules (Kumagai et al. 2007).

To our knowledge, only two ANK proteins have been reported in rice. One was XB3 containing an ANK domain interacting with the kinase domain of XA21, a receptor-like kinase protein in rice having a major role in resistance against *Xanthomonas oryzae pv oryzae* (Wang et al. 2006). The other was OsCBT identified as a transcriptional activator modulated by CaM binding (Choi et al. 2005).

Rice is the main staple food for a large segment of the world population and is an ideal model plant to analyze gene expression and function (Zhang 2007; Zhang et al. 2008). The finished high quality sequences of the rice genome (International Rice Genome Sequencing Project, 2005) and data generated from high-throughput expression analysis provided an excellent opportunity for genome-wide analysis of all the genes belonging to a specified gene family. In the study reported in this paper, we identified 175 *OsANK* genes in rice by database searches, which genes were classified by protein domains. We analyzed the phylogenetic relationship of the ANK genes in rice and *Arabidopsis* as well as segmental and tandem duplications of *OsANK* genes. We also surveyed the expression patterns of *OsANK* genes in the whole life developmental stages of rice and their responses to three representative phytohormone (NAA, KT and GA3) and light/dark treatments. The data generated will be very helpful for studies on the biological functions of each *OsANK* gene.

Materials and methods

Collection and classification of *OsANK* members

First of all, we download the Hidden Markov Model (HMM) profile of ANK from Pfam (http://pfam.sanger.ac.uk/). The consensus protein sequence of ANK was generated by hmmemit (Eddy 1998) (http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=hmmemit). Then BLAST search tools BLASTP and TBLASTN (Altschul

et al. 1997) were used to identify putative *OsANK* with the ANK consensus protein sequence as a query against three databases: TIGR rice genome annotation (http://rice.plantbiology.msu.edu/), National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) and the Knowledge-Based Oryza Molecular Biological Encyclopedia (KOME) (http://www.cdna01.dna.affrc.go.jp/cDNA). The BLASTP and TBLASTN search parameters of the three databases were set as follows: max target sequences-500 and expect value less than ten. In addition, aimed at a more complete collection of putative ANK genes in rice, keywords (ankyrin, ankyrin repeat) and domain (PF00023) search against TIGR were also performed. The SMART (http://smart.embl-heidelberg.de/) and Pfam searches were used to confirm and make classification of each predicted OsANK gene or protein. Based on additional conserved motifs or domains besides ANK, we classified the OsANK proteins into subfamilies and the sample protein structures of each subfamily were drawn manually. Information about the gene structures, transcripts, chromosomal localization, full-length cDNA, BAC accessions for each gene and characteristics of corresponding proteins were procured from TIGR, KOME and GRAMENE (Liang et al. 2008).

Phylogenetic analysis and sequence alignment

By using OsANK protein sequences, an unrooted tree was generated by ClustalX version1.83 (Thompson et al. 1997) with neighbor-joining method (Saitou and Nei 1987) and bootstrap analysis (1,000 replicates). The tree was analyzed and displayed using MEGA software version 4 (Tamura et al. 2007). Another unrooted tree was constructed using the same method with the alignment of OsANK and AtANK protein sequences. We defined two proteins with 100% support as homologous proteins in the same species while those from different species with 100% support as orthologous proteins. Multiple sequence alignments were performed with Alignx in Vector NTI 9.0 (Lu and Moriyama 2004) and were confirmed by ClustalX version1.83 (Thompson et al. 1997).

Localization on chromosomes and duplications

OsANK genes were placed on rice chromosomes according to their positions given in the TIGR rice database. The distribution of *OsANK* genes on the rice chromosomes was drawn by MapInspect (http://www.plantbreeding.wur.nl/UK/software_mapinspect.html) and modified manually with annotation. For detection of large segmental duplications, we referred to the segmental genome duplication of rice on TIGR with a maximum length distance permitted between collinear gene pairs of 100 kb as well as 500 kb (http://rice.plantbiology.msu.edu/segmental_dup/index.shtml). We designated tandem duplicated genes if two *OsANK* genes were separated by five or fewer gene loci according to Rice Genome Annotation Release 6 of TIGR. The software MegAlign 4.03 (Clewley and Arnold 1997) was used to analyze the homology of duplicated *OsANK* genes.

Phytohormone and light/dark treatments

For phytohormone treatments, 7-day-old light-grown rice seedlings (trefoil stage) of two elite hybrid rice parents Minghui 63 and Zhenshan 97 were both transferred to solutions of 0.1 mM NAA (a member of the auxin family), GA3 (first isolated and identificated GA) and KT (a cytokinin) respectively. Samples were harvested at the time points of 5, 15, 30 and 60 min after treatments and the samples under the same phytohormone treatment of different time points were mixed together.

For the light/dark treatments, seedlings of Minghui 63 and Zhenshan 97 at the plumule and radicle stages were placed under 48 h continuous light or darkness and harvested at the indicated time respectively.

Expression profile analysis

Expression profile data of *OsANK* genes in Minghui 63 and Zhenshan 97 were extracted from database CREP (http://crep.ncpgr.cn/) which composed of hybridization of RNA samples from 39 tissues that covered the whole life cycle of rice with the Affymetrix rice microarray. Thirteen vegetative and 11 reproductive tissues of different developmental stages covering the whole life cycle of rice were used for *OsANK* expression profile analysis in this study: (1) germinating seed at 72 h of imbibitions (Seed); (2) seedlings of 3 days after sowing (Seedling 1); (3) seedlings at trefoil stage (Seedling 2); (4) shoots of seedlings with two tillers (Shoot); (5) roots of seedlings with two tillers (Root); (6) leaves at secondary branch primordial stage of young panicle development (Leaf 1); (7) leaves at meiosis stage of young panicle development (Leaf 2) (8) flag leaves at 5 days before heading (Flag leaf 1); (9) flag leaves at 14 days after flowering (Flag leaf 2); (10) sheath at secondary branch primordial stage of young panicle development (Sheath 1); (11) sheath at meiosis stage of young panicle development (Sheath 2); (12) stem at 5 days before heading (Stem 1); (13) stem at heading stage (Stem 2); (14) panicle at secondary branch primordial stage (Panicle 1); (15) panicle at pistil/stamen primordial differentiation stage (Panicle 2); (16) panicle at pollen-mother cell formation stage (Panicle 3); (17) panicle at meiosis stage of young panicle development (Panicle 4); (18) panicle at heading stage (Panicle 5); (19) hull at 1 day before flowering (Hull); (20) stamen at 1 day before flowering (Stamen); (21) spikelet at 3 days after flowering (Spikelet); (22) endosperm at 7 days after pollination (Endosperm 1); (23) endosperm at 14 days after pollination (Endosperm 2);

(24) endosperm at 21 days after pollination (Endosperm 3). In addition, expression profiles of *OsANK* genes under three different phytohormone treatments (NAA, KT and GA3) of seedlings at trefoil stage as well as light/dark treatments of the seedlings (plumule and radicle tissues) were also analyzed. The detailed information of samples used in microarray analysis was listed in Supplemental Table 1.

After normalization, variance stabilization, the average signal value of two biological replicates for each tissue was used for analysis, except for five tissues (2, 3, 14, 15 and 16) which had three biological replicates and two technical replicates. Wherever more than one probe set was available for one gene, the average signal value of these probe sets belonging to the same gene was used for analysis. To identify preferentially expressed genes, a student-*t* test was performed. A gene in a given tissue was defined as preferentially expressed only if the expression value of the gene in this tissue was more than two-fold and had a *P* value less than 0.05 compared to all other 23 tissues. Under phytohormone and light/dark treatments, genes that were up- or down-regulated more than two-fold and with *P* value less than 0.05 compared to control were considered as differentially expressed.

Real-time PCR analysis

Real-time PCR reactions were carried out using the same RNA samples which were used for microarrays. Gene-specific primers were designed using PRIMER EXPRESS version 2.0 (Applied Biosystems) with default parameters. For real-time PCR analysis, the first-strand cDNA was synthesized from total RNA using Superscript III reverse transcriptase (Invitrogen). Real-time PCR was carried out using ABI PRISM 7500 Real-Time PCR system (Applied Biosystems). Each reaction of a volume of 25 μl contained 2.0 μl transcription product, 0.5 μl primers, 12.5 μl 2 × SYBR Premix® Ex Taq™ and 0.5 μl 50 × ROX reference dye II (TaKaRa). The thermal cycle was set as: 95°C for 10 s; 45 cycles of 95°C for 5 s, 60°C for 34 s. Rice Actin1 gene (Accession number X16280) was used as internal control. The relative expression levels were analyzed as described previously (Livak and Schmittgen 2001).

Results

Identification of ANK genes in rice

To identify ANK genes in the rice genome, three public databases: TIGR, NCBI and KOME were used. The consensus protein sequence generated by hmmemit from ANK HMM profile (PF00023) is DGFTPLHLAALRGNLEV VKLLLSQGADLNAQDD. Using this sequence as a query, searching by BLASTP and TBLASTN against rice (*Oryza sativa japonica* subsp. cv Nipponbare) genome was proceeded. By removal of the same sequences from the three databases and different transcripts of the same gene, we identified 172 putative *OsANK* genes including two genes (AK110327 and AK119826) only found in KOME. Keywords and domain searches against TIGR were also performed, resulting in one and eight new members respectively. All the protein sequences of the 181 putative *OsANK* genes were confirmed by SMART and Pfam searching for the presence of ANK domain. Six of them (Os01g28220, Os05g39750, Os07g09160, Os08g28220, Os11g43690 and Os12g13170) had no ANK domain and were excluded in further analysis. Therefore, there were at least 175 *OsANK* genes in the rice genome, of which 173 had corresponding locus IDs in TIGR database. The number of ANK proteins in rice (175 members) is greater than that in *Arabidopsis* (105 members) (Becerra et al. 2004). The detailed information and the gene structure of representative *OsANK* genes could be found in Supplemental Table 2 and Supplemental Fig. 1. For convenience, all the "LOC_" prefix of TIGR locus IDs were omitted in the rest of this paper.

Classification of OsANK proteins

According to the detailed results of SMART and Pfam searches, the 175 OsANK proteins were classified into ten subfamilies based on their domain compositions (Fig. 1). Seventy-three members merely with ANK domain belonged to subfamily ANK-M. Besides ANK domain, OsANK proteins containing several other known functional domains were classified into the following subfamilies. Thirty-seven members containing the transmembrane domain were identified as subfamily ANK-TM; 22 members containing the tetratricopeptide repeat domain were identified as subfamily ANK-TPR; 9 members with the Ring Finger domain were identified as ANK-RF subfamily; ANK-BTB subfamily (6 members) had broad-complex, tramtrack and bric a brac domains; ANK-ZnF subfamily (7 members) contained zinc-finger domain; ANK-BPA subfamily (3 members) had BAR, PH and ArfGap domains; ANK-IQ subfamily (4 members) contained the Calmodulin-binding domain; ANK-PK subfamily (4 members) had S_TKc domain or Pkinase-tyr domain; ANK-O subfamily (10 members) contained other domains including CHROMO, ACBP, RCC1, AAA, PPR, Motile-sperm, HDAC, rve, RVT, CNMP, Ion-trans and RRM (Supplemental Fig. 2). The differences and comparison of classification of *OsANK* and ANK gene family in *Arabidopsis* can be seen in Table 1. The numbers of members in ANK-TM, ANK-ZnF, ANK-BPA, ANK-IQ, ANK-BTB subfamilies are not very

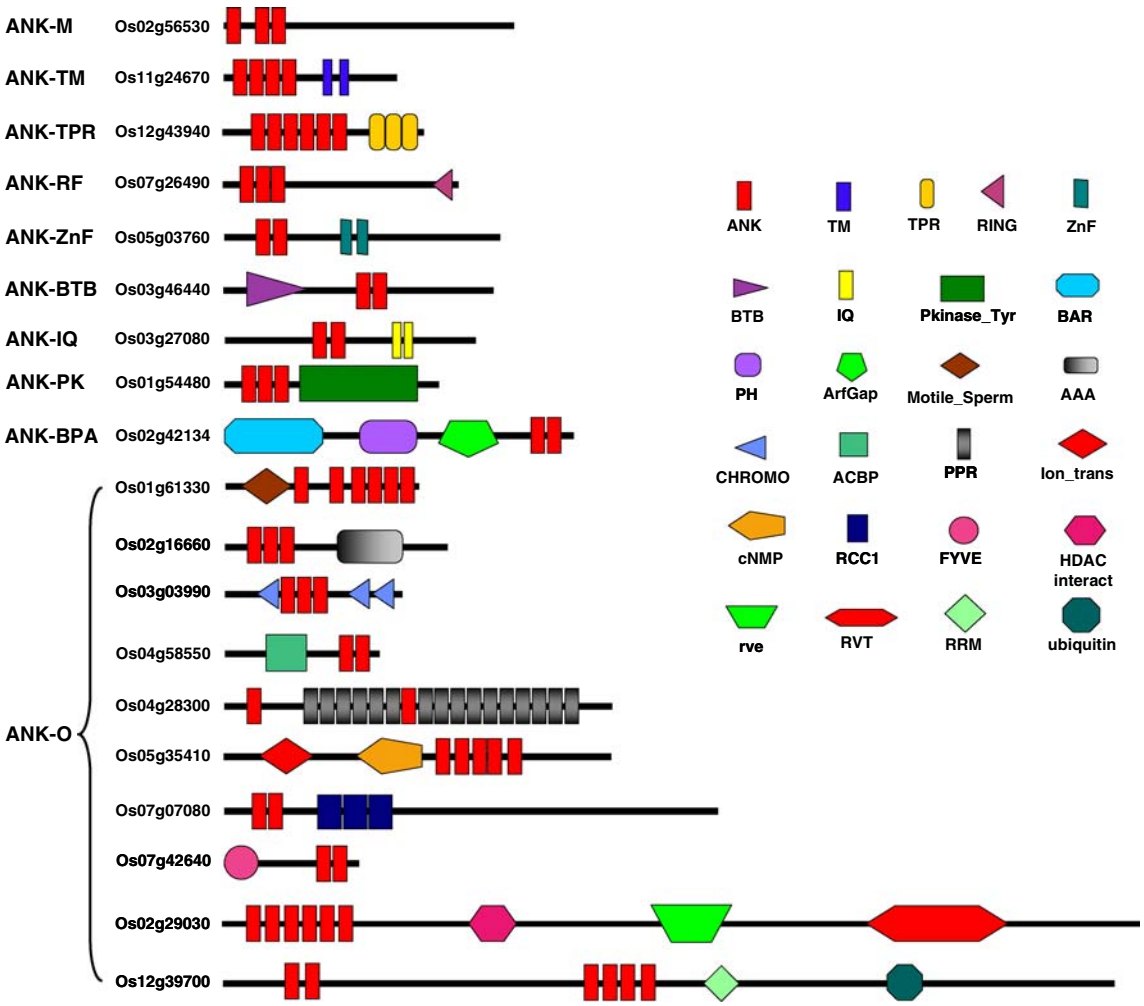


Fig. 1 Structure of representative OsANK protein from each subfamily. Subfamily name of each corresponding protein belonged to and TIGR locus ID are given on the *left*. Domain abbreviations are: *ANK* ankyrin repeat domain; *TM* transmembrane; *TPR* tetratricopeptide repeat domain; *RING* ring finger domain; *ZnF* zinc finger; *BTB* broad-complex, tramtrack and bric a brac domains; *IQ* calmodulin-binding domain; *BAR* BAR domain; *PH* pleckstrin homology domain; *ArfGap* putative GTP-ase activating proteins for the small GTPase, ARF; *Motile_Sperm* MSP (major sperm protein) domain; *AAA* ATPase family associated with various cellular activities; *CHROMO* CHRomatin organisation MOdifier domain; *ACBP* acyl CoA binding protein; *PPR* pentatricopeptide repeats; *Ion_trans* ion transport protein; *cNMP* cyclic nucleotide-binding domain; *RCC1* regulator of chromosome condensation; *FYVE* FYVE zinc finger; *rve* integrase core domain; *RVT* reverse transcriptase (RNA-dependent DNA polymerase); *RRM* RNA recognition motif; *ubiquitin* ubiquitin family. The length and order of domains represent actual situation in each protein

different between the two species. However, the number of protein kinases in OsANK is slightly smaller than that in *Arabidopsis*. The numbers of ANK-M and ANK-TPR in rice are larger than that in *Arabidopsis*, likely because those proteins having only one ANK or more than two ANK domains but separated by more than 20 amino acids were excluded in the previous study of ANK gene family in *Arabidopsis* (Becerra et al. 2004).

Conserved motifs investigation in ANK-M subfamily proteins

Almost half of the OsANK proteins (41.71%) had ANK as the only recognizable domain. In order to find more conserved motifs or domains among the 73 members of ANK-M subfamily, the motif investigation software MEME (Multiple EM for Motif Elicitation) version 4.0.0 (Bailey et al. 2006) (http://meme.sdsc.edu/meme4/cgi-bin/meme.cgi) was employed. The parameters of this analysis were set up as below: number of repetitions—any, maximum number of motifs—20, optimum motif width set to $\geq$6 and $\leq$200. The *E* values of 20 putative motifs identified were less than 1.00E-30 (Fig. 2). The motifs identified by MEME were between 15 and 119 amino acids in length and exist in at least two of predicted ANK-M proteins. Motif 8 existed in as many as 60 ANK-M proteins with *E* value less than 1.00E-100. Motif 17 composed of 119 amino acids was only found in two ANK-M proteins.

These putative motifs obtained from MEME were annotated by searching SMART and InterProScan (Hunter et al. 2009), which showed that most motifs had unknown functions except motif 1, 2, 3, 8, 9, 11 and 17. All the motifs with known functions were annotated as ANK. The log likelihood ratio, *E* value, consensus amino acid sequence and length of each motif are given in Supplemental Table 3. The phylogenetic analysis of *OsANK* gene family (below) showed that those ANK-M genes had very distant evolutionary relationships. Therefore, we inferred that maybe these motifs identified by MEME were important for the function of the proteins although most of their functions were unknown.

Phylogenetic analysis and multiple sequence alignments

To clarify the phylogenetic relationship among the *OsANK* genes and infer the evolutionary history of this gene family, a combined phylogenetic tree was constructed with the aligned OsANK protein sequences (Fig. 3), from which it can be seen that the OsANK proteins fall into eight major groups (group A to group H). Group B, E, F and G were further divided into 4, 3, 2 and 2 subgroups, respectively. In addition, we observed that most members in the same groups or subgroups shared one or more domains outside the ANK domain, further supporting the subfamily definition above. For instance, all the members of ANK-TPR subfamily belonged to group A. Seven out of nine of ANK-RF members were assigned to subgroup B1. Similarly, subgroup B2 contained all the seven members of ANK-RF. Group C consisted of 21 members of ANK-M and 20 members of ANK-TM. All the members of ANK-BTB, ANK-BPA, and three of four members of ANK-IQ subfamily were classified into group E and all the ANK-PK proteins were assigned to subgroup G1. Group F1 consisted of 12 proteins while 10 of them belonged to ANK-M.

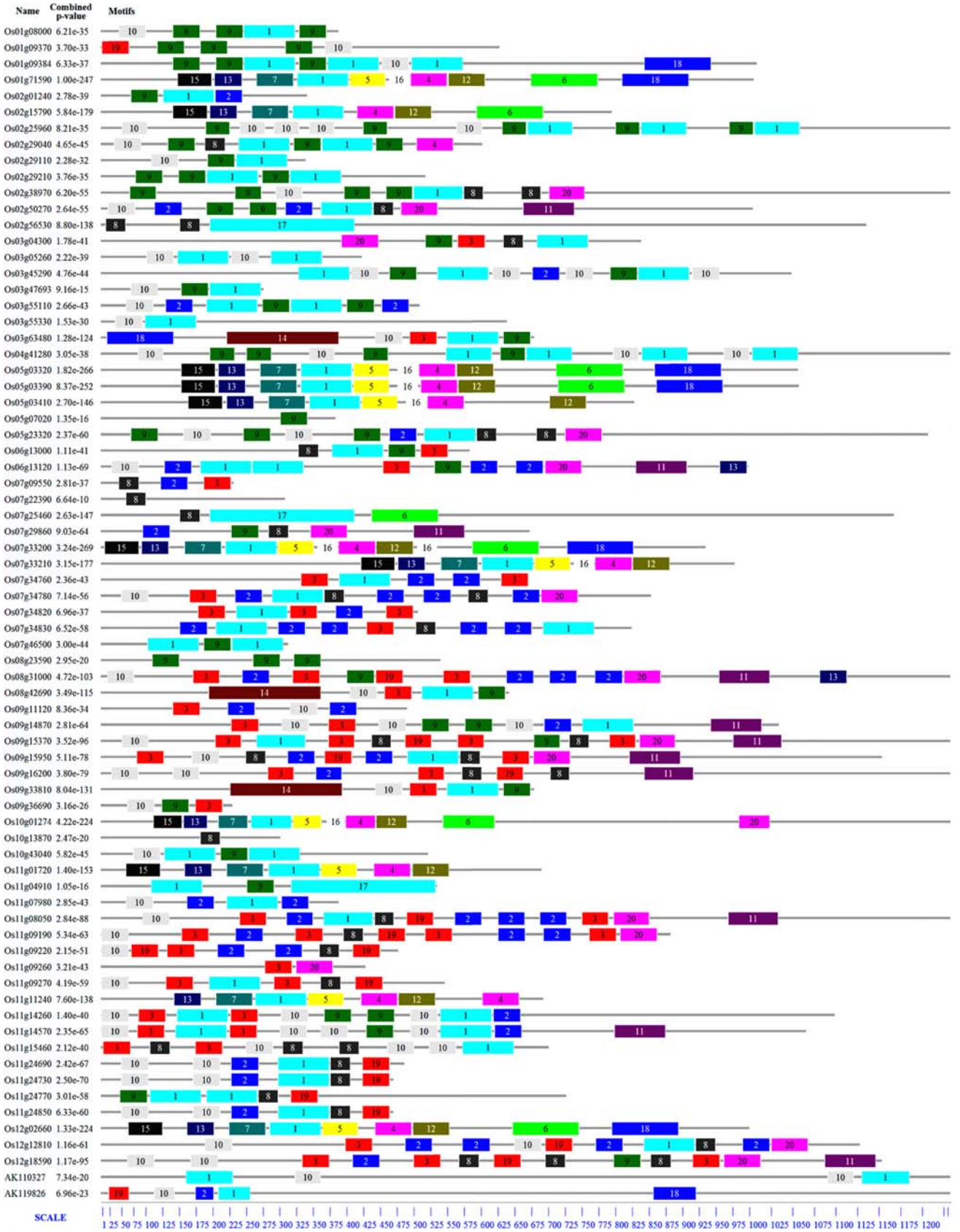
Fig. 2 Distribution of conserved motifs in the members of ANK-M subfamily identified by MEME. The name of each member and combined *P* value are shown on the *left side* of the figure and length of each motif is indicated at the *bottom* of the figure. Different motifs are indicated with different *color boxes* numbered 1–20. The same number in different proteins refers to the same motif. The length and order of motifs represent the actual situation in each protein. The detailed information of the motifs is given in Supplemental Table 3

However, the members of ANK-M were distributed in almost all the groups. It reflected that the functions of ANK-M subfamily genes were diversified. In general, most of the closely related members in the phylogenetic tree had the same or very similar domain composition. However, the classification based on domain composition did not completely match the phylogenetic classification.

With the development of comparative genomics, it is possible to analyze proteins of the same gene family among different species. In order to evaluate the phylogenetic relationship among the ANK proteins in rice and *Arabidopsis*, another unrooted tree was constructed from alignments of all the 280 ANK protein sequences of the two species using the same method mentioned above (Supplemental Fig. 3). The AtANK protein sequences were obtained from The *Arabidopsis* Information Resources (TAIR) (http://www.arabidopsis.org/) according to the IDs provided in the previous article (Becerra et al. 2004). This phylogenetic analysis suggested that most OsANK and AtANK proteins clustered in specie-specific clades with very high bootstrap supporting, except for 17 pairs of orthologous proteins, i.e. Os05g01310 and AT3G04710. We also identified 17 pairs of paralogous proteins in *Arabidopsis* and 20 pairs in rice, such as AT3G23280 and AT4G14365, Os02g29040 and Os02g29130. This result indicated that only a few members of *OsANK* and *AtANK* gene families possibly originated from the same ancestral genes before divergence of dicots and monocots.

Table 1 Numbers of each subfamily of ANK proteins in rice and *Arabidopsis*

Subfamily	Description	Rice	*Arabidopsis*
ANK-M	Proteins with only ankyrin repeats	73	18
ANK-TM	Ankyrin-transmembrane proteins	37	40
ANK-TPR	Protein with tetratricopeptide repeats	22	1
ANK-O	Proteins with other domains	10	13
ANK-RF	Ring finger proteins	9	5
ANK-ZnF	Zinc-finger proteins	7	6
ANK-BTB	Proteins with BTB domain	6	7
ANK-IQ	Calmodulin binding motif-containing protein	4	4
ANK-PK	Protein kinases	4	7
ANK-BPA	ARF GTPase-activating domain-containing protein	3	4
Total		175	105


Name
Combined p-value
Motifs
Os01g08000 6.21e-35
Os01g09370 3.70e-33
Os01g09384 6.33e-37
Os01g71590 1.00e-247
Os02g01240 2.78e-39
Os02g15790 5.84e-179
Os02g25960 8.21e-35
Os02g29040 4.65e-45
Os02g29110 2.28e-32
Os02g29210 3.76e-35
Os02g38970 6.20e-55
Os02g50270 2.64e-55
Os02g56530 8.80e-138
Os03g04300 1.78e-41
Os03g05260 2.22e-39
Os03g45290 4.76e-44
Os03g47693 9.16e-15
Os03g55110 2.66e-43
Os03g55330 1.53e-30
Os03g63480 1.28e-124
Os04g41280 3.05e-38
Os05g03320 1.82e-266
Os05g03390 8.37e-252
Os05g03410 2.70e-146
Os05g07020 1.35e-16
Os05g23320 2.37e-60
Os06g13000 1.11e-41
Os06g13120 1.13e-69
Os07g09550 2.81e-37
Os07g22390 6.64e-10
Os07g25460 2.63e-147
Os07g29860 9.03e-64
Os07g33200 3.24e-269
Os07g33210 3.15e-177
Os07g34760 2.36e-43
Os07g34780 7.14e-56
Os07g34820 6.96e-37
Os07g34830 6.52e-58
Os07g46500 3.00e-44
Os08g23590 2.95e-20
Os08g31000 4.72e-103
Os08g42690 3.49e-115
Os09g11120 8.36e-34
Os09g14870 2.81e-64
Os09g15370 3.52e-96
Os09g15950 5.11e-78
Os09g16200 3.80e-79
Os09g33810 8.04e-131
Os09g36690 3.16e-26
Os10g01274 4.22e-224
Os10g13870 2.47e-20
Os10g43040 5.82e-45
Os11g01720 1.40e-153
Os11g04910 1.05e-16
Os11g07980 2.85e-43
Os11g08050 2.84e-88
Os11g09190 5.34e-63
Os11g09220 2.15e-51
Os11g09260 3.21e-43
Os11g09270 4.19e-59
Os11g11240 7.60e-138
Os11g14260 1.40e-40
Os11g14570 2.35e-65
Os11g15460 2.12e-40
Os11g24690 2.42e-67
Os11g24730 2.50e-70
Os11g24770 3.01e-58
Os11g24850 6.33e-60
Os12g02660 1.33e-224
Os12g12810 1.16e-61
Os12g18590 1.17e-95
AK110327 7.34e-20
AK119826 6.96e-23
SCALE
1 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200

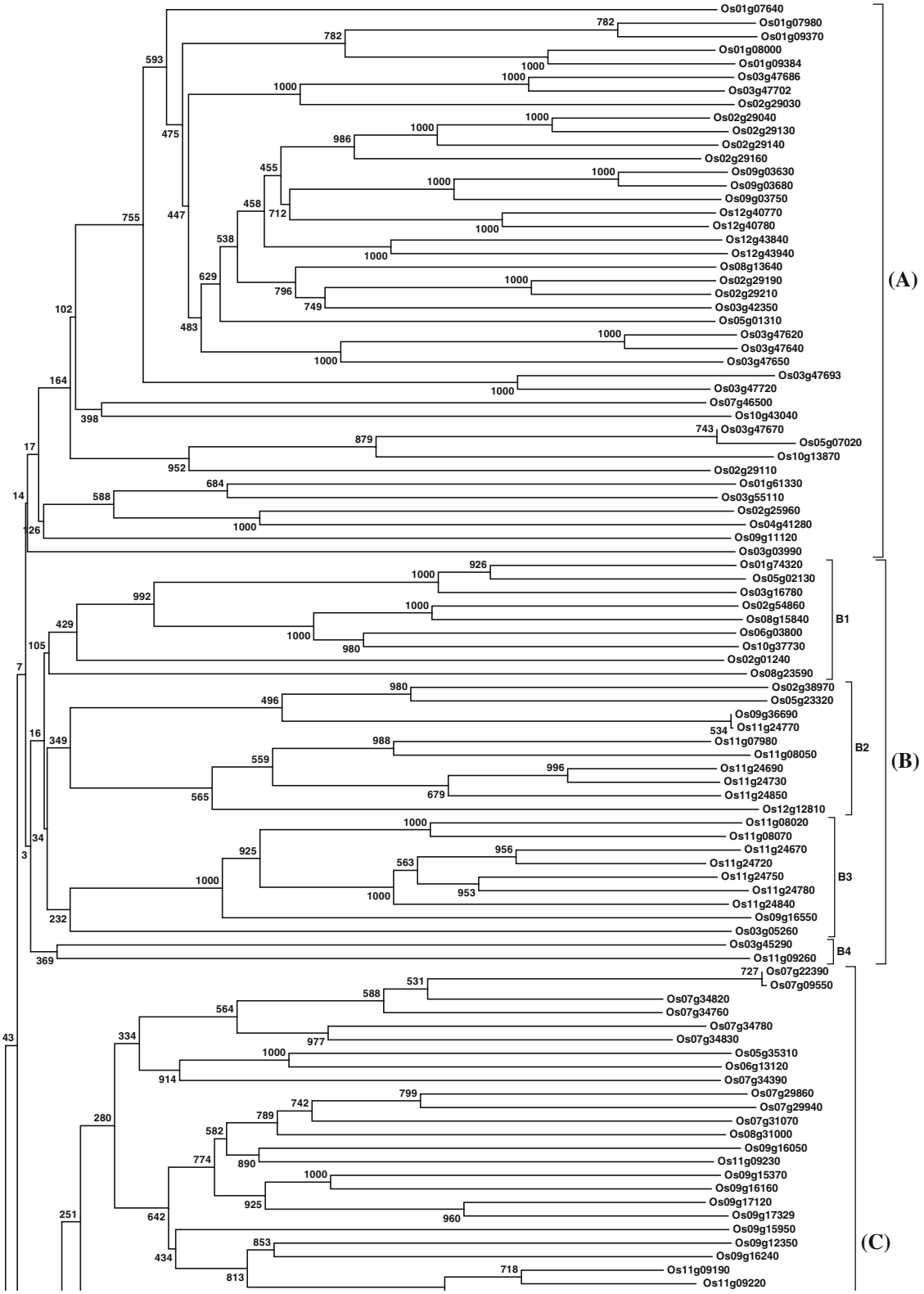


Fig. 3 Evolutionary relationship among the rice OsANK proteins. The unrooted tree was generated using ClustalX program by neighbor-joining method. Bootstrap values from 1,000 replicates are indicated at each node. OsANK proteins are divided into eight distinct groups (A–H). Group B, E, F and G are further divided into 11 subgroups (B1–B4, E1–E3, F1–F2, and G1–G2)

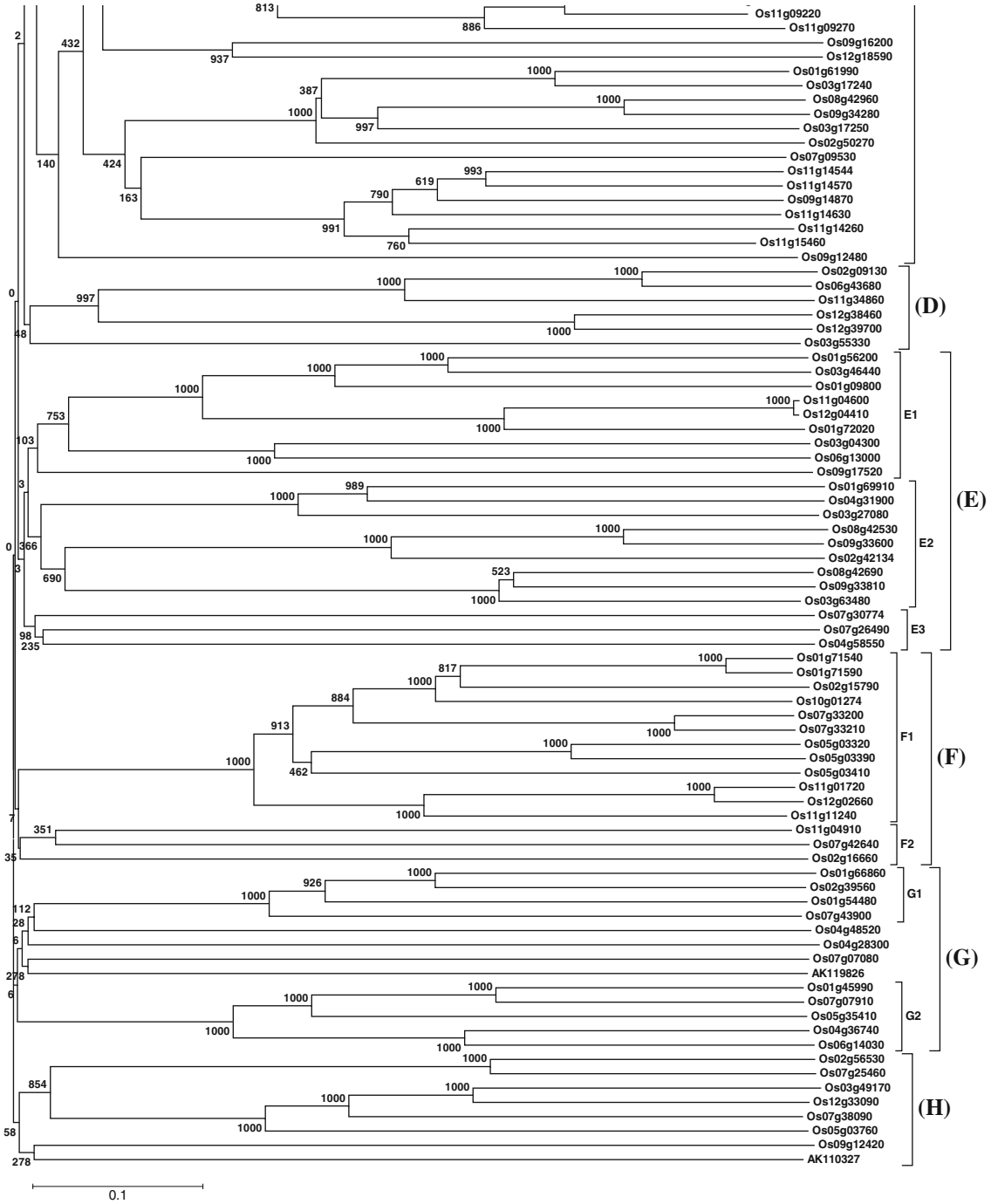


Fig. 3 continued

To gain understanding of the connection between the domain composition and phylogenetic relations of the members in the same subfamily, multiple sequence alignments of each subfamily were proceeded. As an example, the alignment result of ANK-TPR subfamily was list in Supplemental Fig. 4. The amino acid residues corresponding to the ANK and TPR domain could be easily identified from the alignments. As all the members of ANK-TPR subfamily were in the same group of the phylogenetic tree, the result further supported the relationship analysis between subfamily and phylogenetic group above.

Chromosomal locations of the *OsANK* genes

To determine the genomic distribution of *OsANK* genes, we used the DNA sequences of *OsANK* genes to search

against the rice genome database with BLASTN at TIGR. The position of each gene could be found in Supplemental Table 2. Totally, 173 of the 175 *OsANK* genes could be localized on the 12 chromosomes with obviously uneven distribution. Within a single chromosome, *OsANK* genes could be found in all regions: at the telomeric ends, near the centromere, dispersed all over, individually or in clusters (Fig. 4). Chromosome 11 had the largest number of 28 *OsANK* genes followed by 22 on chromosome 3. In contrast, only four *OsANK* genes were found on chromosome 10 and six *OsANK* genes were on the long arm of chromosome 4. Seven out of 10 were located in the short arm of chromosome 5 and seven genes were located in the long arm of chromosome 12. Less than 10 *OsANK* genes were found on chromosome 4, 6, 8 and 10. Two *OsANK* genes (Os07g22390 and Os10g13870) encoding proteins having only the ANK domain were positioned around the centromere, which were considered to contain actively transcribed genes (Cooke 2004; Nagaki et al. 2004). Additionally, six *OsANK* genes, Os01g74320, Os02g01240, Os05g01310, Os10g01274, Os10g43040 and Os12g43940, were located near the telomeric regions.

During the evolution of plant gene family, segmental duplication and tandem duplication play a part in retaining the large number of gene family (Cannon et al. 2004). With the purpose of elucidating the potential mechanism of evolution of *OsANK* gene family, both segmental and tandem duplication events were analyzed. We found nine segmental duplication events (Fig. 4). Although the ANK-M subfamily contained the maximum members of *OsANK*

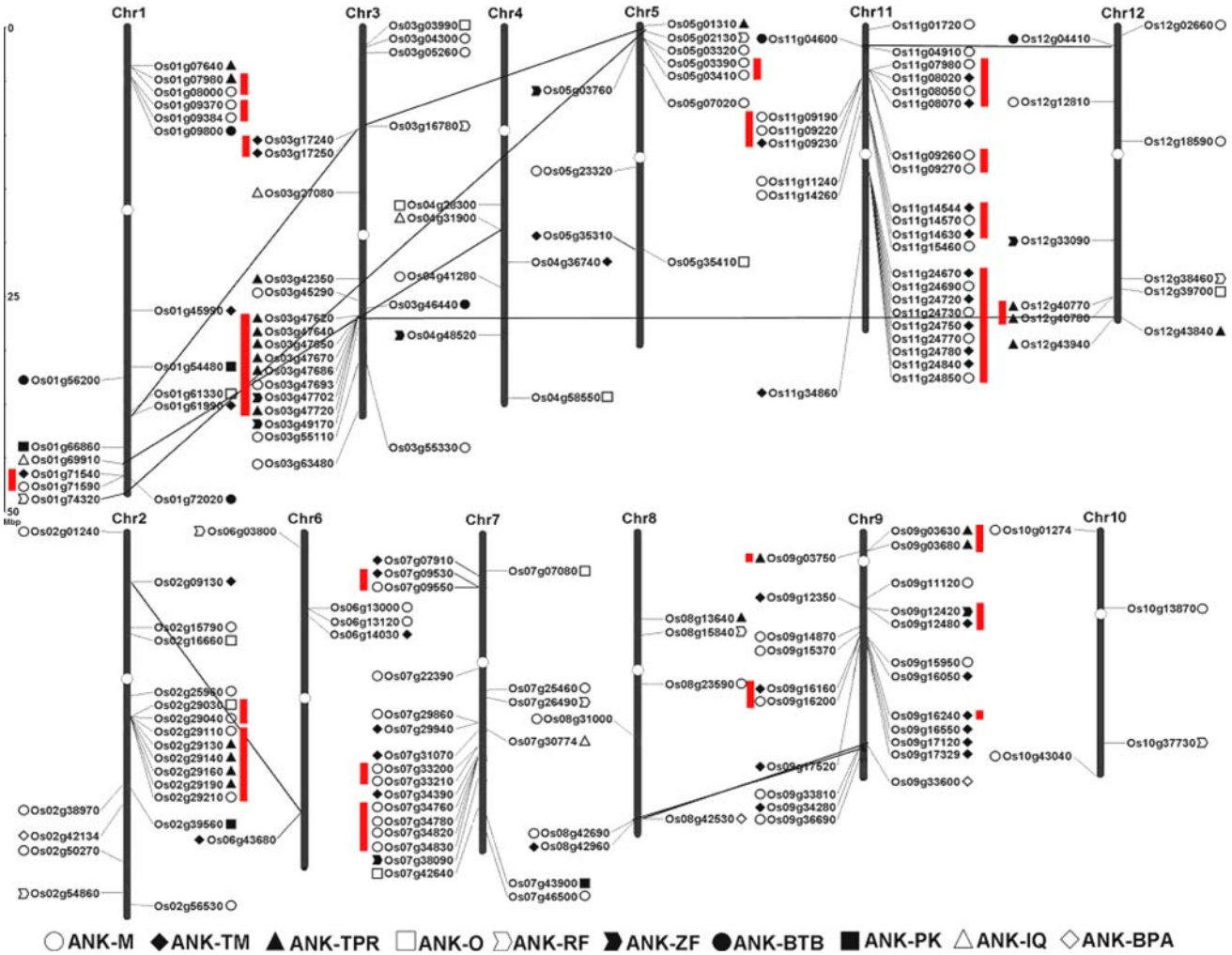


Fig. 4 Distribution of *OsANK* genes on the rice chromosomes. The scale is in megabases (*Mb*). The chromosome numbers are indicated at the *top* of each *bar*. The *white circle* on each chromosome (*vertical bars*) shows the rough position of the centromere. The markers in front of the *OsANK* genes indicate the subfamily of each protein belonged. The ten markers represent the subfamily are mentioned at the *bottom* of the diagram. The genes with open reading frames in opposite orientations are marked on different sides of the chromosome, i.e. on the *left side* represent downward while on the *right* indicate upward. *Straight lines* connect the *OsANK* genes presented on duplicated chromosomal segments and tandem duplicated gene clusters are marked by the *red rectangle*

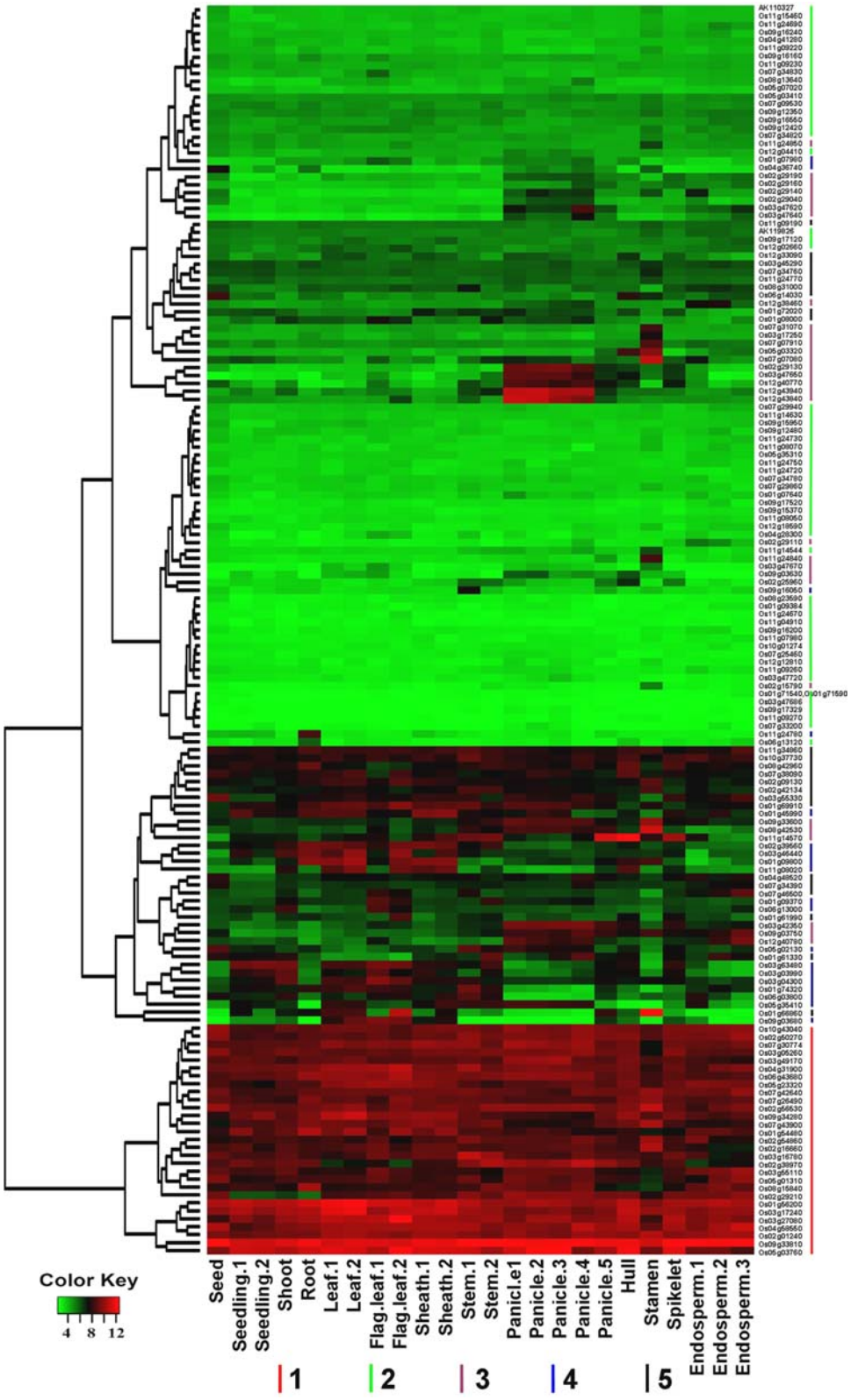


Fig. 5 Expression analysis of *OsANK* genes in the whole life cycle of Minghui 63. Hierarchical cluster displaying expression profiles for 158 *OsANK* genes based on average log2 signal values in 24 tissues. The tissues are indicated at the *bottom*. The *color key* representing log2 signal values is shown at the *bottom*. On the *left side* of expression map, the cluster dendrogram is shown. On the *right side*, five groups have been made for the genes showing discrete expression patterns. Different *color lines* representing different groups are mentioned at the *bottom* of the diagram

genes, only one pair of *OsANK* genes belonging to this subfamily were on duplicated chromosomal segments on chromosome 8 and 9 (Os08g35930 and Os09g27090). Similarly, the second large subfamily ANK-TM held three pairs of segmental duplicated genes. Os05g02130 participated in two duplication events with Os01g74320 and

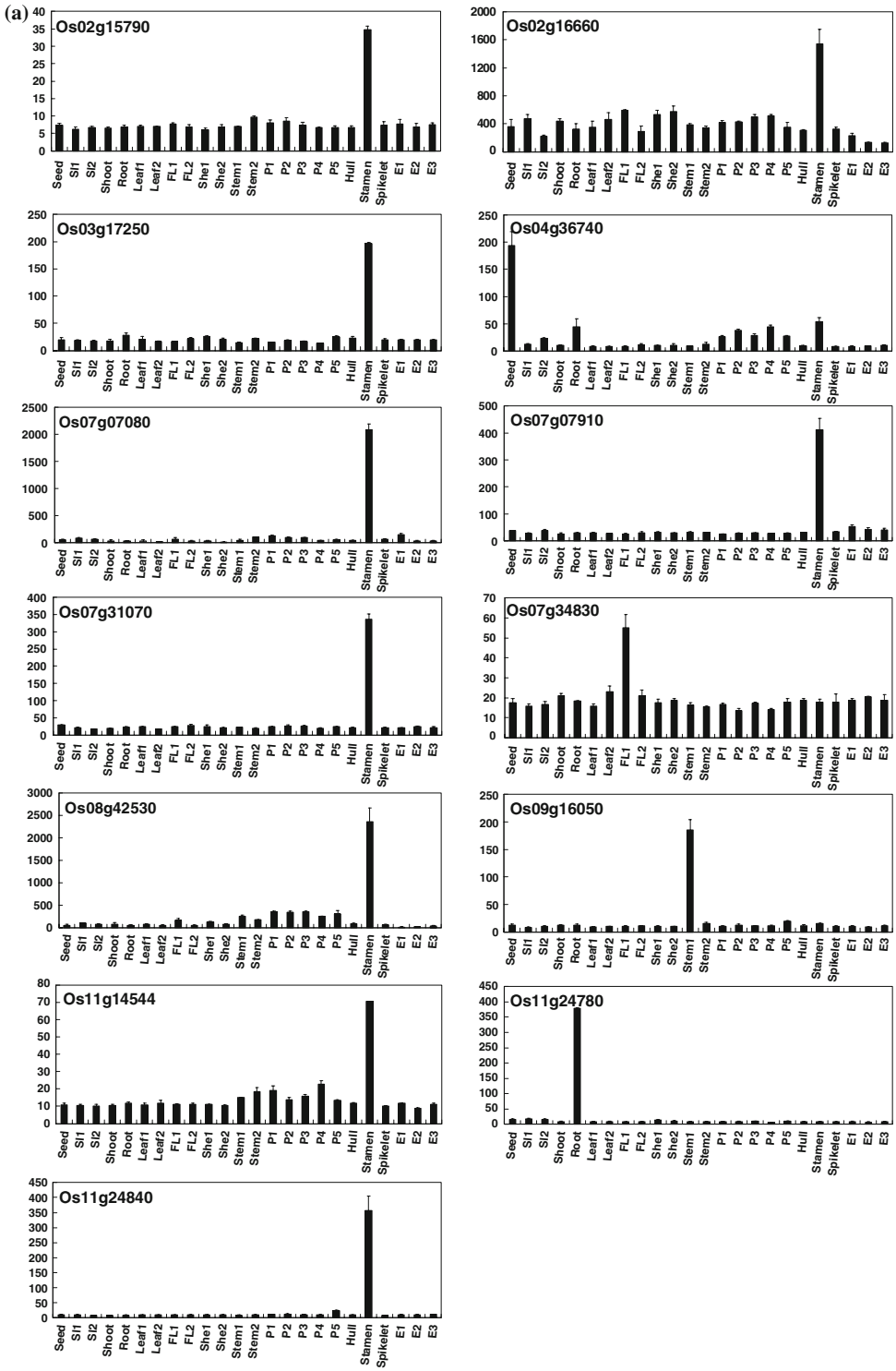
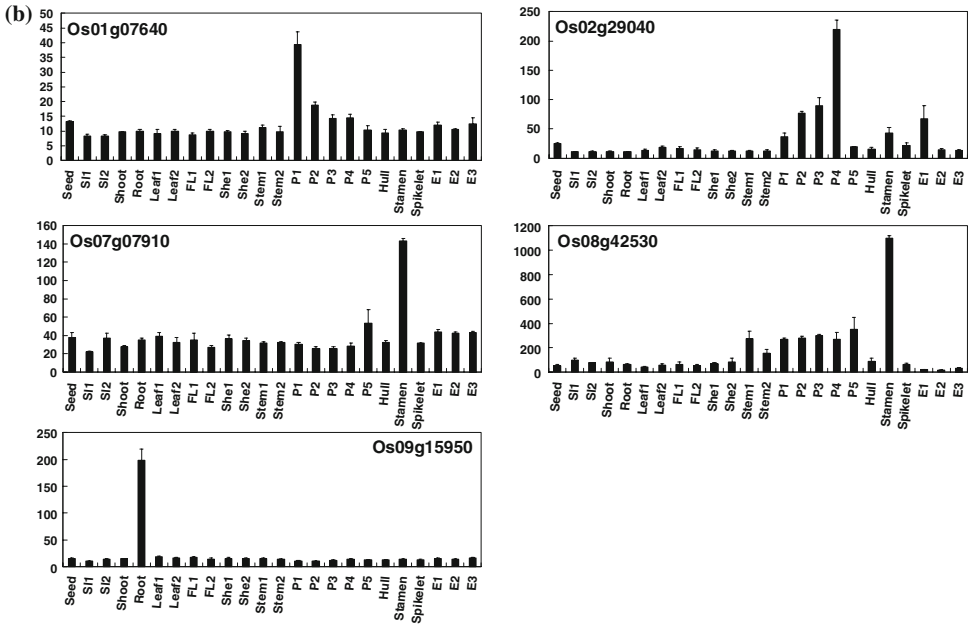


Fig. 6 Expression patterns of 13 preferentially expressed *OsANK* genes in Minghui 63 and 5 genes in Zhenshan 97. **a** 13 preferentially expressed genes in Minghui 63. **b** 5 preferentially expressed genes in Zhenshan 97. The *X*-axis is for the selected tissues and *Y*-axis represents the signal values. Tissue names: *Seed* germinating seed at 72 h of imbibitions; *Sl 1* seedlings of 3 days after sowing; *Sl 2* seedlings at trefoil stage; *Shoot* shoots of seedlings with two tillers; *Root* roots of seedlings with two tillers; *Leaf 1* leaves at secondary branch primordial stage of young panicle development; *Leaf 2* leaves at meiosis stage of young panicle development; *FL 1* flag leaves at 5 days before heading; *FL 2* flag leaves at 14 days after flowering; *She 1* sheath at secondary branch primordial stage of young panicle development; *She 2* sheath at meiosis stage of young panicle development; *Stem 1* stem at 5 days before heading; *Stem 2* stem at heading stage; *P1* panicle at secondary branch primordial stage; *P2* panicle at pistil/stamen primordial differentiation stage; *P3* panicle at pollen-mother cell formation stage; *P4* panicle at meiosis stage of young panicle development; *P5* panicle at heading stage; *Hull* hull at 1 day before flowering; *Stamen* stamen at 1 day before flowering; *Spikelet* spikelet at 3 days after flowering; *E1* endosperm at 7 days after pollination; *E2* endosperm at 14 days after pollination; *E3* endosperm at 21 days after pollination

Os03g16780 and these three genes all belonged to the ANK-RF. Besides, no duplication events occurred among the chromosome 4, 7 and 10.

Totally 65 *OsANK* genes were involved in tandem duplications consisting of 19 clusters or 47 pairs (Fig. 4). The number of *OsANK* genes arranged in the tandem repeats varied from 2 to 9 and most of these genes were assigned in the same orientation on the chromosomes except two pairs (Os09g16200 and Os09g16240; Os11g09230 and Os11g09260). Moreover, the homology of protein sequences of these genes ranged from 21.5 to 99.4% and 17 pairs were juxtaposed with no intervening gene. The detailed information about the duplicated genes could be found in Supplemental Table 4.

Fig. 6 continued



Expression profiling of *OsANK* gene family in the whole life cycle of rice

Microarray analysis was performed using Affymetrix rice microarray for studying the expression pattern of *OsANK* genes. The rice tissues and developmental stages selected for microarray analysis cover the entire life cycle of rice in Minghui 63 and Zhenshan 97. Detailed information on selected stages could be found in methods. The expression profile of the microarray data was confirmed by RT-PCR in the previous articles of our lab (Nayidu et al. 2008; Nuruzzaman et al. 2008; Ye et al. 2009).

Probes for 158 of the 175 *OsANK* genes could be identified in the Affymetrix microarray. Twenty-eight genes had two probe sets and the average signal value of the probe sets was used for analysis. Os01g71540 and Os01g71590 shared the same probe set because of their high sequence homology (99.4%). Os03g47640 had two probe sets in the array and one of the probe sets was shared with Os03g47620. The transcripts of these genes would cross-hybridize in the microarray analyses, which made it difficult to obtain the signal values for each of the genes. We thus used real-time PCR with gene-specific primers to validate the expression patterns of these four genes (Supplemental Fig. 5).

Average expression values for 158 *OsANK* genes of each sample are given in Supplemental Table 5. Based on the signal values, it was obvious that most of the *OsANK* genes were expressed in at least one of the 24 investigated tissues. A hierarchical cluster displaying the logarithm of average signal values for the 158 *OsANK* genes in Minghui 63 is presented in Fig. 5, based on which the expression patterns of *OsANK* genes can be classified into five groups. Twenty-nine genes belonged to group 1, all of them showed high expression levels in all the tissues analyzed. Among the 29 genes, gene Os09g33810 had the highest expression level in the entire life cycle. As the largest group, group 2 consisted of 57 genes, all the genes in this group showed low expression in all the analyzed tissues. Group 3 of 30 genes showed relatively high expression level during reproductive tissues compared to vegetative tissues. Notably, the expression signals of four genes (Os05g03320, Os07g07080, Os09g33600 and Os08g42530) observed in stamen was three to 40-fold higher than that in the vegetative tissues. Five genes (Os02g29130, Os03g47650, Os12g43940, Os12g43840 and Os03g42350) showed the highest expression in developing panicles, one gene (Os11g14570) had the highest expression in hull, and Os12g40780 had high expression level in panicle and endosperm stages. Interestingly, the expression of two genes (Os03g47650 and Os12g43840) which was high in panicles was declined gradually as the panicles matured. Group 4 comprised of 19 genes with predominant expression in vegetative tissues, most of which showed a similar expression pattern in the same tissue at different developmental stages. Os03g46440 had a higher expression signal in flag leaf 2 than in flag leaf 1 and in leaf 2 than in leaf 1. Similarly, Os03g03990 expressed lower in flag leaf 2 than in flag leaf 1 and in leaf 2 than in leaf 1. Another gene (Os11g08020) showed more than 10-fold higher expression in flag leaf 2 than in flag leaf 1. Os01g74320 showed lower expression level in stem 2 than in stem 1. Group 5 consisted of 23 genes showing relatively high expression signal in some vegetative and reproductive tissues compared with others. Os01g74320

had high expression level in flag leaf 2 and stamen. Interestingly this gene showed more than 60-fold higher expression in flag leaf 2 than in flag leaf 1.

Most genes had the same expression pattern in Zhenshan 97 as in Minghui 63 except 13 genes (Os01g09370, Os01g07980, Os02g29160, Os02g29190, Os02g29210, Os03g17250, Os05g01310, Os05g03320, Os09g03630, Os09g03680, Os09g15950, Os09g16050 and Os11g24850). The expression patterns of *OsANK* genes in Zhenshan 97 are given in Supplemental Fig. 6. Notably, Os09g03680 had low expression signal in Minghui 63 while having relatively high expression signal in leaf, flag leaf and sheath in Zhenshan 97. Similarly, Os09g03630 had relatively high expression level in Zhenshan 97 but showed very low expression in Minghui 63 in almost all the vegetative tissues.

With the aim of revealing *OsANK* gene expression features, an analysis of preferential expression was performed. Finally, 13 and 5 *OsANK* genes were identified in Minghui 63 and Zhenshan 97 respectively which showing preferential expression in a given stage (Figs. 6a, b; 7). Among these, only two genes (Os07g07910 and Os08g42530) showed the same expression pattern in two genotypes (Fig. 7). Out of the 16 genes, nine showed preferential expression in stamen at 1 day before flowering; one (Os02g29040) was preferentially expressed in panicle 4 and two (Os09g15950 and Os11g24840) in root; Os01g07640, Os04g36740, Os09g16050 and Os07g34830 were preferentially expressed in panicle 1, seed, stem 1 and flag leaf 1, respectively.

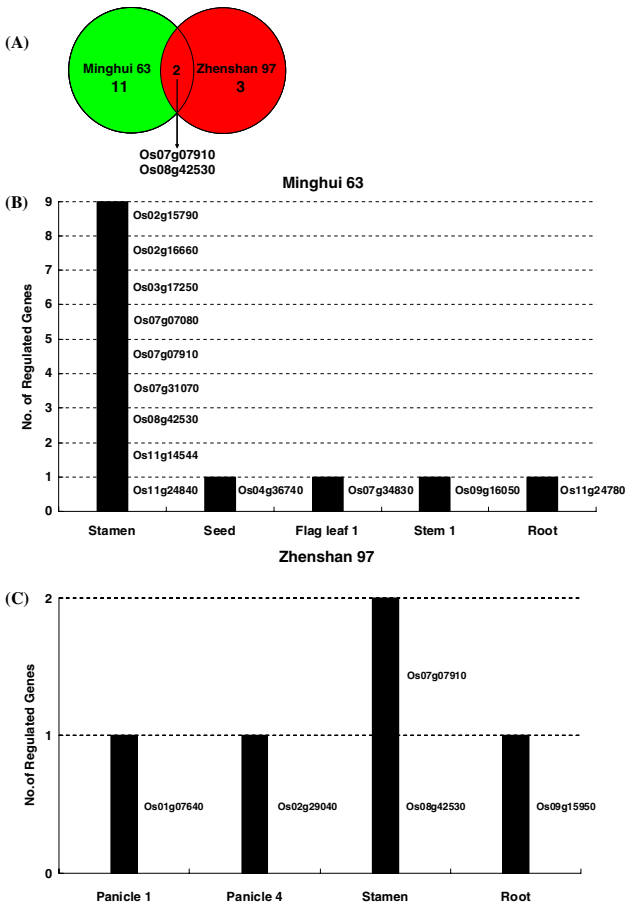


Fig. 7 Summary of expression analyses of *OsANK* genes that show preferential expression in Minghui 63 and Zhenshan 97. **a** Venn diagram indicates the numbers of genes showing preferential expression in Minghui 63, Zhenshan 97 or in both genotypes; **b** and **c** *Black bars* indicate the numbers of genes showing preferential expression in Minghui 63 and Zhenshan 97. Corresponding genes are listed on the *right side* of this column. Tissues are mentioned at the *bottom* of the column

Responses of *OsANK* genes to NAA, KT, GA3 and light/dark treatments

Phytohormones play a critical role in plant growth and development. To investigate the *OsANK* genes in response to phytohormone treatment, microarray analysis was performed. We identified a total of 18 *OsANK* genes that were differentially expressed with treatments of one or more of the phytohormone NAA, KT, GA3 in the two genotypes compared with the control (Fig. 8). The fold change values with respect to control are given in Supplemental Table 6. Among these, 15 genes were up-regulated while three genes were down-regulated. The down-regulated genes were Os05g02130 treated by NAA in Minghui 63 and Os07g07080 and Os09g34280 treated by GA3 in Zhenshan 97. Seven genes showed differential expression with phytohormone treatments in both genotypes. Interestingly, two genes (Os04g48520 and Os06g03800) were both up-regulated with all three phytohormone treatments in both genotypes. However, the expression profiles of the remaining genes in two genotypes were different. For instance, Os01g07980 was up-regulated specifically by GA3 treatment in Zhenshan 97 while it was up-regulated by all three phytohormone treatments in Minghui 63. Conversely, Os06g13000 was up-regulated by all three phytohormone treatments in Zhenshan 97 but up-regulated by GA3 and KT treatments in Minghui 63. Six genes (Os02g29210, Os03g04300, Os05g23320, Os07g07080, Os09g34280 and Os12g43940) were differentially expressed only in Zhenshan 97 while five genes (Os01g08000, Os01g09370, Os03g03990, Os05g02130 and Os11g08020) only in Minghui 63 (Figs. 9, 10). The fold change values with respect to seedlings treated by phytohormone can be obtained in Supplemental Table 6.

To investigate the light regulation of *OsANK* genes, expression profiles of *OsANK* genes in seedlings (plumule and radicle tissues) treated with light or dark for 48 h were also investigated, with the fold changes with respect to control listed in Supplemental Table 6. Nine *OsANK* genes

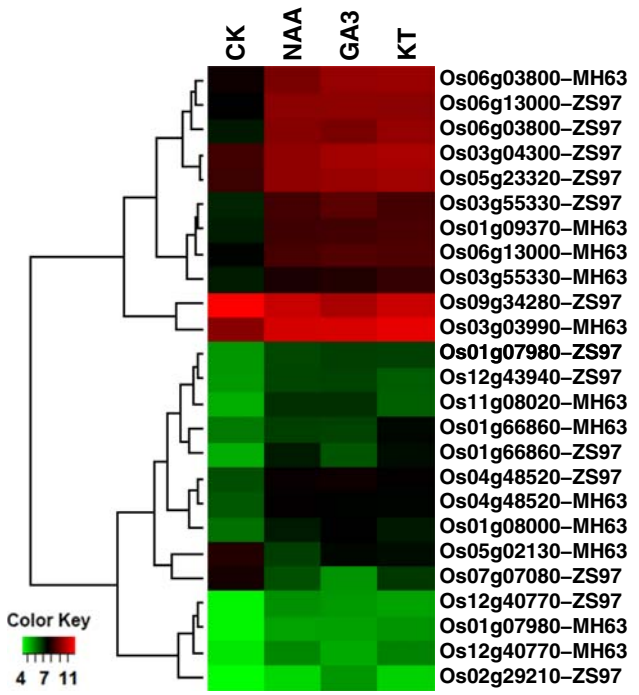


Fig. 8 Clustering of expression profiles of 18 *OsANK* genes showing differential expression in 7-day-old seedlings with three phytohormone (NAA, GA3 and KT) treatments. The *color scale* representing log2 signal values is shown on the *left*. Cluster dendrogram is shown on the *left*. Differentially expressed genes are listed on the *right* of each lane. The different treatments are mentioned on the *top* of each lane

showed differential expression at the plumule stage under light compared to dark in Minghui 63 and Zhenshan 97 (Fig. 11). Four genes (Os03g03990, Os03g04300, Os06g03800 and Os06g13000) were in common in the two genotypes. Amongst the nine genes, eight were up-regulated and one (Os02g29210) was down-regulated in Zhenshan 97. We also found four genes differentially expressed at the radicle stage under light compared to dark in both Minghui 63 and Zhenshan 97 (Fig. 11). Interestingly, three of the four genes (1 in Minghui 63 and 2 in Zhenshan 97) were down-regulated and each was differentially expressed only in one genotype.

Comparisons of expression profiles of the duplicated *OsANK* genes

Duplicated genes had three alternative outcomes in the evolutionary course: nonfunctionalization, neofunctionalization and subfunctionalization (Lynch and Conery 2000). The expression patterns for *OsANK* genes present in segmental and tandem duplicated regions were examined in Minghui 63. Probe sets were available for seven of the nine pairs of genes located in the segmental duplicated regions. Five pairs of genes showed highly similar expression

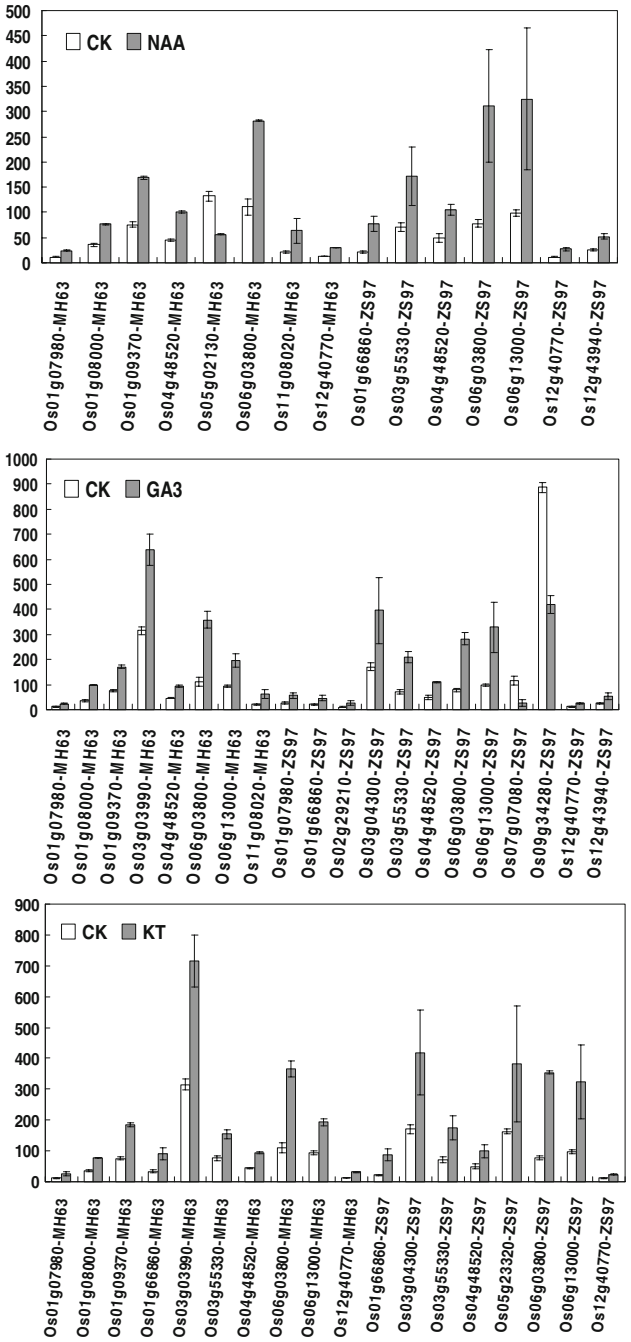


Fig. 9 Expression profiles of 18 differentially expressed *OsANK* genes in 7-day-old seedlings subjected to three phytohormones (NAA, GA3 and KT). *X*-axis represents the differentially expressed genes. *Y*-axis represents average expression values. Each treatment is mentioned at the *top* of each diagram

patterns in most of the tested tissues, indicating subfunctionalization after the duplication events. Os03g47686 was not expressed in all the selected tissues, which may indicate that one of the members lost its function during the course of evolution.

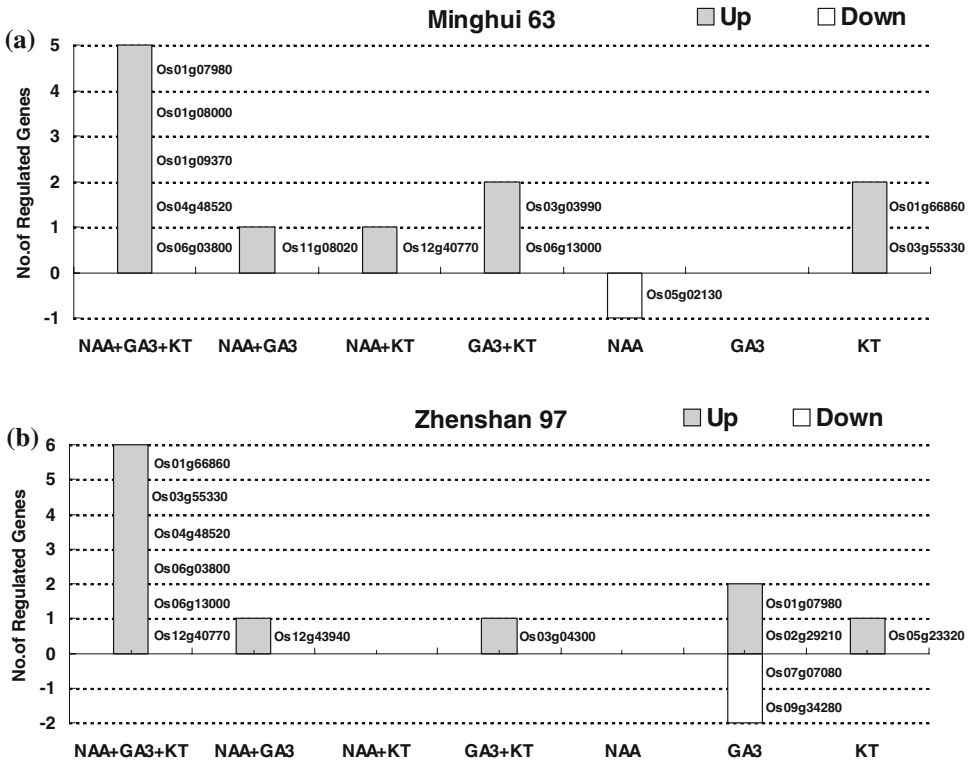


Fig. 10 Summary of expression analyses of *OsANK* genes responsive to phytohormone treatments in Minghui 63 (**a**) and Zhenshan 97 (**b**). *Grey bars* indicate the numbers of genes up-regulated by the phytohormone, and the corresponding genes are listed in the *right* of the column. *White bars* indicate the numbers of genes down-regulated by the phytohormone. Corresponding genes are also listed in the *right* of the column. The types of phytohormone are listed in the *bottom* of the column

Of the 19 clusters of tandem duplicated genes, six genes (Os02g29030, Os03g47693, Os03g47702, Os05g03390, Os07g09550 and Os07g33210) did not have probe sets on Affymetrix microarray. Therefore, we analyzed the rest 14 clusters of tandem duplicated genes. In the largest cluster of nine members on chromosome 11, the expression patterns of two tandem duplicated genes (Os11g24780 and Os11g24840) were different, which may indicate neofunctionalization. Of the four members (Os11g07980, Os11g08020, Os11g08050 and Os11g08070) of the gene cluster on chromosome 11, expression for three of them was nearly complete absence in the stages analyzed, indicating nonfunctionalization after duplications. The expression patterns of some segmental and tandem duplication genes are showed in Fig. 12.

Discussion

Ankyrin repeat containing proteins as identified in different plant species are involved in various developmental processes, including signaling pathways, stress responses, plant defenses and so on. We identified 175 *OsANK* genes in the rice genome, which can be classified in 10 subfamilies and are distributed on all 12 chromosomes.

Gene duplications are one of the primary driving forces in the evolution of genomes and genetic systems (Moore and Purugganan 2003). In addition, segmental duplications occur more often in more slowly evolving gene family, i.e. MYB gene family. Whereas, tandem duplication in local genomic clusters with low retention of segmental duplications is common in the large and rapidly evolving gene family, i.e. NBS-LRR disease resistance family (Cannon et al. 2004). Our analysis revealed that the number of *OsANK* genes arranged in tandem was much larger than that those involved in chromosomal segments duplication events. Therefore, we hypothesized that *OsANK* gene family is a large, rapidly evolving gene family during the process of evolution. Besides, the tandem duplications of chromosomal regions might have played a key role in the expansion of the *OsANK* gene family. This phenomenon was also be observed in some other rice gene families, i.e. F-box family (Jain et al. 2007). Interestingly, most of the tandem duplicated genes belonged to ANK-M, ANK-TM and ANK-TPR subfamilies, providing explanation for the large number of these three subfamilies, especially the ANK-TPR. Since the members of ANK-TPR (22 members) in rice is much more than that in *Arabidopsis* (only one member), ANK-TPR of rice may undergoes multiple duplication events to gain more members during the process of evolution.

Since ANK appears to be conserved in structure rather than in function (Bork 1993), knowing the expression profile of *OsANK* genes may provide clues to the function of each gene. Our analyses suggested that the expression patterns of the 158 *OsANK* genes could be classified into five groups. Most of the *OsANK* genes in group 2 showed

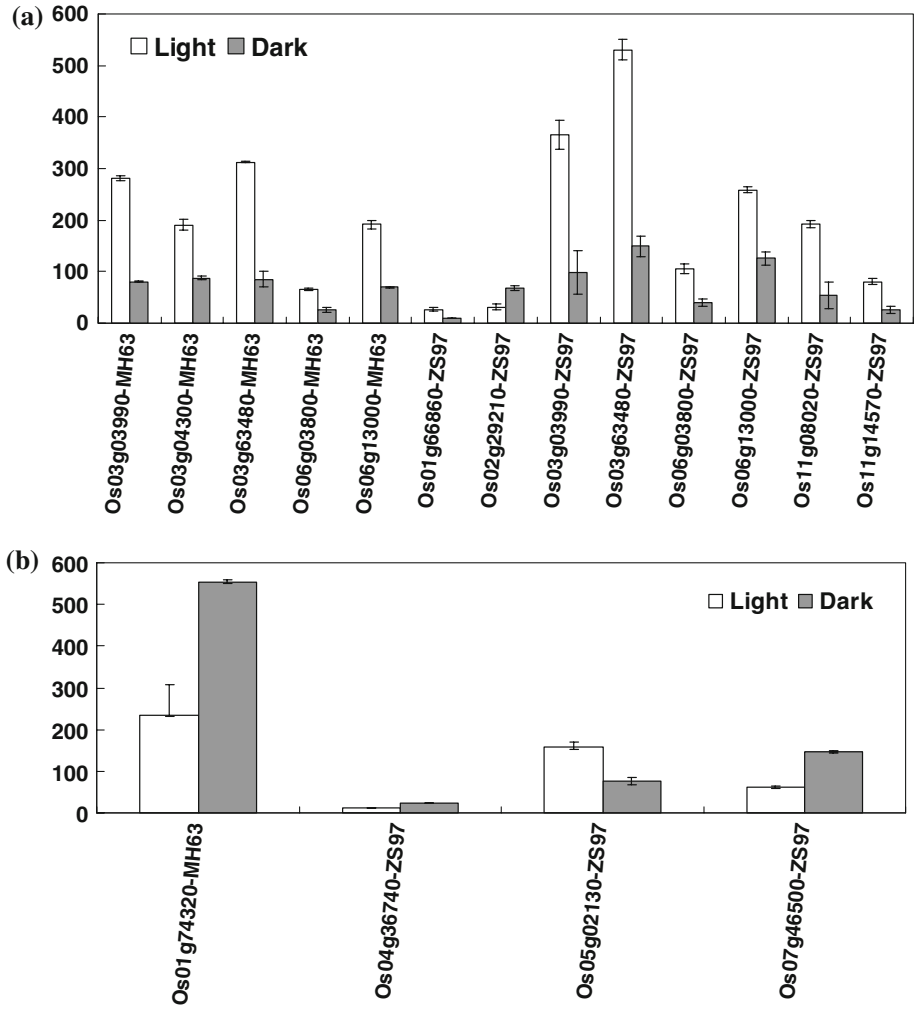


Fig. 11 Histogram diagrams of 13 differentially expressed *OsANK* genes regulated by light and dark treatments. **a** Nine genes showing differential expression at plumule stage (48 h after emergence) with light and dark treatments in Minghui 63 and Zhenshan 97. **b** Four genes showing differential expression in radicle (48 h after emergence) with light and dark treatments in Minghui 63 and Zhenshan 97. *X*-axis represents the differentially expressed genes. *Y*-axis represents the average expression values

low expression level in all the analyzed tissues but not assigned to have non-functions. An ANK Membrane Protein (IGN1) has been reported in *Lotus japonicus* and it was required for differentiation and/or persistence of bacteroids and symbiosomes. The expression of *IGN1* was detected in flowers, shoots, roots and nodules at low levels while the expression levels did not enhance in the nodulation process (Kumagai et al. 2007). Group 3 had relatively high expression level in reproductive tissues compared with vegetative tissues. XB3 belonging to group 3 was required for a pathogen-induced type of programmed cell death (Wang et al. 2006). Through our expression profile analysis, *XB3* showed high transcript accumulation in seed and stem. Therefore, the variability in the expression patterns of genes of the same family indicated that their roles might not be redundant.

Moreover, 16 genes were identified to show preferential expression in specific tissues in two genotypes in this study. Interestingly, nine genes showed preferential expression in stamen of 1 day before flowering suggesting important roles in pollination and fertilization. In lily, *LIANK* was required for pollen germination and pollen tube growth. RNA gel blots evaluated that *LIANK* was expressed in fresh pollen and up-regulated significantly during germination and tube growth (Huang et al. 2006). Furthermore, two *OsANK* genes (Os09g15950 and Os11g24780) were preferentially expressed in root. It is possible that these two genes are involved in the regulation of root growth. Nodzon et al. (2004) reported that the ubiquitin ligase XBAT32 played a role in the initiation of lateral root in *Arabidopsis*. High expression of *XBAT32* was observed in the vascular system of primary roots, particularly in the zone where lateral root initiation occurs (Nodzon et al. 2004). *TIP1* regulated root hair growth and it was expressed in root, leaf, inflorescence stem and floral tissue (Hemsley et al. 2005). We also identified genes preferentially expressed in leaf. It was reported that *AKR* was involved in the regulatory process of chloroplast differentiation in *Arabidopsis*. The *AKR* message was the highest in leaves, less in flowers and subtending floral

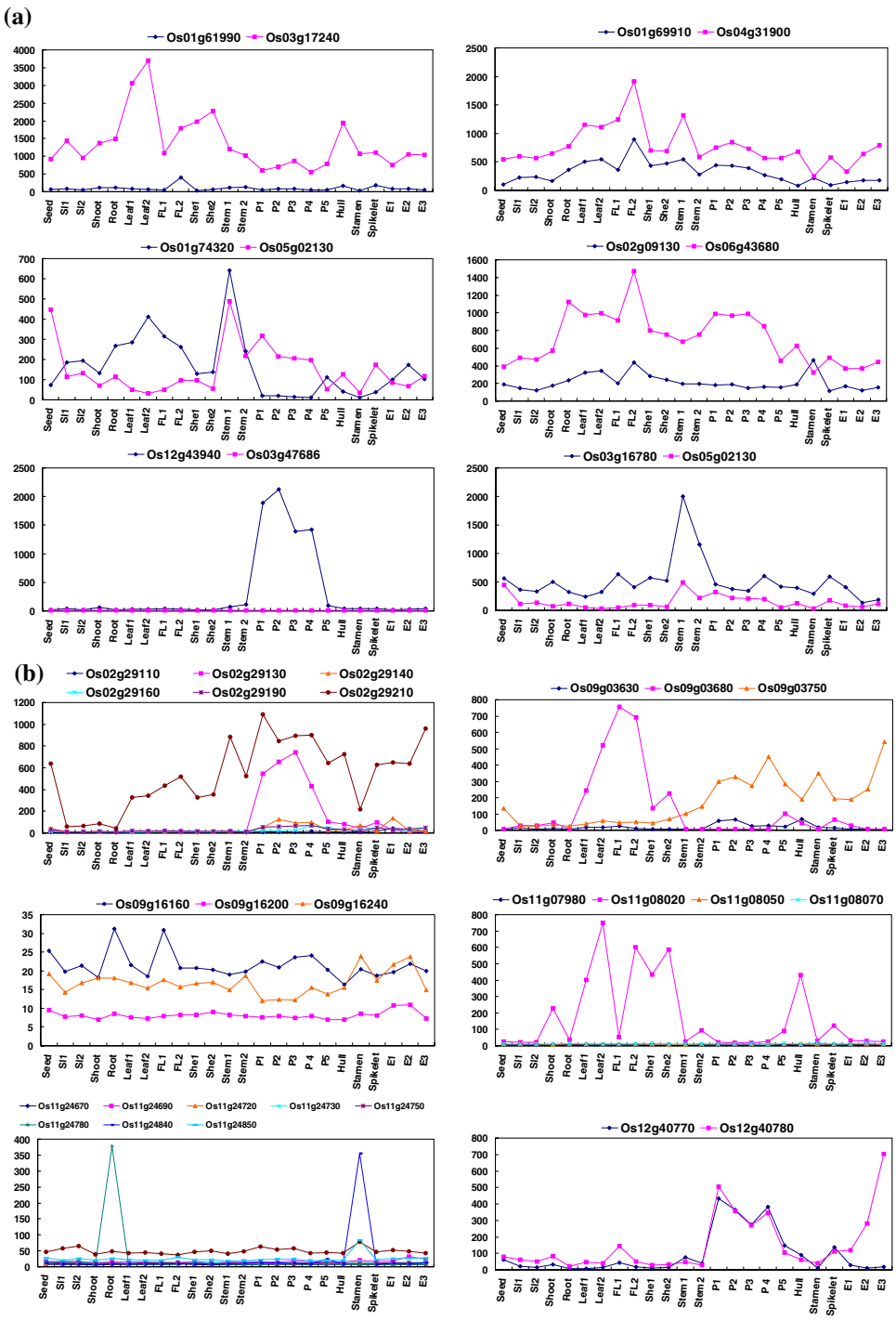


Fig. 12 Examples of expression patterns of *OsANK* genes found in duplicated regions of the rice genome. **a** Expression patterns of some *OsANK* genes found in segmental duplicated regions. **b** Expression patterns of some tandem duplicated *OsANK* genes. *X*-axis represents the developmental stages. *Y*-axis indicates the average expression values. The gene names are listed on the *top* of diagram

bracts, and the least in inflorescence stems (Zhang et al. 1992). Other two genes *BOP1* (Ha et al. 2004; Norberg et al. 2005) and *GBP1* (Wirdnam et al. 2004) were characterized to be involved in leaf development. Therefore, these genes preferentially expressed in specific tissues may deserve special notice for further investigation of their functions.

Some ANK genes in plants were shown to be affected by auxin (Nodzon et al. 2004; Koizumi et al. 2005), salicylic acid (Cao et al. 1997; Lu et al. 2003) and ABA (AbuQamar et al. 2006; Sakamoto et al. 2008). We identified a total of 18 *OsANK* genes that were differentially expressed under one or more phytohormone treatments in two genotypes. Putative orthologs of *Arabidopsis* ANK proteins known to be involved in hormone response could be demarcated in rice based on their phylogenetic relationship, such as Os02g54860 for XBAT32. Although these genes did not show differential expression with the

three phytohormone treatments, the possibility of regulating by phytohormone cannot be completely ruled out.

ANK genes were also shown to be regulated by light. In *Arabidopsis*, the *AKR* (Zhang et al. 1992) and *ACD6* (Lu et al. 2003) genes appeared to be regulated by light in leaves. Our data identified 13 *OsANK* genes as showing response to light/dark treatments indicating that most of the *OsANK* genes were not sensitive to light or dark. Os03g04300, a closely related ortholog of *AKR*, was more than two-fold up-regulated by light treatment in Minghui 63. However, no closely related ortholog of *ACD6* could be predicted in rice. This study provided evidence that thirteen candidate *OsANK* genes are likely to be regulated by light.

In conclusion, the preferentially expression in specified tissues and the response to phytohormone and light treatments of the *OsANK* genes provided clues to the roles of these genes in signaling, growth and development. The systematic sequence analysis and expression profiles of *OsANK* genes can serve a very useful reference for more detailed functional analysis and will be helpful in selection of appropriate candidate genes for further study of ANK genes in rice.

Acknowledgments We thank the reviewers for their careful reading and helpful comments on this manuscript. This research was supported in part by grants from the National Special Key Project of China on Functional Genomics of Major Plants and Animals, and the National Natural Science Foundation of China.

References

AbuQamar S, Chen X, Dhawan R, Bluhm B, Salmeron J, Lam S, Dietrich RA, Mengiste T (2006) Expression profiling and mutant analysis reveals complex regulatory networks involved in Arabidopsis response to Botrytis infection. Plant J 48:28–44

Albert S, Despres B, Guilleminot J, Bechtold N, Pelletier G, Delseny M, Devic M (1999) The EMB 506 gene encodes a novel ankyrin repeat containing protein that is essential for the normal development of Arabidopsis embryos. Plant J 17:169–179

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

Bae W, Lee YJ, Kim DH, Lee J, Kim S, Sohn EJ, Hwang I (2008) AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. Nat Cell Biol 10:220–227

Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34:W369–W373

Becerra C, Jahrmann T, Puigdomenech P, Vicient CM (2004) Ankyrin repeat-containing proteins in Arabidopsis: characterization of a novel and abundant group of genes coding ankyrin-transmembrane proteins. Gene 340:111–121

Bork P (1993) Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? Proteins 17:363–374

Breeden L, Nasmyth K (1987) Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila. Nature 329:651–654

Cannon SB, Mitra A, Baumgarten A, Young ND, 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 H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell 88:57–63

Choi MS, Kim MC, Yoo JH, Moon BC, Koo SC, Park BO, Lee JH, Koo YD, Han HJ, Lee SY, Chung WS, Lim CO, Cho MJ (2005) Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). J Biol Chem 280:40820–40831

Clewley JP, Arnold C (1997) MEGALIGN. The multiple alignment module of LASERGENE. Methods Mol Biol 70:119–129

Cooke HJ (2004) Silence of the centromeres—not. Trends Biotechnol 22:319–321

Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755–763

Garcion C, Guilleminot J, Kroj T, Parcy F, Giraudat J, Devic M (2006) AKRP and EMB506 are two ankyrin repeat proteins essential for plastid differentiation and plant development in Arabidopsis. Plant J 48:895–906

Ha CM, Jun JH, Nam HG, Fletcher JC (2004) BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in *Arabidopsis thaliana*. Plant Cell Physiol 45:1361–1370

Hemsley PA, Kemp AC, Grierson CS (2005) The TIP GROWTH DEFECTIVE1 *S*-acyl transferase regulates plant cell growth in Arabidopsis. Plant Cell 17:2554–2563

Huang J, Chen F, Del Casino C, Autino A, Shen M, Yuan S, Peng J, Shi H, Wang C, Cresti M, Li Y (2006) An ankyrin repeat-containing protein, characterized as a ubiquitin ligase, is closely associated with membrane-enclosed organelles and required for pollen germination and pollen tube growth in lily. Plant Physiol 140:1374–1383

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C (2009) InterPro: the integrative protein signature database. Nucleic Acids Res 37:D211–D215

Jain M, Nijhawan A, Arora R, Agarwal P, Ray S, Sharma P, Kapoor S, Tyagi AK, Khurana JP (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

Ketchum KA, Slayman CW (1996) Isolation of an ion channel gene from *Arabidopsis thaliana* using the H5 signature sequence from voltage-dependent K+ channels. FEBS Lett 378:19–26

Koizumi K, Naramoto S, Sawa S, Yahara N, Ueda T, Nakano A, Sugiyama M, Fukuda H (2005) VAN3 ARF-GAP-mediated vesicle transport is involved in leaf vascular network formation. Development 132:1699–1711

Kumagai H, Hakoyama T, Umehara Y, Sato S, Kaneko T, Tabata S, Kouchi H (2007) A novel ankyrin-repeat membrane protein, IGN1, is required for persistence of nitrogen-fixing symbiosis in root nodules of *Lotus japonicus*. Plant Physiol 143:1293–1305

Li HY, Chye ML (2004) Arabidopsis acyl-CoA-binding protein ACBP2 interacts with an ethylene-responsive element-binding protein, AtEBP, via its ankyrin repeats. Plant Mol Biol 54:233–243

Liang C, Jaiswal P, Hebbard C, Avraham S, Buckler ES, Casstevens T, Hurwitz B, McCouch S, Ni J, Pujar A, Ravenscroft D, Ren L, Spooner W, Tecle I, Thomason J, Tung CW, Wei X, Yap I, Youens-Clark K, Ware D, Stein L (2008) Gramene: a growing plant comparative genomics resource. Nucleic Acids Res 36:D947–D953

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25:402–408

Lu G, Moriyama EN (2004) Vector NTI, a balanced all-in-one sequence analysis suite. Brief Bioinform 5:378–388

Lu H, Rate DN, Song JT, Greenberg JT (2003) ACD6, a novel ankyrin protein, is a regulator and an effector of salicylic acid signaling in the Arabidopsis defense response. Plant Cell 15:2408–2420

Lu H, Liu Y, Greenberg JT (2005) Structure-function analysis of the plasma membrane-localized Arabidopsis defense component ACD6. Plant J 44:798–809

Lux SE, John KM, Bennett V (1990) Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. Nature 344:36–42

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

Michaely P, Bennett V (1992) The ANK repeat: a ubiquitous motif involved in macromolecular recognition. Trends Cell Biol 2:127–129

Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. Proc Natl Acad Sci USA 100:15682–15687

Mosavi LK, Cammett TJ, Desrosiers DC, Peng ZY (2004) The ankyrin repeat as molecular architecture for protein recognition. Protein Sci 13:1435–1448

Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J (2004) Sequencing of a rice centromere uncovers active genes. Nat Genet 36:138–145

Nayidu NK, Wang L, Xie W, Zhang C, Fan C, Lian X, Zhang Q, Xiong L (2008) Comprehensive sequence and expression profile analysis of PEX11 gene family in rice. Gene 412:59–70

Nodzon LA, Xu WH, Wang Y, Pi LY, Chakrabarty PK, Song WY (2004) The ubiquitin ligase XBAT32 regulates lateral root development in Arabidopsis. Plant J 40:996–1006

Norberg M, Holmlund M, Nilsson O (2005) The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. Development 132:2203–2213

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

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

Sakamoto H, Matsuda O, Iba K (2008) ITN1, a novel gene encoding an ankyrin-repeat protein that affects the ABA-mediated production of reactive oxygen species and is involved in salt-stress tolerance in *Arabidopsis thaliana*. Plant J 56:411–422

Sedgwick SG, Smerdon SJ (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem Sci 24:311–316

Seong ES, Cho HS, Choi D, Joung YH, Lim CK, Hur JH, Wang MH (2007a) Tomato plants overexpressing CaKR1 enhanced tolerance to salt and oxidative stress. Biochem Biophys Res Commun 363:983–988

Seong ES, Choi D, Cho HS, Lim CK, Cho HJ, Wang MH (2007b) Characterization of a stress-responsive ankyrin repeat-containing zinc finger protein of *Capsicum annuum* (CaKR1). J Biochem Mol Biol 40:952–958

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

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

Wang YS, Pi LY, Chen X, Chakrabarty PK, Jiang J, De Leon AL, Liu GZ, Li L, Benny U, Oard J, Ronald PC, Song WY (2006) Rice XA21 binding protein 3 is a ubiquitin ligase required for full Xa21-mediated disease resistance. Plant Cell 18:3635–3646

Wirdnam C, Motoyama A, Arn-Bouldoires E, van Eeden S, Iglesias A, Meins F Jr (2004) Altered expression of an ankyrin-repeat protein results in leaf abnormalities, necrotic lesions, and the elaboration of a systemic signal. Plant Mol Biol 56:717–730

Xu X, Jiang CZ, Donnelly L, Reid MS (2007) Functional analysis of a RING domain ankyrin repeat protein that is highly expressed during flower senescence. J Exp Bot 58:3623–3630

Yan J, Wang J, Zhang H (2002) An ankyrin repeat-containing protein plays a role in both disease resistance and antioxidation metabolism. Plant J 29:193–202

Ye S, Wang L, Xie W, Wan B, Li X, Lin Y (2009) Expression profile of calcium-dependent protein kinase (CDPKs) genes during the whole lifespan and under phytohormone treatment conditions in rice (*Oryza sativa* L. ssp. *indica*). Plant Mol Biol 70:311–325

Zhang Q (2007) Strategies for developing green super rice. Proc Natl Acad Sci USA 104:16402–16409

Zhang H, Scheirer DC, Fowle WH, Goodman HM (1992) Expression of antisense or sense RNA of an ankyrin repeat-containing gene blocks chloroplast differentiation in Arabidopsis. Plant Cell 4:1575–1588

Zhang Q, Li J, Xue Y, Han B, Deng XW (2008) Rice 2020: A call for an international coordinated effort in rice functional genomics. Mol Plant 1:715–719