DNA Research **20**, 67–78, (2013) doi:10.1093/dnares/dss034
Advance Access publication on 17 December 2012

Structural, Expression and Interaction Analysis of Rice *SKP1-Like* Genes

Senda Kahloul[1,2], Imen HajSalah El Beji[1], Aurélia Boulaflous[1], Ali Ferchichi[2], Hongzhi Kong[3], Said Mouzeyar[1,*], and Mohamed Fouad Bouzidi[1]

Clermont Université, Université Blaise Pascal, UMR INRA-UBP 1095 GDEC, BP 10448, Clermont-Ferrand F-63000, France[1]; Laboratoires d'Aridocultures et Cultures Oasiennes-IRA, Médenine 4119, Tunisie[2] and Institute of Botany, The Chinese Academy of Sciences, Beijing 100093, China[3]

*To whom correspondence should be addressed. Tel. +33-(0)4-73-40-79-13. Fax. +33-(0)473-40-79-14. Email: said.mouzeyar@univ-bpclermont.fr

(Received 12 July 2012; accepted 13 November 2012)

Abstract

The degradation of proteins by the 26S proteasome is initiated by protein polyubiquitination mediated by a three-step cascade. The specific ubiquitination of different target proteins is mediated by different classes of E3 ubiquitin ligases, among which the best known are Skp1-Cullin-F-box complexes. Whereas protists, fungi and some vertebrates have a single *SKP1* gene, many animal and plant species possess multiple *SKP1* homologues. In this paper, we report on the structure, phylogeny and expression of the complete set of rice *SKP1* genes (*OSKs*, *Oryza sativa SKP1*-like genes). Our analyses indicated that *OSK1* and *OSK20* belong to a class of *SKP1* genes that contain one intron at a conserved position and are highly expressed. In addition, our yeast two-hybrid results revealed that OSK proteins display a differing ability to interact with F-box proteins. However, OSK1 and OSK20 seemed to interact with most of the nine F-box proteins tested. We suggest that rice *OSK1* and *OSK20* are likely to have functions similar to the *Arabidopsis ASK1* and *ASK2* genes.

Key words: SKP1; rice; phylogenetic analysis; structural analysis; yeast two-hybrid screening

1. Introduction

Protein degradation is an important post-transcriptional regulatory process that allows cells to respond rapidly to intracellular signals and changing environmental conditions by adjusting the levels of key proteins. One major proteolytic pathway in eukaryotes involves the Ubiquitin/26S Proteasome System (UPS) that utilizes the post-translational modification of proteins by ubiquitin. The UPS involves the enzymatic activities required for the polyubiquitylation of target proteins, and then for the proteolysis of these tagged proteins by the proteasome. This polyubiquitylation is achieved by the sequential actions of three enzymes: an E1 ubiquitin-activating enzyme, an E2 ubiquitin-conjugating enzyme and an E3 ubiquitin ligase. E1 activates and links free ubiquitins that will be transferred to one of the E2 enzymes, whereas E3 ligase enzymes promote ubiquitylation by mediating a specific interaction between E2 enzymes and the target protein.[1–4]

The specificity of the ubiquitination process is ensured by the E3 ubiquitin ligases encoded by a large multigene family. A survey of sequenced plant genomes showed that they contain more than 1000 E3 ubiquitin ligases. For example, the annotated *Arabidopsis thaliana*, rice and soybean genomes contain 1415, 1332 and 1402 E3 ligase-encoding genes, respectively.[5,6] E3s comprise a diverse family of proteins or protein complexes that can be distinguished in terms of their mechanisms of action and types of interactions. They act either as single subunits, such as homology to E6-AP C terminus, really interesting new gene and U-box or multisubunit complexes such as SkP1 S-phase Kinase-associated Protein 1–CulL1 (Cullin 1)–F-box (SCF), Cullin3-Bric

© The Author 2012. Published by Oxford University Press on behalf of Kazusa DNA Research Institute.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

a brac-Tramtrack-Broad, Cullin 4-DNA damage-binding 1 and anaphase-promoting complex/cyclosome.[3,5]

SCF complexes are a major type of E3 ubiquitin ligase consisting of two constant subunits (ROC1 or Rbx1 and Cullin-1), a moderately variable subunit (SKP1) and a highly variable substrate recognition subunit known as F-box protein.[7,8] F-box proteins contain a conserved F-box domain (40–50 amino acids) at the N-terminus, which interacts with Skp1 proteins. They have highly variable protein–protein interaction domains at the C-terminus, which confer substrate specificity for ubiquitination.[7,9] Beside the canonical SCF, it appears that in some cases, Skp1 and F-box proteins may function in non-SCF complexes and that some F-box proteins have functions of their own.[10] F-box proteins are represented by large families in different organisms. There are 11 F-box protein-encoding genes in budding yeast, 68 in humans, 74 in *Caenorhabditis elegans*, 337 in poplar, 539 in medicago, 692 in *Arabidopsis* and 779 in rice.[11,12] Different research groups have classified *Arabidopsis* and rice F-box proteins in several subfamilies or subgroups based on their phylogenetic relationships or unique functional domains.[7,9,13] Whereas humans and yeasts have a single *SKP1* gene, many animal and plant species possess multiple *SKP1* homologues[14], for example, 7 in *Drosophila*, 21 in *C. elegans* denoted *Skr*, 21 in *C. elegans* and 21 in *Arabidopsis* denoted ASK and 32 in rice denoted OSK.[3,15–17] Based on the finding that the *C. elegans* genome contains 21 *SKP1-related* genes,[15,16] these authors questioned whether all these genes could interact with other components in the SCF complex (i.e. Cullin1 and F-box proteins). Both studies used the Y2H system to demonstrate that Skr proteins displayed varied patterns of interaction with Cullins and F-box proteins. While Skr1 (which is thought to have a function similar to that of the human SKP1) interacted with most of the F-box tested, and the remaining Skr displayed differing interaction capabilities.[16] In addition, both studies showed that the 21 *C. elegans* Skrs displayed various expression profiles and tissue-specific patterns and different RNAi phenotypes, indicating their involvement in a variety of pathways.[15,16] Subsequently, several exhaustive yeast two-hybrid screens were used to analyse the *Arabidopsis* SKP1-like family (ASKs). It was demonstrated that the 21 *Arabidopsis* ASK proteins displayed substantial differences in their abilities to interact with different F-box proteins and that ASK1, ASK2, ASK11 and ASK12 could interact with COI1, FKF1, UFO and other F-box proteins.[18] The ASK1 and ASK2 genes seem to be the most important SKP1 homologues in *A. thaliana*; both of them are expressed widely and at high levels.[8,14,18]

During a phylogenetic study, Kong et al. analysed the evolution of *SKP1-like* genes in a wide range of plant species, but with particular emphasis on *Arabidopsis* and rice and suggested that all the *SKP1* genes found in these species derived from a single ancestral gene represented by *ASK1* in *Arabidopsis* and *OSK1* in rice, and that these genes could have similar functions.[14]

Despite the importance of the SCF complex, there have been few reports of systematic surveys of interactions between the dozens of SKP1-like proteins and the hundreds of F-box proteins. The objective of the present study was, therefore, to uncover the various capabilities of SKP1-like proteins to interact with F-box proteins in rice. Nine F-box proteins representative of the most frequent domains containing F-box proteins, and the whole set of SKP1-like proteins (OSK), were studied. In total, 540 binary interactions were tested using the yeast two-hybrid approach. We showed that the 30 rice SKP1-like proteins displayed various interaction patterns with the tested F-box proteins and that OSK1 and OSK20 exhibited the most frequent interaction capabilities. Our results also suggest that rice OSK1 and OSK20 might be functionally equivalent to ASK1 and ASK2 in *Arabidopsis*.

2. Methods

2.1. Data retrieval

Kong et al.[17] evoked the presence of 32 *SKP1-like* genes in the rice genome, but described only 28 genes in their phylogenetic study. We followed the same nomenclature during our study, although we added three *OSK* genes (*OSK8*, *OSK24* and *OSK31)* and replaced Kong's *OSK14* with a more accurate accession. For details, see Supplementary Table S1. In addition, *OSK18*, which contains an internal retrotransposon, was excluded from subsequent studies. Thus, a total of 30 *OSK* genes were retained for further analyses. Nine F-box proteins belonging to F-box families in the rice genome were selected from the inventory compiled by Jain et al.[7] The Rice Genome Annotation Project [RICECHIP.ORG: Support for Annotation & Functional Analysis of the *Oryza sativa* (Rice) Genome (http://www.ricechip.org/)] was used to identify the number of Expressed Sequence Tags (EST) for *OSK* and F-box genes in various rice tissues (July 2011).

The PLAZA platform (http://bioinformatics.psb.ugent.be/plaza/, version 2.5) was queried to retrieve *SKP1-like* genes from the sequenced genomes of 11 eudicot, 5monocot and 1 moss species. All retrieved sequences were checked manually for consistency and for the presence of *SKP1* signatures, using the InterProScan program.[19]

2.2. Structural analysis of the SKP1 family

Multiple protein sequence alignments were generated using the Muscle software[20] implemented under MEGA5.[21] The maximum likelihood method was used, and phylogeny was tested with a bootstrap of 500 (Poisson's correction model) to construct a phylogenetic tree with MEGA5.[21] The entire set of 288 SKP1-like proteins from the moss, monocot and eudicot species retrieved from Plaza (http://bioinformatics.psb.ugent.be/plaza/) were checked for the presence and positions of introns within the gene. The conservation of intron position was used as a criterion to identify putative ancestral *SKP1-like* genes, as suggested by Kong et al.[17] In this study, an intron was considered to be conserved, if it occurred between two aligning bases in the alignment of the coding sequences.[22]

2.3. Meta-analysis of the expression of rice SKP1-like genes

The Affymetrix probe sets corresponding to rice *SKP1-like* genes were retrieved from the Affymetrix website (http://www.affymetrix.com). Probe sets that recognized multiple genes were excluded from the analysis. Organ distribution (anatomy analysis) of the expression of rice *SKP1* genes was determined using Genevistegator (https://www.genevestigator.com/gv/).[23] Hierarchical clustering analyses were performed using Robust Multi-array Average (RMA)-normalized data corresponding to 871 publicly available experiments (http://bioinf.mind.meiji.ac.jp/Rice_network_public).[24] The data corresponding to *SKP1 genes* were extracted (Supplementary Table S2) and analysed using the MeV 4_5_1 software (http://www.tm4.org/mev/). The *SKP1* genes were then clustered according to their expression profile using the Euclidean distance and complete linkage methods.

2.4. Total RNA extraction and isolation of OSKs and F-box cDNAs

Total RNA was isolated from 3-week-old leaf rice (*O. sativa* L. var. Nipponbare) following the method described by Bogorad et al.,[25] with slight adjustments for small samples(0.5 g). The RNA was then treated with DNaseI (Invitrogen). The SuperScript II reverse transcriptase and oligo-dT primer (Invitrogen) were used to synthesize the first strand cDNA. The forward and reverse primer pairs used to amplify the 13 OSK cDNAs are shown in Supplementary Table S3, as are the primer pairs used to amplify the 5 F-box cDNAs. Then, 1% agarose gel and the GFX Purification Kit (Amersham) were used to separate and purify the PCR products. Because we failed to amplify and then clone all the SKP1 and F-box coding sequences from the tissues available, 17 OSK and 4 F-box cDNA were chemically synthesized by Proteogenix (France). Both the synthesized and amplified cDNAs were then cloned and treated alike.

2.5. cDNA cloning

pENTR™/D-TOPO® cloning kits (Invitrogen) were used to clone the inserts. Under this system, PCR products are cloned directionally by adding four bases to the forward primer (CACC). Entry clones flanked with *att*L1 and *att*L2 sites necessary for recombination into the destination vectors are then obtained. All constructs were checked by sequencing, and a total of 30 OSK genes and 9 F-box genes were cloned and sequenced.

2.6. Binary yeast two-hybrid analysis

2.6.1. Plasmid constructs Gateway Cloning Technology (Invitrogen, Carlsbad, CA, USA) was used to produce hybrid proteins. The yeast expression vectors pDEST™32 and pDEST™22 were used to generate GAL 4 DNA-binding domain and GAL4 DNA activation domain fusion proteins. The LR reaction (Gateway Technology, Invitrogen) was performed to clone 30 OSK and 9 F-box in both directions into the destination vectors.

2.6.2. Yeast strain The MaV203 *Saccharomyces cerevisiae* strain (ProQuest™ Two-hybrid System, Invitrogen) contains deletions in its endogenous GAL4 and GAL80 genes for use with GAL4-based two-hybrid systems. This strain also has leu2 and trp1 mutations for the selection of ProQuest™ bait and prey vectors. MaV203 contains three GAL4 inducible reporter genes, providing four phenotypes, to enable the easier identification of true interactors.

2.6.3. Yeast two-hybrid screening and assays Bait and prey vectors were transformed together in the MaV203 yeast strain using the ProQuest™ Two-hybrid System (Invitrogen), according to the manufacturer's manual. Each co-transformation was plated onto a nutritionally selective plate that was deficient in tryptophan and leucine (Sc-Leu-Trp), to test for positive transformation. X-gal assays based on induction of the lacZ reporter gene were used to verify the interactions between fusion proteins. Constructs that gave a positive interaction were checked for self-activation according to the manufacturer's instructions (Invitrogen). The yeast cells were co-transformed with a bait or prey plasmid and an empty plasmid. Self-activation tests were performed using the HIS3 reporter gene encoding an enzyme involved in histidine biosynthesis. This enzyme can be inhibited by a competitive inhibitor, 3-aminotriazole (3AT). The addition of 3AT at concentrations of 10, 25, 50 or 100 mM in a histidine-free culture

medium increases the stringency of this test. The empty vectors and pEXP32/Krev1 and pEXP22/RalGDS-m2 vectors were used as negative controls, whereas yeast strains which were co-transformed with the pEXP32/krev1 and pEXP22/RalGDS-wt vectors were used as positive controls (Invitrogen).

2.7. Protein extraction and immunoblotting

To check for the actual production of fusion proteins, total yeast proteins were extracted according to the method described by Printen and Sprague[26] and separated by electrophoresis with the Mini-PROTEAN TGX (Tris-Glycine eXtended, BIO-RAD) on either precast 10% gels or 8% SDS-polyacrylamide gels. Protein samples were blotted onto a nitrocellulose membrane (Amersham Hybond™ ECL™, GE Healthcare Life Sciences) using a liquid transfer system (Criterion™ Blotter, BIO-RAD) in a transfer buffer (25 mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3). The fusion protein blots were probed with primary anti-GAL4-BD mouse monoclonal antibody (sc-510, Santa Cruz Biotech, Santa Cruz, CA, USA) and secondary goat anti-mouse IgG-HRP antibody (sc-2005, Santa Cruz Biotech, CA, USA) at dilutions of 1/500 and 1/10 000, respectively and visualized using chemiluminescence as instructed by the manufacturer (ECL, Amersham, Pharmacia). Each immunoblot assay was repeated at least twice.

3. Results

3.1. Structural analysis of rice OSK genes and other plant SKP1-like genes: SKP1-like gene intron identification

PLAZA (http://bioinformatics.psb.ugent.be/plaza/) was used to help in identifying the presence and position of introns within each *OSK* gene. In a first step, the rice *OSK* genes were placed in three different classes according to Kong et al.[17]: type Ia corresponding to *OSK* genes containing one intron, type Ib corresponding to intronless OSK genes and type II corresponding to *OSK* genes containing more than one intron. Details of the structure and a description of the rice *OSK* genes are given in Supplementary Table S1. In a second step, we retrieved and analysed 288 *SKP1-like* genes belonging to 17 species including the moss *Physcomitrella patens*, 5 monocots (*Brachypodium distachyon*, *O. sativa* ssp. Japonica, *O. sativa* ssp. Indica, *Sorghum bicolor* and *Zea mays*) and 11 eudicots (*Arabidopsis lyrata*, *A. thaliana*, *Carica papaya*, *Fragaria vesca*, *Glycine max*, *Lotus japonicus*, *Manihot esculenta*, *Medicago truncatula*, *Populus trichocarpa*, *Ricinus communis* and *Vitis vinifera*) (Supplementary Table S4). Genes belonging to the three classes (type Ia, type Ib and type II according to Kong et al.[17]) were found in each species. However, careful examination of nucleotide and protein alignments revealed that *SKP1-like* genes containing a single intron could be split into two subclasses. In the first subclass, the intron was in a conserved position, i.e. occurring between two aligning bases in the alignment of the coding sequences (Fig. 1).[22] In addition, whereas the first exon and the single intron were variable in length, the second exon was constant and was 174 bp (or 58 amino acid) long (Fig. 1). The second subclass of plant *SKP1-like* genes with a single intron contained *SKP1* genes, where both the intron and the two exons were of variable length. A survey of the whole set of plant and moss *SKP1-like* genes indicated that each species contained at least one gene with a conserved single intron (*A. lyrata*, *F. vesca* and *L. japonicus*). However, the majority of plant species contained two or three genes with a single conserved intron. When the different classes of *SKP1-like* genes were considered, it was clear that the majority of plant *SKP1-like* genes were intronless and represented 50% of the entire set (144 out of 288). The results of this survey are summarized in Supplementary Table S5. Phylogenetical analysis of the 40 plant *SKP1-like* genes with a conserved single intron indicated that these genes formed 3 main clusters corresponding to the moss, monocot and eudicot groups (Fig. 2). The grouping, thus, observed was mainly determined by the first exon, which was more variable between species than the second exon (Fig. 1). Interestingly, *SKP1-like* genes originating from the same species generally clustered. However, some genes originating from different species also clustered together, indicating possible orthology relationships. For example, in the monocots, SB05G012740 and SB02G031280 from sorghum clustered, respectively, with ZM04G05180 and ZM07G31190 from maize while RC28962 G00160 and RC30170G05850 from *Ricinus* clustered, respectively, with ME07520G02920 and ME02943G00470 in *M. esculenta*.

3.2. Expression profiles of the rice Skp1-like genes

Gene expression profiles can provide indications concerning gene functions. We looked for evidence of the expression of all *OSK* genes in the EST databases and found that EST sequences were available for 12 out of the 31 *OSK* genes. The number of EST also varied significantly in various tissues/organs, indicating the differential expression of *OSK* genes (Supplementary Table S6). Within the set of expressed OSK genes, OSK1 was the most strongly expressed and was represented by 43.6% of ESTs, whereas OSK20, OSK8, OSK22 and OSK28 were represented by 16, 12.7, 9 and 7% of ESTs, respectively. The cumulative percentage of these genes accounted for about

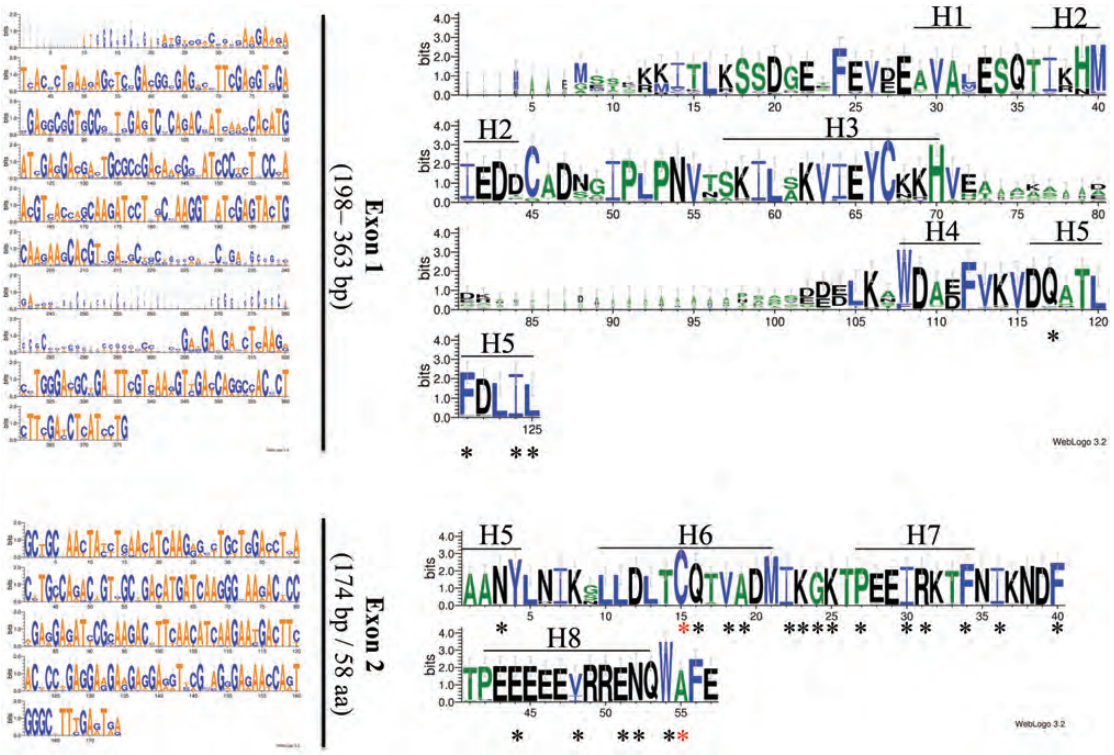

Figure 1. Multiple alignment of exons in plant *SKP1-like* genes generated with weblogo 3.2 (http://weblogo.threeplusone.com/manual.html). The sequence logo was derived from the alignment of 40 exons in plant *SKP1-like* gene with a conserved single intron, showing the high conservation of exons 2. The vertical axis shows the data content in bits. The overall height of the stack indicates sequence conservation at that position, whereas the height of the nucleotide within a stack indicates its relative frequency at that position. The eight helixes (H1 through H8) identified in human SKP1 are indicated by bars.

88.3% of all the ESTs detected (Supplementary Table S6).

In addition to this EST-based analysis of expression, we surveyed the expression profiles of OSK genes during various developmental stages using the Genevestigator platform (https://www.genevestigator.com/gv/).[24] The microarray data from 1475 arrays reflecting different developmental stages throughout the life cycle of rice were analysed. *OSK* genes displayed varied expression patterns with some *OSK* genes being strongly and widely expressed such as *OSK1*, *OSK8*, *OSK11*, *OSK20* and *OSK23*. However, the majority of *OSK* genes are expressed in inflorescence parts of the plants (Fig. 3).

In addition, RMA-normalized data from 871 microarrays were retrieved[27,28] and the Affymetrix probe sets corresponding to *OSK* genes enabled the identification of 20 non-ambiguous Probe set OSK pairs. The hierarchical clustering of 20 OSK genes based on their average log2 signal values in the 871 experiments revealed 3 different clusters. One cluster contained widely and strongly expressed genes including *OSK1*, *OSK8*, *OSK20* and *OSK24*, a second cluster contained moderately expressed genes and the third cluster contained weakly expressed genes (Fig. 3; Supplementary Table S2). Overall, the expression data indicated that *OSK1* is the most widely and strongly expressed rice SKP1 gene.

3.3. *Yeast two-hybrid interaction between OSKs, F-box proteins and western blots*

An experiment was designed to test interactions between 30 rice OSK and 9 F-box proteins. For each OSK-F-box pair, the interaction was tested in both directions, resulting in a total of 540 binary interactions.

Prior to the actual interaction tests, each protein was examined individually for transactivation (activation of reporter gene transcription), and we observed transactivation for OSK4, OSK7, OSK12, OSK15 and OSK31 in yeast when fused to the Gal4 BD domain (i.e. the pDEST^{TM}32 expression vector). No transactivation was seen with these proteins when they were fused to the AD domain (i.e. the pDEST^{TM}22 expression vector).

Of the 30 OSK proteins examined, 8 were shown to interact with at least 1 F-box protein in yeast (Table 1). Some OSK proteins interacted preferentially with certain subfamilies of F-box proteins. For example, OSK4, OSK12 and OSK16 interacted with an F-box protein (02g0260200) belonging to the

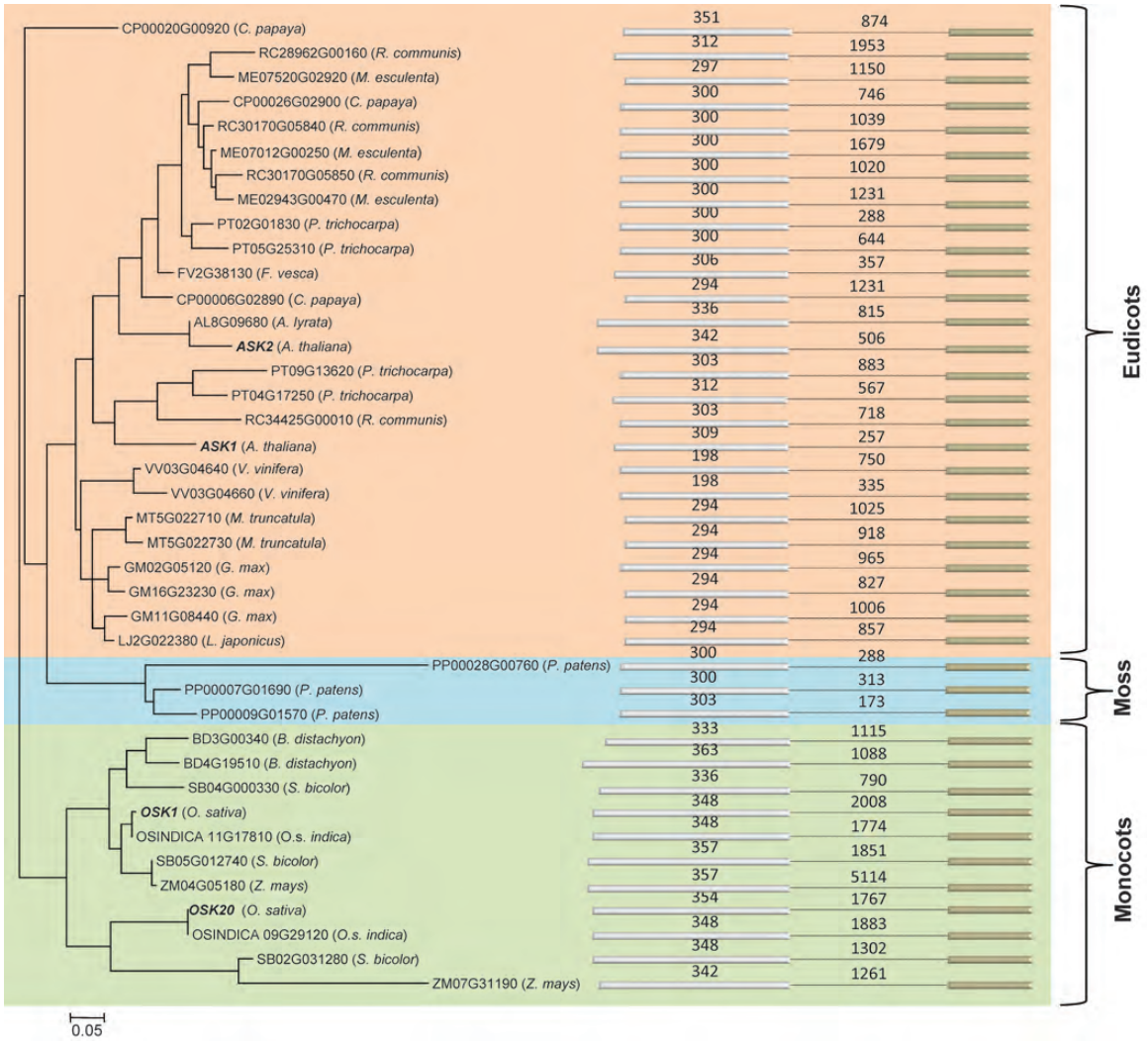

Figure 2. Molecular phylogenetic analysis using the maximum likelihood method. A phylogenetic analysis of 40 plant *SKP1-like* genes with a conserved single intron indicated that these genes form 3 main clusters corresponding to moss, monocot and eudicot groups. Their evolutionary history was inferred using the maximum likelihood method based on the Poisson correction model. The tree with the highest log likelihood (−3400.1299) is shown. This tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 32 amino acid sequences. There were a total of 213 positions in the final dataset. Evolutionary analyses were conducted under MEGA5.[20]

FBX subfamily, but did not interact with members of the other groups (Fig. 4). In the case of OSK5, OSK6, OSK9 and OSK30, the expression was very low and below the detectable level by protein blotting and hence it was indicated as 'not determined' for combinations involving these proteins (Table 1). On the other hand, the same F-box protein (02g0260200) interacted with other members of the OSK family such as OSK1, OSK15 and OSK20. It appears, therefore, that the majority of interactions were observed with the FBX group (Fig. 4). In contrast, the FBL2 (Os05g0425700) F-box member interacted with only two OSK proteins: OSK1 and OSK20. On the other hand, OSK1 and OSK20 were found to interact with all the F-box proteins tested and belonging to three different subfamilies: FBX, FBL and FBK (Fig. 4). In addition, the F-box protein Gibberellin-insensitive dwarf2 (GID2) , which is essential for GA-mediated DELLA protein degradation, interacted with OSK1, OSK13, OSK20 and OSK25 (Table 1; Fig. 4).

Concerning the F-box proteins, four belonging to the FBA, FBD, FBT and FBDUF subfamilies did not display any detectable interaction under our experimental conditions, even after 24 h of incubation at 37°C during the X-gal assays (Table 1). To verify that the negative interactions were not merely due to a lack of protein production by the yeast, we examined the expression of OSK and F-box proteins fused to the Gal4 DNA-binding domain in yeast by western blotting with the Gal4 monoclonal antibody. We observed the expression of fusion proteins in 26 OSK (Fig. 5). However, no fusion proteins for OSK5, OSK6, OSK9

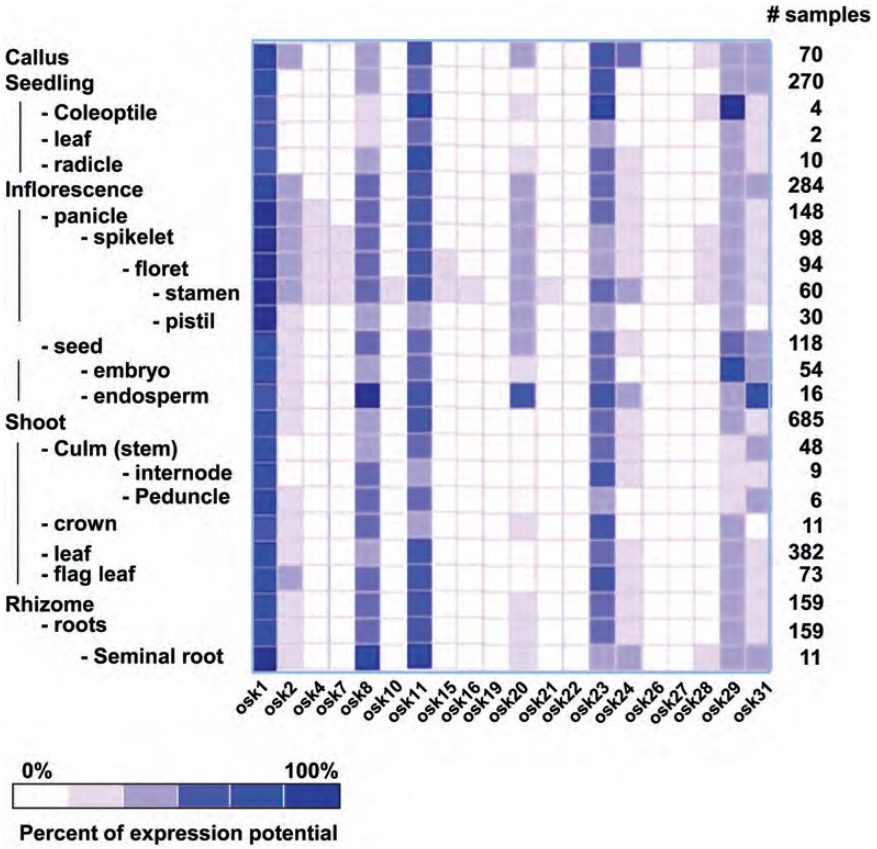

Figure 3. Meta-analysis of the expression of rice *SKP1-like* genes**.** (A) Anatomical analysis of *OSK* gene expression using the Genevistegator platform (https://www.genevestigator.com/gv/). For a given gene and developmental stage, the level of expression is shown as a percentage of the maximum among all the microarray results found for this gene. The *OSK* genes display various patterns of expression. Most *OSK* genes are expressed in the inflorescence parts. (B) Expression profile of *OSK* genes. The colour scale (representing log signal values) is shown on the *y*-axis and the cluster tree at the top of the *x*-axis. The differential expression of 20 *OSK* genes is grouped into 3 different clusters (cluster 1: widely and strongly expressed genes, cluster 2: moderately expressed genes and cluster 3: weakly expressed genes) using Euclidean distance and the complete linkage method.

or OSK30 were detected. For all F-box fusion proteins, the presence of a fusion protein in yeast cells was confirmed by western blotting (Fig. 5). Therefore, we considered that the negative interactions observed with some of the OSK and F-box proteins reflected the genuine behaviour of these proteins in Y2H, rather than a technical artefact due to the absence of interactors.

4. Discussion

Comprehensive analyses of the SKP1 protein family in *C. elegans* demonstrated that the 21 *SKP1-related* genes displayed various interaction capabilities with different F-box proteins, with SKR1 being the most prone to interact *in vitro*.[16] In plants, similar comprehensive analyses performed to date have only concerned the model plant *A. thaliana*.[8,13,18] In contrast, although it is clear that all sequenced plant genomes contain several *SKP1-like* genes, no systematic analysis of their binding capabilities has been performed so far. We, therefore, decided to carry out a comprehensive study in rice, a monocot model.

4.1. *OSK1 and OSK20 belong to a class of SKP1 genes that contain one intron in a conserved position*

Phylogenetic studies have suggested that all *Arabidopsis* and rice *SKP1* genes derived from a single ancestral gene.[14] This hypothesis was supported by the findings of Kong et al.[17] who further suggested that *SKP1* genes could be classified as either type Ia, with one intron and two exons, type Ib that are intronless and type II that contain several introns. They suggested that type Ib could have derived from type Ia by retroposition. Their study also stated that the position of the single intron could vary between *SKP1* genes, but some of them have a conserved position. Interestingly, Kong et al.[17] showed that *ASK* and *OSK* genes with a conserved intron occupy basal positions in phylogenetic trees, whereas intronless genes often form terminal clades. To extend their study, we examined

Table 1. Yeast two-hybrid interaction between 30 OSK proteins and 9 F-box proteins

OSK proteins		F-box proteins								
		GID2	FBA	FBX	FBL2	FBL1	FBK	FBD	FBT	FBDUF
Accession	Name	AB100246	Os09g0479100	Os02g0260200	Os05g0425700	AAV32196	Os07g47650	Os08g35960	Os01g48370	Os01g08830
Os11g26910	OSK1	**+**	−	+	+	−	+	−	−	−
Os10g30200	OSK2	−	−	−	−	−	−	−	−	−
Os02g13180	OSK3	−	−	−	−	−	−	−	−	−
Os09g10200	OSK4	−	−	**+**	−	−	−	−	−	−
Os09g10260	OSK5	ND	ND	ND	ND	ND	ND	ND	ND	ND
Os07g05180	OSK6	ND	ND	ND	ND	ND	ND	ND	ND	ND
Os09g10300	OSK7	−	−	−	−	−	−	−	−	−
Os11g48030	OSK8	−	−	−	−	−	−	−	−	−
Os07g05150	OSK9	ND	ND	ND	ND	ND	ND	ND	ND	ND
Os06g02360	OSK10	−	−	−	−	−	−	−	−	−
Os06g02350	OSK11	−	−	−	−	−	−	−	−	−
Os09g10270	OSK12	−	−	**+**	−	−	−	−	−	−
Os09g10230	OSK13	**+**	−	+	−	−	−	−	−	−
Os02g01160	OSK14	−	−	−	−	−	−	−	−	−
Os08g28820	OSK15	−	−	**+**	−	−	−	−	−	−
Os07g05160	OSK16	−	−	+	−	−	**+**	−	−	−
Os07g43180	OSK17	−	−	−	−	−	−	−	−	−
Os07g43260	OSK19	−	−	−	−	−	−	−	−	−
Os09g36830	OSK20	+	−	+	+	−	+	−	−	−
Os07g22680	OSK21	−	−	−	−	−	−	−	−	−
Os07g43250	OSK22	−	−	−	−	−	−	−	−	−
Os07g43270	OSK23	−	−	−	−	−	−	−	−	−
Os12g40300	OSK24	−	−	−	−	−	−	−	−	−
Os08g28800	OSK25	**+**	−	+	−	−	−	−	−	−
Os07g43220	OSK26	−	−	−	−	−	−	−	−	−
Os07g43230	OSK27	−	−	−	−	−	−	−	−	−
Os07g43240	OSK28	−	−	−	−	−	−	−	−	−
Os08g28780	OSK29	−	−	−	−	−	−	−	−	−
Os09g10020	OSK30	ND	ND	ND	ND	ND	ND	ND	ND	ND
Os03g01660	OSK31	−	−	−	−	−	−	−	−	−

Summary of the results of the 540 interactions. Plus (+) and minus (−) indicate positive and negative interaction, respectively. Italics (+) indicates that interaction was detected in both combinations. Bold (**+**) indicates that interaction was detected in only one combination. ND, not determined.

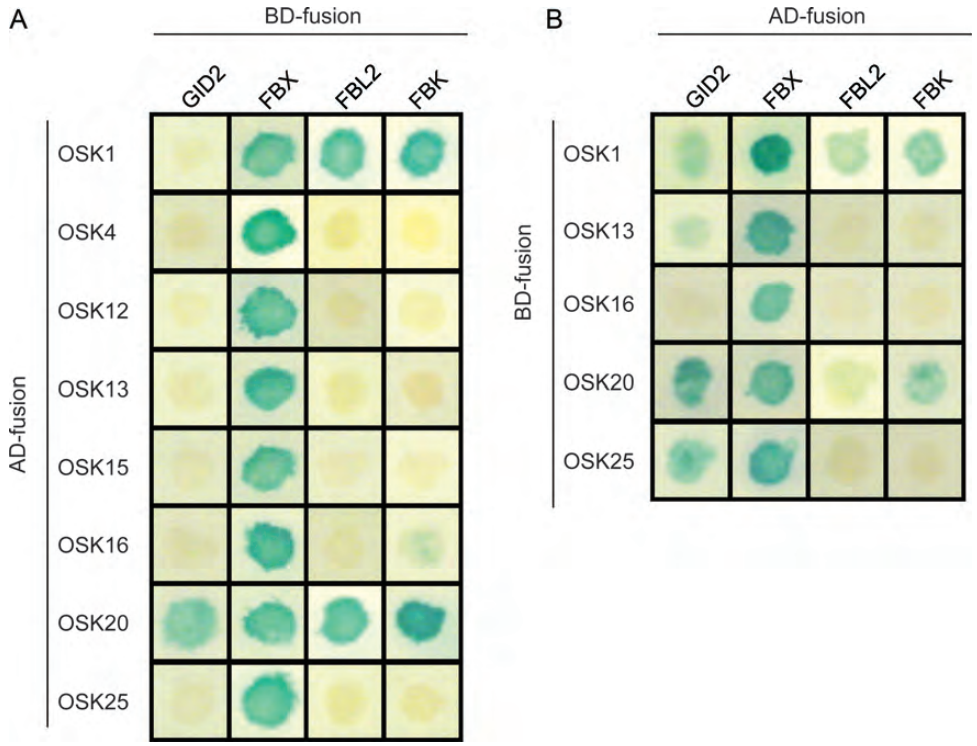

Figure 4. Interactions of eight rice OSK proteins with representative F-box proteins were determined using yeast two-hybrid analysis. To generate prey (AD-fusion) and bait (BD-fusion) constructs, we fused full-length OSK and F-box proteins with the GAL4 activation domain and GAL4 DNA binding-domain, respectively. The constructs were tested with both combinations. Positive AD-X and DB-X interacting clones were examined for self-activation. Each co-transformation was plated onto a nutritionally selective plate. X-gal assays was used. Strong interacting clones and weak interacting clones produced a dark blue and a light blue colour, respectively, whereas non-interacting clones remained white. We observed self-activation for OSK4, OSK12 and OSK15 in bait (BD-fusion) constructs. Most interactions were observed with the FBX (02g0260200) F-box protein with both combinations.

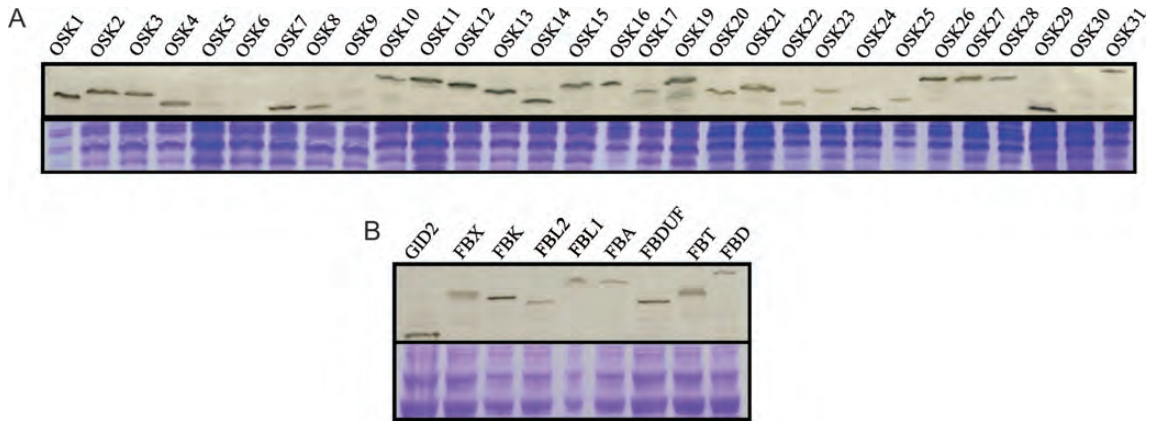

Figure 5. Yeast two-hybrid interactions between 30 OSK proteins and 9 F-box proteins. Immunoblot analysis of OSK and F-box proteins fused to the Gal4 DNA-binding domain. We observed the expression of fusion proteins for 26 OSK and all the F-box proteins tested.

the genomic organization of *SKP1* genes in the moss *P. patens* and in 16 monocot and eudicot genomes and defined 4 subclasses of *SKP1* genes, among which 1 subclass corresponds to *SKP1* genes with 1 intron at a conserved position and contains the rice OSK1 and OSK20, suggesting orthology relationships between these genes. By means of a detailed analysis of certain animal, insect and plant genomes, Henricson et al. suggested that genuine orthologous genes tend to have more conserved intron positions than non-orthologous genes.[29] In contrast, the majority (140/288) of the plant *SKP1* genes were intronless, suggesting an active retroposition phenomenon.

4.2. *OSK1 and OSK20 belong to a subclass of highly expressed SKP1 genes*

Our EST survey indicated that *OSK1* and *OSK20* are the most widely represented genes in public EST databases, (Supplementary Table S6). These results agreed with Kong et al.'s general finding that animal and plant members of this gene family are not uniformly represented by ESTs in the databases, indicating varied degrees of expression.[14] In particular, they showed that 63 (96.9%) of the 65 significantly slowly evolving members have ESTs, whereas 89 (93.7%) of the 95 moderately and rapidly evolving members of animals and plant species did not have ESTs.[14] During the present study, our comprehensive comparison of the large collection of microarray data in rice (i.e. Genevestigator platform[23]) clearly indicated that rice *SKP1* genes exhibited an expression profile that was heterogeneous in terms of tissues, conditions and overall intensity. This was in complete agreement with the reports by Takahashi et al. and Kong et al. in *Arabidopsis*.[14,18] Moreover, in all species, where sufficiently large and accessible data were available (*A. thaliana*, rice and *P. patens*), *SKP1-like* genes containing single conserved intron (*ASK1* and *ASK2* in *A. thaliana*, *OSK1* and *OSK20* in rice; PP00028G00760, PP00007G01690 and PP00009G01570 in *P. patens*) exhibited broad and strong levels of expression when compared with the other *SKP1-like* genes, suggesting their involvement in various physiological processes, even though other members of the *SKP1* gene family (intronless or containing non-conserved introns) could also be strongly expressed under some conditions.

4.3. *Yeast two-hybrid analysis: OSK proteins interact specifically with F-box proteins*

The yeast two-hybrid matrix method involved a subset of 9 F-box and 30 OSK proteins and uncovered selective and distinctive interactions between the OSK proteins. This result strongly suggests that a combination of OSK proteins with F-box proteins produces diverse substrate selectivity. This has already been shown in many studies.[8,16,18] It was pointed out that in *C.s elegans*, SKP1 proteins (SKR) displayed diverse patterns of interaction with F-box proteins.[16] In *A. thaliana*, some ASK proteins (e.g. ASK1, ASK2, ASK11 and ASK12) were able to interact with a broad spectrum of F-box proteins, whereas others only interacted with a few types of proteins.[8,13,18] However, some ASK proteins cannot interact with any F-box proteins.[18] Likewise, here, we detected no interaction between 22 of the OSK proteins and F-box proteins tested, which could indicate that these interactions may require additional partners and/or the Y2H technique is not appropriate for the detection of interactions between these proteins. Alternatively, these F-box proteins could be involved in protein complexes other than the SCF, as suggested by Hermand.[10] The presence of a large number of F-box proteins in *O. sativa* implies that a plethora of SCF complexes (which recognize a broad array of substrates) are possible. Yeast two-hybrid analysis showed that several putative F-box proteins interacted with one or more members of the *A. thaliana* SKP1-like (ASK) family.[9] In fact, our study revealed that OSK1 is not the only one partner of GID2. The Gomi studies reported that GID2 preferentially interacted with OsSKP15 (OSK1).[30] Our results may, therefore, indicate that GID2 is a component of the SCF complex, not only through its interaction with OSK1 but also through its interaction with other OSK members (OSK13, OSK20 and OSK25). Interestingly, we observed that the majority of interactions were observed with the (Os02g 0260200 = LOC_Os02g15950) F-box protein that belongs to the FBX subgroup. This gene is homologous to—and belongs to—the same orthology group as *Arabidopsis* gene At3g61590 that has been shown to interact with all, but two of the ASK tested.[18] This result may suggest conservation in these genes between rice and *Arabidopsis* at the level of interaction capabilities.

4.4. *Comparison of key amino acid sequences between* Oryza *OSK proteins and the human Skp1 protein*

Apart from some specific interactions, we noted that the majority of the positive interactions observed involved OSK1 or OSK20. In this respect, they resembled Skr1 and Skr2 in *C. elegans*,[16] ASK1 and ASK2 and possibly ASK11 that have been found to be implicated in most positive interactions with F-box proteins in *A. thaliana*.[8,18] It has been suggested that these non-specific interacting SKP1 proteins should be named the 'master components' of SCF complexes.[8] Kong et al.[14] and Risseeuw et al.[8] also noted that these proteins shared the majority of the 26 amino acids that have been shown to contact the human SKP2 F-box protein.[31,32] In this study, we found that OSK1 and OSK20 shared, respectively, 23 and 24 of the 26 key amino acids with Hs-SKP1 (Supplementary Fig. S1).

4.5. *Conclusion*

The plant SKP1 genes could be split into different subfamilies, among which one contains presumably ancestral members from which the other *SKP1-like* genes have evolved within each species.[17] From a compilation of publicly available data, results from previous studies and the Y2H results described in this study, we can suggest that: (i) typical members of this subclass contain two exons, the first exon

being of variable length and the second exon being 174 bp (58 amino acids) long, with one intron in a conserved position, (ii) members of this subclass are widely and strongly expressed throughout the whole plant and (iii) they are able to interact with the majority of F-box proteins and may, therefore, participate in the formation of most of the SCF complexes found in plant cells. Based on these criteria, we suggest that rice *OSK1* and *OSK20* are likely to be functionally equivalent to *Arabidopsis ASK1* and *ASK2* genes, and we propose functional orthologues for 15 other plants (Supplementary Table S4).

Acknowledgements: The authors would like to thank Monique Sibaud for her technical assistance, Dominique Marcon for images of the Y2H interactions and Nelly Lajoinie for growing the plants. We would like to thank K. Gomi for providing us with the unpublished OsSKP sequences for comparison.

Supplementary data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was supported by a grant from the Tunisian Ministry of Higher Education and Research for a Doctoral Scholarship concerning the first author.

References

1. Gieffers, C., Schleiffer, A. and Peters, J.M. 2000, Cullins and cell cycle control, *Protoplasma*, **211**, 20–8.
2. Vierstra, R.D. 2003, The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins, *Trends Plant Sci.*, **8**, 135–42.
3. Vierstra, R.D. 2009, The ubiquitin-26S proteasome system at the nexus of plant biology, *Nat. Rev. Mol. Cell. Biol.*, **10**, 385–97.
4. Hua, Z. and Vierstra, R.D. 2011, The Cullin-RING Ubiquitin-Protein Ligases, *Annu. Rev. Plant Biol.*, **62**, 299–334.
5. Mazzucotelli, E., Belloni, S., Marone, D., et al. 2006, The E3 ubiquitin ligase gene family in plants: regulation by degradation, *Curr. Genomics*, **7**, 509–22.
6. Du, Z., Zhou, X., Li, L. and Su, Z. 2009, Plants UPS: a database of plants' Ubiquitin Proteasome System, *BMC Genomics*, **10**, 227.
7. Jain, M., Nijhawan, A., Arora, R., 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**, 1467–83.
8. Risseeuw, E.P., Daskalchuk, T.E., Banks, T.W., et al. 2003, Protein interaction analysis of SCF ubiquitin E3 ligase subunits from Arabidopsis, *Plant J.*, **34**, 753–67.
9. Kuroda, H., Takahashi, N., Shimada, H., Seki, M., Shinozaki, K. and Matsui, M. 2002, Classification and expression analysis of Arabidopsis F-box-containing protein genes, *Plant Cell Physiol.*, **43**, 1073–85.
10. Hermand, D. 2006, F-box proteins: more than baits for the SCF?, *Cell Division*, **13**, 1–30.
11. Jin, J., Cardozo, T., Lovering, R.C., Elledge, S.J., Pagano, M. and Harper, J.W. 2004, Systematic analysis and nomenclature of mammalian F-box proteins, *Genes Dev.*, **18**, 2573–80.
12. Xu, G., Ma, H., Nei, M. and Kong, H. 2009, Evolution of F-box genes in plants: different modes of sequence divergence and their relationships with functional diversification, *Proc. Nat. Acad. Sci. USA.*, **106**, 835–40.
13. Gagne, J.M., Downes, B.P., Shiu, S.H., Durski, A.M. and Vierstra, R.D. 2002, The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis, *Proc. Nat. Acad. Sci. USA.*, **99**, 11519–24.
14. Kong, H., Leebens-Mack, J., Ni, W., de Pamphilis, C.W. and Ma, H. 2004, Highly heterogeneous rates of evolution in the *SKP1* gene family in plants and animals: functional and evolutionary implications, *Mol. Biol. Evol.*, **21**, 117–28.
15. Nayak, S., Santiago, F.E., Jin, H., Lin, D., Schedl, T. and Kipreos, E.T. 2002, The *Caenorhabditis elegans* Skp1-related gene family: diverse functions in cell proliferation, morphogenesis, and meiosis, *Curr. Biol.*, **12**, 277–87.
16. Yamanaka, A., Yada, M., Imaki, H., Koga, M., Ohshima, Y. and Nakayama, K. 2002, Multiple Skp1-related proteins in *Caenorhabditis elegans*: diverse patterns of interaction with Cullins and F-box proteins, *Curr. Biol.*, **12**, 267–75.
17. Kong, H., Landherr, L.L., Frohlich, M.W., Leebens-Mack, J., Ma, H. and de Pamphilis, C.W. 2007, Patterns of gene duplication in the plant SKP1 gene family in angiosperms: evidence for multiple mechanisms of rapid gene birth, *Plant J.*, **50**, 873–85.
18. Takahashi, N., Kuroda, H., Kuromori, T., et al. 2004, Expression and interaction analysis of *Arabidopsis Skp1*-related genes, *Plant Cell Physiol.*, **45**, 83–91.
19. Quevillon, E., Silventoinen, V., Pillai, S., et al. 2005, InterProScan: protein domains identifier, *Nucleic Acids Res.*, **33**, 116–20.
20. Edgar, R.C. 2004, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.*, **321**, 792–1797.
21. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony, *Methods Mol. Biol. Evol.*, **28**, 2731–9.
22. Zheng, J., Rogozin, I.B., Koonin, E.V. and Przytycka, T.M. 2007, Support for the Coelomata clade of animals from a rigorous analysis of the pattern of intron conservation, *Mol. Biol. Evol.*, **24**, 2583–92.
23. Hruz, T., Laule, O., Szabo, G., et al. 2008, Genevestigator v3: a reference expression database for the meta-

analysis of transcriptomes, *Adv. Bioinformatics*, **2008**, 420747.
24. Kazuki, H., Kohei, H., Keita, S., et al. 2011, Oryza Express: an integrated database of gene expression networks and omics annotations in rice, *Plant Cell Physiol.*, **52**, 220–9.
25. Bogorad, L., Gubbins, E.J., Krebbers, E., et al. 1983, Cloning and physical mapping of maize plastid genes, *Methods Enzymol.*, **97**, 524–54.
26. Printen, J.A. and Sprague, G.F. Jr. 1994, Protein-protein interactions in the yeast pheromone response pathway: Ste5p interacts with all members of the MAP kinase cascade, *Genetics*, **138**, 609–19.
27. Wu, Z., Irizarry, R.A., Gentleman, R., Martinez-Murillo, F. and Spencer, F. 2004, A model-based background adjustment for oligonucleotide expression arrays, *J. Am. Stat. Assoc.*, **99**, 909–17.
28. Irizarry, R.A., Wu, Z. and Jaffee, H.A. 2006, Comparison of Affymetrix GeneChip expression measures, *Bioinformatics*, **22**, 789–94.
29. Henricson, A., Forslund, K. and Sonnhamme, E. 2010, Orthology confers intron position conservation, *BMC Genomics*, **11**, 412.
30. Gomi, K., Sasaki, A., Itoh, H., et al. 2004, GID2, an F-box subunit of the SCF e3 complex, specifically interacts with phosphorylated SLR1 protein and regulates the gibberellin-dependent degradation of SLR1 in rice, *Plant J.*, **37**, 626–34.
31. Schulman, B.A., Carrano, A.C. and Jeffrey, P.D. 2000, Insights into SCF ubiquitin ligases from the structure of the Skp1–Skp2 complex, *Nature*, **408**, 381–6.
32. Zheng, N., Schulman, B.A., Song, L., et al. 2002, Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex, *Nature*, **416**, 703–9.