Plant Molecular Biology **52:** 579–590, 2003.

OsBP-73, a rice gene, encodes a novel DNA-binding protein with a SAP-like domain and its genetic interference by double-stranded RNA inhibits rice growth

Jun Chen, Wei-Hua Tang[1], Meng-Min Hong and Zong-Yang Wang*
*National Laboratory of Plant Molecular Genetics, Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, China (*author for correspondence, e-mail zywang@iris.sipp.ac.cn); [1]current address: Plant Gene Expression Center, USDA/ARS-UC-Berkeley, 800 Buchanan Street, Albany, CA 94710, USA*

Received 21 November 2002; accepted in revised form 26 January 2003

Key words: DNA-binding protein, rice, RNAi, SAP domain, *Wx* gene, yeast one-hybrid system

Abstract

The SAP domain is a recently defined DNA binding domain that forms a helix-extended-helix structure. SAP proteins have been implicated in nuclear architecture and/or RNA metabolism. In this paper, we describe the cloning and characterization of a rice gene, *OsBP-73*, encoding a 375 amino acid protein with a SAP-like domain. We identified the binding sequence of OsBP-73 by gel retardation assays and southwestern blotting. Northern blot analysis demonstrated that *OsBP-73* is weakly expressed in root, leaf and immature seed. *OsBP-73* gene expression was also examined by histochemical studies of transgenic rice plants carrying an *OsBP-73* 5′/GUS reporter gene. The reporter gene is mainly expressed in the tissues with high cell division activities, such as root tip, stem node, panicle and immature seed. Genetic interference of *OsBP-73* gene expression by double-stranded RNA strikingly inhibits the whole plant growth but does not affect the passage from the juvenile to adult phase. These results suggest that *OsBP-73* may play an important role in the regulation of cell proliferation.

Introduction

It is becoming increasingly clear that DNA-protein interactions play central roles in a wide range of cellular processes. Examples include transcriptional regulation, chromosome organization, DNA replication and DNA repair (Guille and Kneale, 1997). DNA-protein interactions are mediated by various DNA binding domains or modules. The most familiar modules are the DNA binding domains of transcription factors (Meshi and Iwabuchi, 1995). Some DNA-binding domains are proposed to organize the chromatin structure and be important for regulating gene expression.

The SAP domain (after SAF-A/B, Acinus and PIAS) is a recently defined DNA-binding domain (Aravind and Koonin, 2000; Kipp *et al*., 2000b). Proteins with a SAP domain (also named SAF box) have been identified in yeast, mammals and plants (Aravind and Koonin, 2000; Kipp *et al*., 2000b). The SAP domain is composed of 35 amino acids residues and comprises two amphipathic helices separated by a glycine-containing region. Some positions in this domain are enriched in positively charged amino acids (R, K) which are thought to contact the DNA backbone. Recent studies showed that the SAP domain forms a helix-extended-helix (HEH) structure and that some prokaryotic proteins, such as transcription terminator RHO protein, are also predicted to contain SAP domains (Aravind and Koonin, 2001).

SAP domain-containing proteins have been implicated in various functions related to their interactions with DNA and/or RNA. For example, human scaffold attachment factor A (SAF-A) is an abundant component of the nuclear scaffold (nuclear matrix) and is also present in heterogeneous nuclear ribonucleoprotein complexes, which have been implicated in nuclear

organization and RNA processing (Fackelmayer *et al*., 1994; Nayler *et al*., 1998). Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation (Sahara *et al*., 1999). Members of the PIAS proteins family combining SAP domains and MIZ Zn-finger motifs are the protein inhibitors of activated STATs (signal transducer and activator of transcription) (Chung *et al*., 1997; Liu *et al*., 1998). Yeast Tho1p protein, another SAP-containing protein, plays a role in regulating elongation of transcription by RNA polymerase II (Piruat and Aguilera, 1998). Plant PARP, for poly(ADP-ribose) polymerase, is a nuclear enzyme activated by DNA breaks and catalyzing poly-ADP ribosylation of chromosomal proteins. Many studies on plant PARP have indicated it has roles in DNA repair and DNA replication (Lepiniec *et al*., 1995; Amor *et al*., 1998). Recent studies suggest that plant PARP also plays a role in transcriptional regulation (Storozhenko *et al*., 2001; Chabeaud *et al*., 2001). Although the functions of many SAP domain-containing proteins are clear, the characteristics of the SAP domains in these proteins have been scarcely investigated with the exception of SAF-A. Sequence blast and structure prediction revealed that there are many other SAP proteins both in eukaryotic and prokaryotic cells (Aravind and Koonin, 2000, 2001; Kipp *et al*., 2000b). However, the function of most of these proteins is not known at present because they have been identified through genome sequencing projects rather than through biochemical or biological activities (Kipp *et al*., 2000b). Much work is required to elucidate the biological function of these SAP proteins and their SAP domains.

The rice *Wx* gene encodes a starch granule-bound starch synthase, which is responsible for synthesis of amylose in the rice endosperm, pollen and embryo sac (Wang *et al*., 1995). The Rice *Wx* gene is spatially and temporally expressed and may be regulated at the transcriptional and post-transcriptional level (Hirano *et al*., 1995; Wang *et al*., 1995). In order to investigate the transcriptional regulation of the rice *Wx* gene, we demonstrated by gel retardation assays that one or more proteins in nuclear extracts of immature rice seeds interact with a 31 bp DNA sequence in the promoter region of the rice *Wx* gene (Chen *et al*., 1997). Further experiments with transgenic lines showed that the *Wx* promoter containing the 31 bp sequence enhances *GUS* reporter gene expression 2–3 times as high as the *Wx* promoter without this 31 bp sequence (Ge *et al*., 2000). We suggest that the 31 bp sequence in the upstream of the rice *Wx* gene contains a key *cis* element. To identify transcription factors that bind to this sequence, we performed a yeast one-hybrid screen, with the 31 bp DNA sequence as bait. Three groups of cDNA clones were isolated from a rice cDNA library (Chen *et al*., 1999). One is*OsEBP-89*, encoding an Ap2/EREBP protein which binds to the GCC-like box in the 31 bp sequence (Yang *et al*., 2002). Another is *OsBP-5*, encoding a bHLH protein which binds to a CAACGTG box in the 31 bp sequence (unpublished). Here we report that one group of the cDNA clones encodes OsBP-73, a protein with a SAP-like domain. We used a gel retardation assay and southwestern blotting to demonstrate that OsBP-73 can specifically bind to the 5′-ACGCACGCTAACGTGA-3′ region within the 31 bp DNA sequence. *OsBP-73* is ubiquitously expressed in the tissues we examined. Using double-stranded RNA interference, we observed a dramatic decrease of whole plant growth correlated with a decrease in the levels of *OsBP-73* mRNA.

Materials and methods

Plant material

Oryza sativa ssp. *japonica* cv. Zhonghua 11 was grown under greenhouse conditions.

Isolation and sequencing of OsBP-73 *cDNA and genomic clones*

The rice immature seed cDNA library was constructed previously (Wang *et al*., 1995), and the rice genomic library was purchased from Clontech. About 2×10^5 plaques were hybridized with a ^{32}P-labeled c73 probe as described by Maniatis *et al.* (1982). Three and two positive plaques were purified. DNA fragments were cloned into phagemid vector pUC18 (Yanisch-Perron *et al*., 1985), and both DNA strands were sequenced by the dideoxy chain-termination method (Sanger *et al*., 1977) with ABI733 Sequencers (Genecore Biological Company). Sequence analysis was performed with the software (version 7.0) developed by the Wisconsin Genetics Computer Group. Databases were searched by the BLAST network service of the National Institutes of Health, USA. Secondary structures of the protein was predicted by PCGENE5.0. The presence of domains in protein was predicted by ExPASy Molecular Biology Server and at the website of www.wustl.edu.

RT-PCR

First-strand cDNAs were synthesized with Oligo(dT) from 2.5–5 μg total RNA in rice immature seed, using Superscript II RNase H$^-$ reverse transcriptase (BRL), following the manufacturer's instructions. PCR reactions were carried out with three forward primers UP1, UP2, UP3 and one reverse primer 3ORF. Their sequences were indicated by arrow in Figure 1A.

Protein expression in Escherichia coli *and purification of the series deletion proteins*

The 724 bp *Sal*I-*Not*I restriction fragment containing the c73 cDNA was inserted in frame into vector pET28-C(+) (Novagen) and obtained plasmid pEB73. pEB73 was digested with *Sac*I and *Sal*I, then a series of 5′ deletion plasmids was generated by *Exo*III (Maniatis *et al*., 1982). Five deletion plasmids, pEB-N1, pEB-N2, pEB-N3, pEB-N4 and pEB-N5 were obtained. One 3′ deletion plasmid, pEB-C1, was obtained by deleting the *Xho*I fragment of pEB73. Plasmid pEB-C1 lost the stop codon TGA at +3801 and uses the translation stop codon of vector pET28C(+). The 3′-untranslated region of c73 cDNA was amplified with primers ch3z and ch5f whose sequences are indicated by an arrow in Figure 1A. The amplification product was then cloned into the *Xho*I site of pEB-C1 and produced the other 3′ deletion plasmid, pEB-C2, which contains the 3′-untranslated region of *OsBP-73* and use the stop codon TGA at +3801.

pEB73 and all deletion plasmids were expressed in *E. coli* strain BL21 (DE3). The resulting His-tagged fusion proteins were extracted and purified using Ni-NTA resin according to methods described previously (Tang *et al*., 1999).

Gel retardation assays and competition experiments

Both strands of the 31 bp nucleotide sequence, the left (L fragment) and right part (R fragment), of the 31 bp nucleotide sequence and the seven mutants of the R fragment were synthesized, annealed and inserted into vector pUC18. An insertion containing two tandem copies of the 31 bp nucleotide sequence and insertions containing one copy of L, R or mutated R DNA fragments respectively were chosen for use in the gel retardation assays. The 31 bp nucleotide sequence-containing DNA fragment was digested with *Eco*RI and *Bam*HI and labeled by using Klenow fragment of DNA polymerase I to fill in with ^{32}P-dATP (Chen *et al*., 1997). A 10 fmol portion of labeled

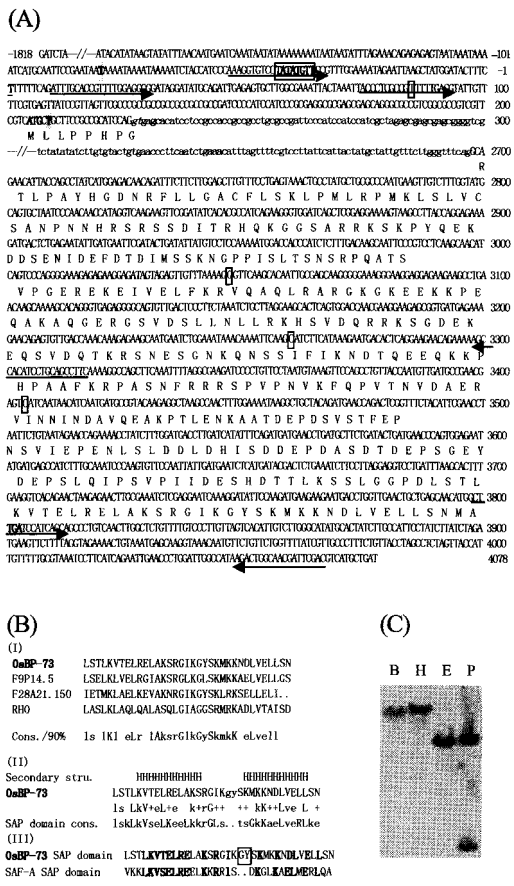

Figure 1. Characterization of *OsBP-73* gene. A. Nucleotide and deduced amino acid sequence of *OsBP-73* gene. The putative TATA box at position −39 is boxed and in bold letter. The putative transcription start site is underlined and in bold. The splicing site gt/ag is in italics. The first bases of c332 (T_{85}), c33 (G_{576}), c31 (C_{790}) and c73 (C_{933}) cDNA are boxed. The first base of *OsBP-73* gene in pasmids p13-D16 (T_{-80}) and p13-D6 (T_{+209}) are shaded and in bold letters. The PCR primers UP1 (5′-TACCCTCGCCGTTTTTGACG-3′), UP2 (5′-ATTTGCACCGTTTTGGAGGG-3′), UP3 (5′-AAAGGTGTC-CTATATGTTCCG-3′), 3ORF (5′-GCCACATCCTGCAGCCT-TC-3′), ch5z (5′-CTTGATCCATCAGCAG-3′) and ch3f (5′-AGACTGGCAACGATTCGA-3′) are indicated by arrow. B. Structure prediction for OsBP-73 protein. (I) Comparison of the C-terminal amino acids of the OsBP-73 protein with other proteins. The protein sequences involved in the comparison are from *Arabidopsis*, F9P14.5 (AC025290.3), F28A21.150 (T04866) and *Micrococcus luteus*, RHO (P52154). Consensus amino acid residues are shown and those perfectly conserved amino acid residues are shown in capitals. (II) A SAP domain was predicted at the C-terminus of OsBP-73 protein by Pfam domain prediction at the website www.wustl.edu. Secondary structure was predicted by PCGENE5.0. (III) Comparison of the SAP-like domain of OsBP-73 with the SAP domain of human SAF-A. Consensus amino acids are in bold. Two additional amino acids GY in OsBP-73 are boxed. C. Southern blot analysis of *OsBP-73* gene. Total DNA of rice cultivar zhonghua11 were digested with restriction enzymes *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H) and *Pst*I (P) and transferred onto Lynon membrane for Southern blotting with the ^{32}P-labeled *Eco*RI/*Xba*I fragment of c332 cDNA as a probe.

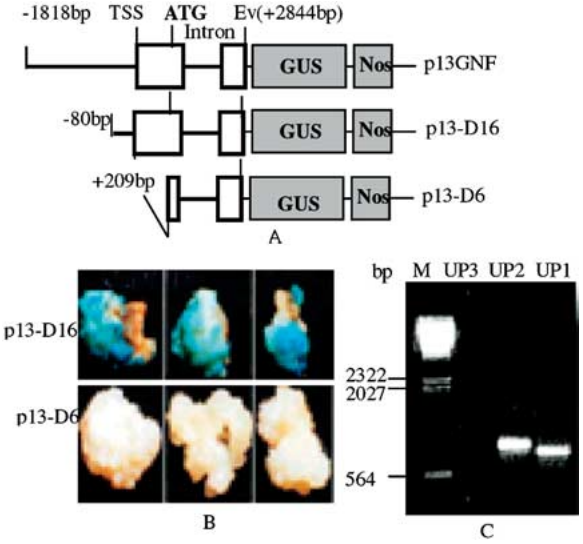

Figure 2. Analysis of the transcription start site of the *OsBP-73* gene. A. Schematic diagram of the constructs for promoter deletion analysis. B. GUS histochemical staining of resistant calluses. C. RT-PCR was carried out using total RNA from rice immature seed and amplified by UP1, UP2 and UP3 respectively with reverse primer 3ORF. M, λ DNA/*Hin*dIII marker.

probe was mixed with purified fusion protein and 4 μg calf thymus DNA in a total volume of 20 μl, with or without competitor DNA in binding buffer (10 mmol/l Tris-Cl pH 8.0, 50 mmol/l NaCl, 1 mmol/l EDTA, 1 mmol/l DTT, 10% glycerol, 0.05% NP40). After incubation for 15 min at 25 °C, the reaction mixtures were subjected to electrophoresis on 4% polyacrylamide gels in 0.25× TBE (22.5 mmol/l Tris-borate pH 8.0, 0.25 mmol/l EDTA). The gel was dried and exposed to X-ray film to produce an autoradiogram.

Southwestern blot analysis

Protein fractions were separated by SDS-PAGE and southwestern blot analysis was performed with the [^{32}P]-labeled 31 bp nucleotide sequence as probe following the method described by Quayle and Feix (1992).

DNA and RNA gel blot analysis

DNA was extracted from the young leaves of *Oryza sativa* ssp. *japonica* cv. Zhonghua 11 and a gel blot was performed as described previously (Maniatis *et al*., 1982).

Total RNAs were isolated and RNA gel blots were carried out as previously described (Wang *et al*., 1995).

Construction of the OsBP-73 *5′/GUS chimeric plasmid*

The DNA fragment containing the GUS-coding region and the NOS terminator was obtained from *Agrobacterium tumefaciens* binary vector pBI101.1 by digestion with *Eco*RI and *Hin*dIII, then cloning this DNA fragment to another *A. tumefaciens* binary vector, pCAMBIA1300 (http://www.cambia.org.au/) to produce plasmid p13GN. A ca. 4.6 kb DNA fragment from −1818 to +2844 containing ca. 4.5 kb of 5′-flanking sequence of second exon of *OsBP-73* plus the 148 bp second exon was then inserted into the polycloning site of p13GN to produce plasmid p13GNF(*OsBP-73* 5′/GUS). Sixty-five amino acids of the OsBP-73 amino terminus are present in this fusion protein. Nine 5′ deletion plasmids were generated from p13GNF by the method of *Exo*III. In turn, there are p13-D10, -D7, -D4, -D11, -D21, -D29, -D16, -D6 and p13GNK. Two of them, p13-D16 and p13-D6, were chosen as the representatives (Figure 2A). Their first base pairs are indicated in Figure 1A.

Transformation of rice

A. tumefaciens-mediated transformation of rice was performed as previously described (Hiei *et al*., 1994).

Histochemical assays of OsBP-73 *5′/GUS reporter gene expression*

Localization of GUS activity in transgenic rice plants was examined by the method described previously (Yang *et al*., 2002).

Construction for double-stranded RNA interference

A schematic diagram of DNA construct is shown in Figure 6A. A 2.0 kb promoter of the *Ubi* gene was inserted in the *Hin*dIII and *Bam*HI sites of p13GN and generated plasmid p13UGN. The 371 bp gene-specific fragment of c332 cDNA (1–371) was ligated to the p13UGN *Sac*I site in sense configuration and the same fragment was ligated to the *Xba*I and *Bam*HI sites in antisense configuration to generate the pUbi::A-GUS-S plasmid. The 1.86 kb GUS-coding region is as a linker between gene-specific fragments in the antisense and the sense configuration.

Results

Isolation and sequence analysis of OsBP-73 *cDNA and genomic DNA*

To identify proteins that recognize a 31 bp DNA sequence in the 5′-upstream region of the rice *Wx* gene, we used a yeast one-hybrid screen (Chen *et al*., 1999). The positive cDNA clones we isolated from a rice cDNA library fall into three groups based on cross-hybridization. The largest group contains eight members. Sequence analysis of three members, c33, c31 and c73, showed that they were transcribed from the same gene albeit of different length (Figure 1A). However, the longest cDNA c33 was still truncated at the 5′ end. To obtain full-length cDNA, another rice cDNA library was screened with the labeled c73 cDNA as probe. Three positive cDNA clones were isolated. Sequence analysis revealed that the longest cDNA clone c332 was 1524 bp long. The sequence of c332 predicts a protein of 375 amino acids with a methionine residue at the amino terminus. The corresponding gene is named *OsBP-73* (*Oryza sativa* binding protein −73).

Southern blot analysis of rice genome DNA was performed to determine the number of genomic DNA fragments with homology to the *OsBP-73* cDNA. Genomic DNA digested with various restriction enzymes and probed with a c332 cDNA fragment between *Eco*RI and *Xba*I. It is known that there is one *Pst*I site but there are no *Bam*HI, *Eco*RI or *Hin*dIII sites in the probe sequence. Genomic DNA digested with *Bam*HI, *Eco*RI or *Hin*dIII restriction enzymes produced only one strong hybridized band, and samples digested with *Pst*I produced two hybridized bands (Figure 1C). These results show that the *OsBP-73* gene is present as a single copy in the rice genome.

With cDNA c73 as probe, the genomic DNA of the *OsBP-73* gene was isolated. A ca. 6.0 kb insertion fragment was subcloned and sequenced. The nucleotide sequence of the genomic clone is shown in Figure 1A. Alignment of the genomic and cDNA sequences showed that the *OsBP-73* gene is composed of two exons (+1 to +226, +2698 to +4078) and a large intron (+227 to +2697) with standard GT/AG termini (Figure 1A). The genomic DNA and cDNA sequences have been submitted to EMBL/GenBank (accession numbers AJ315790 and AJ315791 respectively).

Analysis of the transcription start site of the OsBP-73 *gene*

Promoter deletion was carried out to analyze the putative transcription start site. A chimeric gene containing the 5′-flanking region of the second exon of *OsBP-73* gene (from −1818 to +2844) and the coding region of GUS was constructed in frame, designated p13GNF (*OsBP-73* 5′/GUS) (Figure 2A). Nine 5′ series deletion constructs were generated from p13GNF by using *Exo*III. The shortest construct with the ability to drive *GUS* expression was p13-D16 whose first base was at position T_{-80} (Figure 2B). The first base of the longest cDNA c332 is known to be T_{+85} (Figure1A). This result indicates that the TATA box is located between T_{-80} and T_{+85}. RT-PCR was carried out with three forward primers, UP1, UP2 and UP3, within the region of T_{-80} to T_{+85}, respectively, with the reverse primer 3ORF, which is in the second exon. Figure 2C shows that with UP1 and UP2 a specific band was amplified, whereas with UP3 no band was amplified. We also confirmed that the genomic sequences could be amplified with all the three pair of primers (data not shown). Sequence analysis of the PCR product amplified by UP2 and 3ORF showed that the sequence was identical to the sequence of the *OsBP-73* gene. Therefore the TATA box should be in the region of T_{-80} to G_{+29}. As predicted by the PCGENE5.0 program software, there is only one putative TATA box within this region. Therefore, the TATATGTT at −39 is the TATA box of this gene, T_1indicated in Figure 1A is the transcription start site.

Structure prediction for OsBP-73

Comparison of the OsBP-73 protein sequence with those available in sequence database by using the BLAST network service of the National Institutes of Health showed that the OsBP-73 protein has no significant match with the exception of the 35 amino acids (339–373) at the C-terminus. This 35 amino acid region shares significant similarity with three other proteins, two hypothetical *Arabidopsis* proteins and a *Micrococcus luteus* RHO protein (Figure 1B.I). The similarities between the C-terminus of OsBP-73 and these proteins are 72.5% to 94%. Secondary structure of the OsBP-73 protein was predicted by PCGENE5.0. A 35 amino acid region (339–373) at the C-terminus was composed of two helices separated by seven amino acids. A domain prediction program at www.wustl.edu revealed that a 35 amino acid region at the C-terminus of the OsBP-73 protein

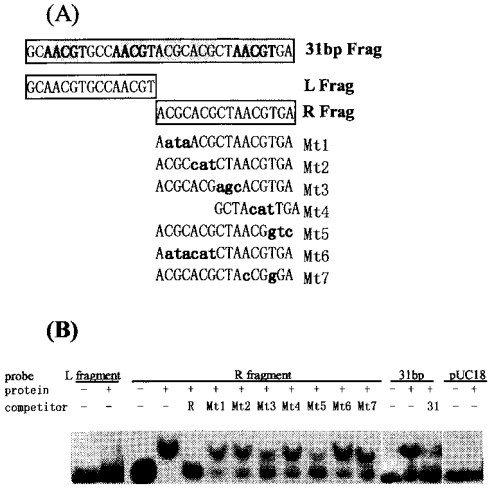

Figure 3. Gel retardation assay of the EB73 protein and the 31 bp nucleotide sequence. A. Sequence of 31 bp, L, R and seven mutated R fragments. B. Binding ability analysis of EB73 protein with L and R fragments, the 31 bp nucleotide sequence and the pUC18 polylinker.

was similar to a recently defined SAP DNA-binding domain (Figure 1B. II). The same prediction was also obtained at ExPASy Molecular Biology Server (http://www.expasy.org). Recent studies showed that the two hypothetical *Arabidopsis* proteins and the RHO protein we found in the database form an HEH structure at this region, and it is structurally homologous to the SAP DNA binding domain (Aravind and Koonin, 2001). All these predictions demonstrate that the 35 amino acid region at the C-terminus of OsBP-73 shares sequence and domain similarity with SAP domain. We designated it as a SAP-like domain.

Specific DNA-binding activity of OsBP-73

Since c33, c31 and c73 were isolated in a yeast one-hybrid screen, the DNA-binding specificity of OsBP-73 needed to be confirmed by an independent method. The c73 cDNA was cloned into the *E. coli* expression vector pET28c(+), named pEB73. Plasmid pEB73 was then introduced into *E. coli* Bl21 (DE3). The His-tagged EB73 protein was expressed by IPTG induction and purified on Ni-NTA resin. Gel retardation assays were carried out with purified EB73 protein with the ^{32}P-labeled 31 bp nucleotide sequence. The purified EB73 protein produced a strong shift band which could be competed by 100-fold excesses of unlabeled probe (Figure 3B). Since the probe used in the experiments contained both the 31 bp nucleotide sequence and part of the polylinker sequence of the pUC18 vector (see Materials and methods), the binding ability of EB73 with the polylinker sequence of pUC18 was also assayed. No shifted band was detected (Figure 3B). These results show that EB73 protein can bind specifically to the 31 bp nucleotide sequence.

The DNA-binding ability of EB73 with the 31 bp nucleotide sequence was also detected by southwestern blotting with the radioactively labeled 31 bp nucleotide sequence as probe (Figure 4C). Distinct hybridized bands were observed in the lanes of purified protein EB73 and the total protein from *E. coli* Bl21 (DE3)/pEB73 after IPTG induction. No hybridized signal was detected in the control lanes.

To analyze the binding site in the 31 bp nucleotide sequence, additional gel retardation assays were performed to examine the binding ability of EB73 to the right part (R, 16 bp) and left part (L, 15 bp) of the 31 bp nucleotide sequence (Figure 3B). The R fragment showed a very strong band shift that could be competed by 100-fold excesses of unlabeled R fragment whereas the L fragment did not show a band shift at all (Figure 3B). These results show that the R fragment contains the binding site for EB73 protein. Additional competition experiments were performed using seven synthesized mutants of the R DNA fragment, Mt1-Mt7, as competitors in gel retardation assays (Figure 3B). The shifted band produced by the R fragments and EB73 protein was competed completely by 100-fold excesses of fragment R. The R mutated fragments Mt1-Mt4, Mt6, and Mt7 showed little or no competition. However, Mt5 seemed to be capable of partially competing away band shift. This result suggests that all of the bases might be responsible for the R fragment binding to the EB73 and the

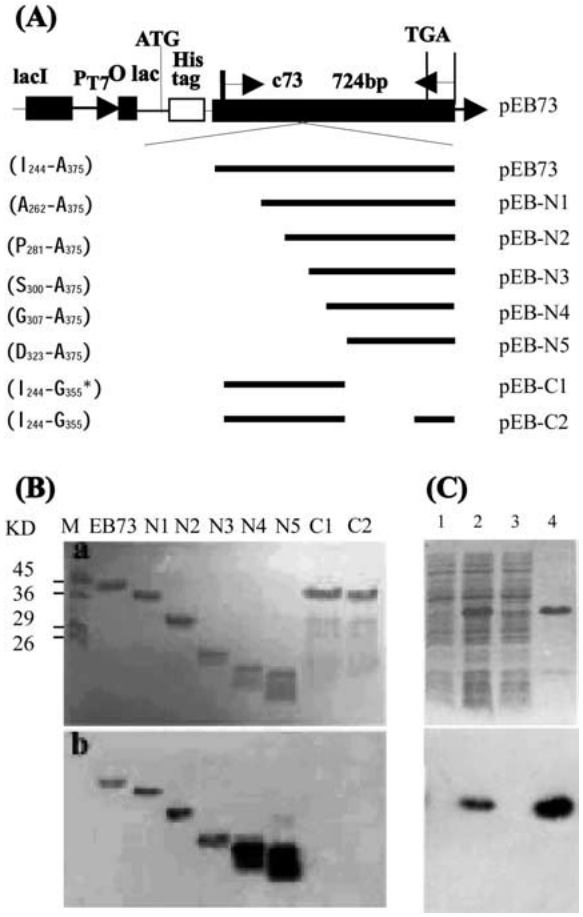

Figure 4. Southwestern blotting assay of various EB73 proteins with 31 bp fragments. A. Schematic diagram of series deletion plasmids. B. Southwestern assay for DNA-binding activity of a series of deletion proteins with the labeled 31 bp nucleotide sequence. The deletion proteins were stained with Brilliant Blue (a) and transferred onto a PVDF membrane for southwestern assay (b). Lanes: M, protein molecular weight marker; 1–8, purified deletion proteins: EB73, EB-N1, EB-N2, EB-N3, EB-N4, EB-N5, EB-C1 and EB-C2. C. Southwestern assay of the binding activity of EB73 protein with the labeled 31 bp nucleotide sequence. The bacterially expressed protein is stained with Coomassie Brilliant Blue (upper panel) and transferred onto PVDF membrane for southwestern assay (bottom panel). Lanes: 1, total protein extracts from *E. coli* BL21(DE3)/pET28c(+) after IPTG induction; 2 and 3, total protein extracts from *E. coli* BL21(DE3)/pEB73 after and before IPTG induction respectively; 4, purified fusion protein EB73.

3′-end three base pairs, TGA, was less critical than other bases in the 16 bp R fragment.

Mapping the DNA-binding domain of OsBP-73 protein

Because the protein encoded by the shortest c73 cDNA could bind to the 31 bp nucleotide sequence specifically, this delimits the DNA-binding region of OsBP-73 to the C-terminal 132 amino acids. The DNA-binding domain of OsBP-73 was further fine-mapped by partial deletion. With pEB73 as the start plasmid, a series of 5′ deletion plasmids were generated in frame by *Exo*III, named pEB-N1, pEB-N2, pEB-N3, pEB-N4 and pEB-N5. Two 3′ deletion plasmids, named pEB-C1 and pEB-C2, were obtained by restriction digestion (Figure 4A). All the deletion fusion proteins were expressed in *E. coli* and purified on Ni-NTA resin. Southwestern blot assay was performed to analyze the binding ability of these deletion proteins with the 31 bp nucleotide sequence (Figure 4B). The results show that EB73, EB-N1, EB-N2, EB-N3, EB-N4 and EB-N5 produced strongly hybridizing bands and that EB-C1, EB-C2 and the protein molecular markers did not produce any hybridizing signal. This result indicates that the 53 amino acid region at the C-terminus of the OsBP-73 protein is responsible for the DNA-binding ability to the 31 bp nucleotide sequence. Gel retardation assays of these deletion proteins with the 31 bp nucleotide sequence also produced a similar result (data not shown).

Expression pattern of the OsBP-73 *gene in rice*

With total RNA from various tissues, a northern blot was performed to analyze the expression of the *OsBP-73* gene. The *OsBP-73* gene was expressed in all tissues analyzed (Figure 5A). Although about 100 μg of total RNA was loaded per lane, a weakly hybridizing band was produced on each lane. This result suggests that the *OsBP-73* gene is expressed at low levels in all tissues analyzed. The ca. 1.6 kb transcript detected by northern blotting matches the predicted size of the full-length cDNA, indicating that the c332 cDNA is a nearly full-length transcript and that the predicted transcription start site is correct.

The expression of the *OsBP-73* gene at different stages of rice seed development was further investigated. Figure 5B shows that the transcript is expressed during all of the stages examined. However, the *OsBP-73* transcript is reduced in the late stages of seed development. In order to understand the *OsBP-73* gene expression pattern in more detail, the chimeric gene p13GNF (*OsBP-73* 5′/GUS) was chosen to transform rice plants by *Agrobacterium*-mediated transformation. Five independent transgenic lines were obtained and a histochemical GUS staining assay was performed on three plants of each independent transgenic line to visualize *GUS* expression in various tissues.

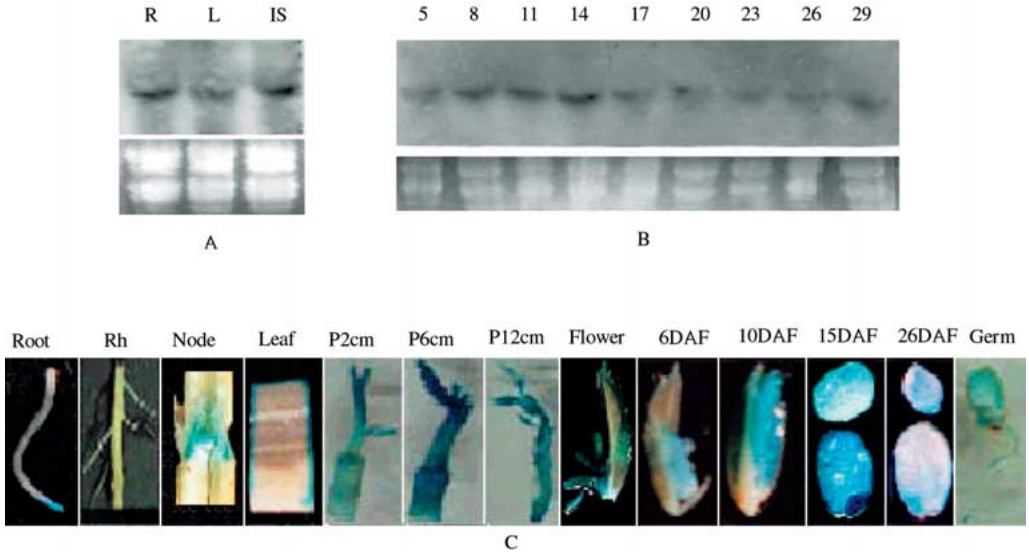

Figure 5. Expression pattern of the *OsBP-73* gene. A, B. Northern blot analysis of *OsBP-73* gene expression in rice with the ^{32}P-labeled c73 cDNA as probe. A. Northern blot analysis was carried out with about 100 μg total RNA from rice tissues including root (R), leaf (L), immature seed (IS) 14 days after flowering (DAF). B. Northern blot was performed with the total RNA from rice seed at 5–29 DAF. C. Histochemical GUS assays of the expression of a chimerical *OsBP-73* 5′/GUS reporter gene in various tissues (root, node and leaf), the developing panicles (P2cm, P6cm and P12cm), the developing seeds (Flower, 6DAF, 10DAF, 15DAF and 26DAF) and the germinating seeds of transgenic plants. Rh, Root hair; P2cm, P6cm and P12cm, panicle with a length of about 2, 6 and 12 cm respectively; Germ, germination.

GUS activities were commonly observed in root tip, stem node, panicle and immature seed, but not in root hair and culm (Figure 5C). Most of the tissues in which GUS activity can be detected have high cell-division activities.

Genetic interference of OsBP-73 *by double-stranded RNA inhibits rice plant growth*

Double-stranded RNA interference is an effective way to discover gene function in *Caenorhabditis elegans* and *Drosophila* (Montgomery *et al*., 1998; Misquitta and Paterson, 1999), and it also works in plants (Waterhouse *et al*., 1998; Chuang and Meyerowitz, 2000). We investigated the biological function of *OsBP-73* gene in rice by double-stranded RNA interference. To make a construct that produces dsRNA, the 371 bp gene-specific sequences of cDNA c332 in the antisense and sense configurations were linked with the GUS-coding region and placed under control of the constitutive *Ubi* promoter (pUbi::A-GUS-S). A single RNA transcribed from the chimeric gene in pUbi::A-GUS-S can potentially form a dsRNA stem with a single-stranded loop structure (Figure 6A). We obtained 25 independent transgenic lines by *Agrobacterium*-mediated transformation. Among them, 18 lines showed a decrease in plant growth rate. PCR analysis indicated that the T_1 generation plants with a growth inhibition phenotype from these RNAi transgenic lines contained the pUbi::A-GUS-S construct, whereas those segregated to non-transgenic plants whose phenotype is similar to the wild-type plants did not harbor the RNAi construct (data not shown). Northern blot analysis showed that the *OsBP-73* transcripts in T_1 plants with growth inhibition phenotype from four independent transgenic lines are significantly reduced than those in the wild type plants (Figure 6B). Effects of the double-stranded RNA interference of *OsBP-73* were investigated during the growth process of T_1 progeny plants with growth inhibition phenotype from four independent transgenic lines. In comparison with the wild type, RNAi plants showed a ca. 2-fold reduction in height, light green leaves and reduced tiller number, root number and panicle size (Figure 7A–E). The reduction in plant height was mainly due to shortening in the culm. The length of all the internodes analyzed at the adult phase was affected (Figure 7F). The panicle axis in RNAi plants was also 2-fold smaller than the wild type, which may also contribute to the effect on plant height. The RNAi plants showed reduced size in blade width, blade length and sheath length. The panicles of RNAi plants also had a reduced number of primary branches, second branches and spikelets (Table 1).

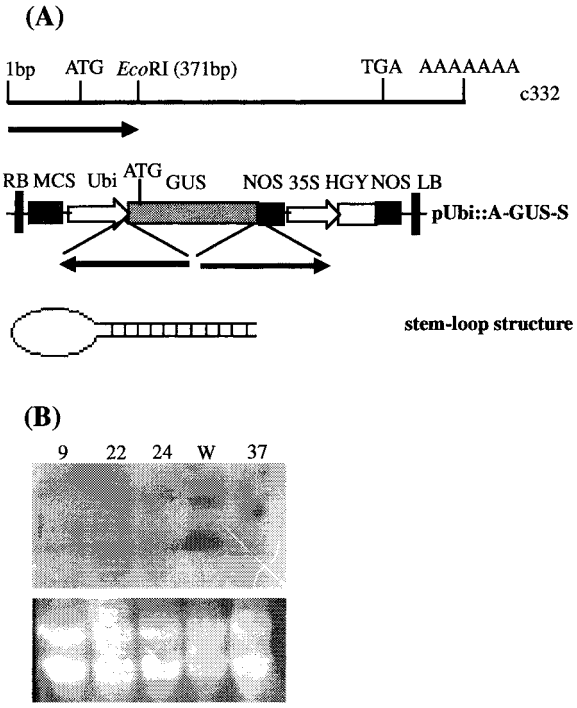

Figure 6. RNAi construct for transgenic plants. A. Schematic diagram of pUbi::A-GUS-S construct for double-stranded RNA interference. In pUbi::A-GUS-S, the 371 bp gene-specific fragment in cDNA c332 was linked with the 1.86 kb GUS coding region both in antisense and sense orientation. A schematic structure of the predicted dsRNA stem with a single-stranded loop generated by pUbi::A-GUS-S construct is shown. B. Northern blot analysis of *OsBP-73* gene expression in RNAi plants. Northern blot analysis was carried out with ^{32}P-labeled c73 cDNA as probe. About 100 μg total RNA from line 9, 22, 24 and 37 RNAi plants and the wild-type plants ZH-11 (w) at the four-leaf phase were loaded per track.

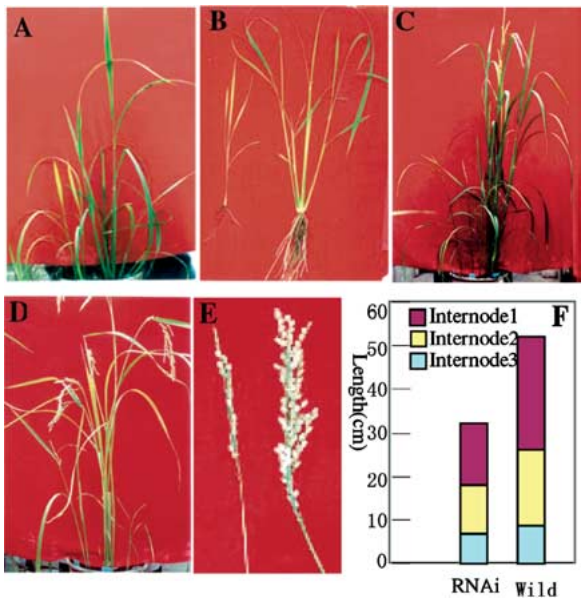

Figure 7. Effects of double-stranded RNA interference of the *OsBP-73* gene on rice plant growth. Comparison between the RNAi plants (left) and the wild-type plants (right). A. RNAi plants with reduced plant height, reduced tiller number, light green and small leaf. B. RNAi plants with reduced root number. C. RNAi plants with small panicles, but flowering nearly simultaneously with wild type. D. RNAi plants with reduced panicle axis length and small panicles. E. RNAi plants with small panicles and reduced number of spiklets per panicle. F. RNAi plants with reduced internode length. The length of internodes of wild-type and RNAi plants were measured 16 DAF starting from the internode below the panicle node (internode 1) downwards.

However, the passage from the juvenile to adult phase was unaffected, because the RNAi plants flowered nearly simultaneously with the non-transgenic plants. Flowers also had all organs and were normal in size. We also noticed that the size of the root tip cells and the leaf cells from the RNAi plants are similar to those of wild type plants (data not shown).

Discussion

We have identified an *OsBP-73* gene isolated by a yeast one-hybrid screen using the 31 bp nucleotide sequence in the 5′-upstream region of rice *Wx* gene as a probe. The *OsBP-73* gene contains two exons interrupted by a large intron and encodes a protein of 375 amino acids. The 2471 bp intron of the *OsBP-73* gene may have an effect on gene expression (data not shown). A database search revealed that the *OsBP-73* gene shares 99% identity with part of the sequences on the OSJNBa0032G08 BAC clone from rice chromosome III. It also reveals two rice EST sequences (accession numbers AU076093 and AY072715) that are identical to the 3′ end of the *OsBP-73* gene. The deduced OsBP-73 protein shares no significant sequence similarities with other known proteins with the exception of the C-terminal 35 amino acids (339–373). The C-terminal 35 amino acid sequence shares high sequence homologues to two hypothetical *Arabidopsis* proteins and a *Micrococcus luteus* RHO protein which are known to form an HEH structure, and it is structurally homologous to the SAP DNA-binding domain (Aravind and Koonin *et al.*, 2001). Domain prediction also showed that the 35 amino acid region (339–373) at the C-terminus of OsBP-73 is homologous to SAP DNA-binding domain (Figure 1B.II). The DNA-binding ability of the OsBP-73 protein to the 31 bp nucleotide sequence was demonstrated by gel retardation and southwestern blot assays. Southwest-

Table 1. Plant growth inhibition of RNAi plants.

Items	Wild type[1]	RNAi				
		Line 9[2]	Line22[2]	Line24[2]	Line37[2]	Average[3]
Leaf blade width						
L1	1.26±0.05	0.87	0.89	1.00	0.71	0.87±0.08
L2	1.01±0.03	0.65	0.78	0.61	0.50	0.64±0.08
L3	0.83±0.01	0.58	0.61	0.47	0.41	0.52±0.08
Leaf blade length						
L1	31.2 ±2.13	17.6	14.0	18.0	11.0	15.2 ±2.9*
L2	40.1 ±1.62	23.2	22.2	24.0	18.2	21.9 ±1.9*
L3	38.7 ±2.52	27.0	29.2	28.0	14.0	24.6 ±5.3*
Sheath length						
L1	25.4 ±1.09	16.8	15.1	17.8	9.2	14.7 ±2.7*
L2	22.0 ±0.64	14.1	12.2	14	8.6	12.2 ±1.8*
L3	20.1 ±0.67	11.6	9.0	12.4	7.1	10.0 ±1.9*
Culm length	55.0 ±0.6	32.6	28.5	33	21.0	28.8 ±4.1*
Panicle axis length	19.1 ±0.75	11.0	10.4	10.8	6.7	9.7 ±1.5*
Number of primary branches	8.4 ±0.4	4	3	3	1	2.8 ±0.9*
Number of secondary branches	15 ±0.55	3	2	1	0	1.5 ±1.0*
Number of spiklets per panicle	94.2 ±2.94	33	25	24	7	22.3 ±7.7*
Tiller number	3.8 ±0.37	0	1	0	0	0.3 ±0.3*

[1]Mean value of five wild-type plants.
[2]Mean value of two RNAi dwarf plants from each independent transgenic line.
[3]Mean value of total eight RNAi dwarf plants from four independent transgenic lines. Statistical significance was determined by t-test. Significance was determined at $P<0.001$ (*). Length and width are in cm and all data are presented as mean ± SE. L1 is the flag leaf, L2 and L3 are the leaves from the phytomers below.

ern blot assays of deleted OsBP-73 proteins confirmed that the 53 amino acid region from 323 to 375 at the C-terminus of OsBP-73 was critical for DNA-binding activity. Taken together, the results of both experiments and predictions show that there is a SAP-like DNA-binding domain at the C-terminus of OsBP-73.

The SAP domain is a recently defined DNA-binding domain that forms an HEH structure (Aravind and Koonin, 2000, 2001). As far as we know, the SAP domain in SAF-A is the best characterized SAP domain. It has been reported that the SAP domain in SAF-A does not tolerate additional insertions between two helices and mutations of the conserved amino acids. Both insertions and mutations will result in a complete loss of DNA-binding activity to SAR DNA elements (Gohring *et al*., 1997; Kipp *et al*., 2000b). However, compared with the SAP domain in SAF-A, the SAP-like domain in OsBP-73 has some different amino acids in the helices and two additional amino acids insertions between two helices (Figure 1B.III). Aravind and Koonin (2001) reported that the HEH structure at the N-terminus of transcription terminator RHO also has additional two amino acids insertion between two helices. These two additional amino acids seem to have no effect on the formation of the HEH structure. Two *Arabidopsis* hypothetical proteins we found in database search also have two additional amino acids between two helices respectively. This result suggests that the SAP-like domain of OsBP-73 shares structural similarity with the SAP domain of RHO rather than with that of the SAF-A. SAF-A was reported to bind to the SAR element that was rich in AT nucleotide sequences. The DNA-binding activity of the SAP domain in SAF-A was almost undetectable when tested in solution, such as a gel retardation assay, suggesting that SAF-A binds to DNA by the mass binding mode (Gohring *et al*., 1997; Kipp *et al*., 2000b). The results of gel retardation showed that OsBP-73 bound specifically to the R fragment of the 31 bp nucleotide sequence. Further competition experiment with mutated R fragments showed that OsBP-73 binds specifically to nucleotide sequence 5′-ACGCACGCTAACGTGA-3′. Since the OsBP-73 protein could bind specifically to the R fragment of the 31 bp nucleotide sequence in both gel retardation and southwestern blot assays, it suggests that the SAP-

like domain in OsBP-73 may behave differently from the SAP domain in SAF-A in DNA binding.

Recently, at least five SAP-domain-containing proteins have been identified. They have been implicated in a variety of functions, including regulation of cell proliferation (Galande and Kohwi-Shigematsu, 1999), programmed cell death (Sahara *et al*., 1999; Kipp *et al*., 2000a) and DNA repair and replication (Babiychuk *et al*., 1994; Lepiniec *et al*., 1995; Amor *et al*., 1998; Tian *et al*., 2000). SAP proteins have also been reported to regulate transcription (Nayler *et al*., 1998; Chabeaud *et al*., 2001; Storozhenko *et al*., 2001) and signal transduction (Liu *et al*., 1998). To elucidate the function of the *OsBP-73* gene, the expression pattern analysis of the *OsBP-73* gene and double-stranded RNA interference were carried out. Northern blot analysis revealed that the *OsBP-73* gene is weakly expressed in root, leaf and immature seed. More detailed information about the pattern of *OsBP-73* gene expression was obtained by histochemical studies of transgenic rice plants carrying an *OsBP-73* 5′/GUS reporter gene. The reporter gene is mainly expressed in the tissues with high cell-division activities, such as root tip, stem node, panicle and immature seed. Genetic interference of *OsBP-73* gene by double-stranded RNA strikingly inhibited the growth of the whole rice plant resulting in dwarf plants with small leaves and panicles and reduced root and tiller number. We also observed that the size of root tip cells and leaf cells from RNAi plants are similar to those of wild-type plants. These results suggest that the *OsBP-73* gene may play an important role in the regulation of cell proliferation. *OsBP-73* was originally isolated as the putative regulator of the rice *Wx* gene (Chen *et al*., 1999). In our study we demonstrated that OsBP-73 binds specifically to the 31 bp nucleotide sequence in the upstream of the rice *Wx* gene, and that the expression pattern of *OsBP-73* and *Wx* overlaps in the late stages during endosperm development (Wang *et al*., 1995). These results implicate that OsBP-73 protein might be also involved in the regulation of the rice *Wx* gene. Further experiments will be needed to elucidate the relationship between *OsBP-73* and the rice *Wx* gene.

Acknowledgements

We thank Michele Engel for her insightful comments on the manuscript. This work was supported by the National Natural Science Foundation of China (No. 39893322) and the Major State Basic Research Program of China: Functional genomics of important traits of rice (No. G1999011604).

References

Amor, Y., Babiychuk, E., Inzé, D. and Levine, A. 1998. The involvement of poly (ADP-ribose) polymerase in the oxidative stress responses in plants. FEBS Lett. 440: 1–7.

Aravind, L. and Koonin, E.V. 2000. SAP: a putative DNA-binding motif involved in chromosomal organization. Trends Biochem. Sci. 25: 112–114.

Aravind, L. and Koonin, E.V. 2001. Prokarotic homologus of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Res. 11: 1365–1374.

Babiychuk, E., Cottrill, P.B., Storozhenko, S., Fuangthong, M., Chen, Y., O'Fareell, M.K., Van Montagu, M., Inzé, D. and Kushnir, S. 1998. Higher plants possess two structurally different poly (ADP-ribose) polymerases. Plant J. 15: 635–645.

Babiychuk, E., Kushnir, S., Van Montagu, M. and Inzé, D. 1994. The *Arabidopsis thaliana* apurinic endonuclease Arp reduces human transcription factors Fos and Jun. Proc. Natl. Acad. Sci. USA 91: 3299–3303.

Chabeaud, G.D., Godon, C., Brutesco, C., Murcia, G.D. and Kazmaier, M. 2001. Ionising radiation induces the expression of PARP-1 and PARP-2 genes in *Arabidopsis*. Mol. Genet. Genomics 265: 954–963.

Chen, L., Wang, Z.Y., Zhang, J.L. and Hong, M.M. 1997. A 31 bp DNA fragment interacting specifically with nuclear proteins extracted form immature rice seeds. Acta Phytophysiol. Sin. 23: 313–318.

Chen, L., Xing, Y.Y., Zhang, J.L., Wang, Z.Y. and Hong, M.M. 1999. Isolation of the genes encoding DNA binding protein interacting with 5′ upstream region of the rice waxy gene by using yeast one-hybrid system. Acta Phytophysiol. Sin. 25: 110–114.

Chuang, C.F. and Meyerowitz, E.M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 97: 4985–4990.

Chung, C.D., Liao, J., Liu, B., Rao, X., Jay, P., Berta, P. and Shuai, K. 1997. Specific inhibition of Stat3 signal transduction by PIAS3. Science 278: 1803–1805.

Fackelmayer, F.O., Dahm, K., Renz, A., Ramsperger, U. and Richter, A. 1994. Nucleic acid binding properties of hnRNP-U/SAF-A, a nuclear matrix protein which binds DNA and RNA in vivo and *in vitro*. Eur. J. Biochem. 211: 749–757.

Galande, S. and Kohwi-Shigematsu, T. 1999. Poly(ADP-ribose)polymerase and Ku autoantigen form complex and synergistically bind to matrix attachment sequences. J. Biol. Chem. 274: 20521–20528.

Ge, H.F., Wang, Z.Y. and Hong, M.M. 2000. A 31 bp fragment within the 5′ upstream region of rice *Waxy* gene enhances gene expression. Acta Phytophysiol. Sin. 26: 159–163.

Gohring, F., Schwab, B.L., Nicotera, P., Leist, M. and Fackelmayer, F.O. 1997. The novel SAR-binding domain of scaffold attachment factor A (SAF-A) is a target in apoptotic nuclear breakdown. EMBO J. 16: 7361–7371.

Guille, M.J. and Kneale, G.G. 1997. Methods for the analysis of DNA-protein interactions. Mol. Biotechnol. 8: 35–52.

Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobac-*

terium and sequence analysis of the boundaries of the T-DNA. Plant J. 6: 271–282.

Hirano, H.Y., Tabayashi, N., Matsumura, T., Tanida, M., Koneda, Y. and Sano, Y. 1995. Tissue-dependent expression of the rice *wx+* gene promoter in transgenic rice and petunia. Plant Cell Physiol. 36: 37–44.

Kipp, M., Schwab, B.L., Przybylski, M., Nicotera, P. and Fackelmayer, F.O. 2000a. Apoptotic cleavage of scaffold attachment factor A (SAF-A) by caspase-3 occurs at a noncanonical cleavage site. J. Biol. Chem. 275: 5031–5036.

Kipp, M., Gohring, F., Ostendorp, T., van Drunen, C.M., van Driel, R., Przybylski, M. and Fackelmayer, F.O. 2000b. SAF-Box, a conserved protein domain that specifically recognizes scaffold attachment region DNA. Mol. Cell Biol. 20: 7480–7489.

Lepiniec, L., Babiychuk, E., Kushnir, S., Van Montagu, M. and Inzé, D. 1995. Characterization of an *Arabidopsis thaliana* cDNA homologue to animal poly(ADP-ribose) polymerases. FEBS Lett. 364: 103–108.

Liu, B., Liao, J., Rao, X., Kushner, S.A., Chung, C.D., Chang, D.D. and Shuai, K. 1998. Inhibition of Stat1-mediated gene activation by PIAS1. Proc. Natl. Acad. Sci. USA 95: 10626–10631.

Maniatis, T., Fritsch, E.F. and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Plainview, NY.

Meshi, T. and Iwabuchi, M. 1995. Plant transcription factors. Plant Cell Physiol. 36: 1405–1420.

Misquitta, L. and Paterson, B.M. 1999. Targeted disruption of gene function in *Drosophila* by RNA interference (RNA-i): a role for nautilus in embryonic somatic muscle formation. Proc. Natl. Acad. Sci. USA 96: 1451–1456.

Montgomery, M.K., Xu, S. and Fire, A. 1998. RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 95: 15502–15507.

Nayler, O., Stratling, W., Bourquin, J.P., Stagljar, I., Lindemann, L., Jasper, H., Hartmann, A.M., Fackelmayer, F.O., Ullrich, A. and Stamm, S. 1998. SAF-B protein couples transcription and pre-mRNA splicing to SAR/MAR elements. Nucl. Acids Res. 26: 3542–3549.

Piruat, J.I. and Aguilera, A. 1998. A novel yeast gene, *Tho2*, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. EMBO J. 17: 4859–4872.

Quayle, T. and Feix, G. 1992. Functional analysis of the −300 region of maize zein genes. Mol. Gen. Genet. 231: 369–374.

Sahara, S., Aoto, M., Eguchi, Y., Imamoto, N., Yoneda, Y. and Tsujimoto, Y. 1999. Acinus is acaspase-3-activated protein required for apoptotic chromatin condensation. Nature 401: 168–172.

Sanger, F., Nicklen, S. and Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463–5467.

Storozhenko, S., Inzé, D., Van Montagu, M. and Kushnir, S. 2001. *Arabidopsis* coactivator ALY-like proteins, DIP1 and DIP2, interact physically with the DNA-binding domain of the Zn-finger poly (ADP-ribose) polymerase. J. Exp. Bot. 52: 1375–1380.

Tian, R.H., Zhang, G.Y., Yan, C.H. and Dai, Y.R. 2000. Involvement of poly (ADP-ribose) polymerase and activation of caspase-3-like protease in heat shock-induced apoptosis in tobacco suspension cells. FEBS Lett. 474: 11–15.

Tang, W., Qiu, Z., Wang, Z., Zhang, J. and Hong, M. 1999. Expression of a rice cDNA c73 in *E. coli* cells and purification of its fusion protein. Acta Phytophysiol. Sin. 25: 401–407.

Wang, Z.Y., Zheng, F.Q., Shen, G.Z., Gao, J.P., Snustad, D.P., Li, M.G., Zhang, J.L. and Hong, M.M. 1995. The amylose content in rice endosperm is related to the post transcriptional regulation of the *Wx* gene. Plant J. 7: 613–622.

Waterhouse, P.M., Graham, M.W. and Wang, M.B. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. USA 95: 13959–13964.

Yang, H.J., Shen, H., Chen, L., Xing, Y.Y., Wang, Z.Y., Zhang, J.L. and Hong, M.M. 2002. The *OsEBP-89* gene of rice encodes a putative EREBP transcription factor and is temporally expressed in developing endosperm and intercalary meristem. Plant Mol. Biol. 50: 379–391.

Yanisch-Perron, C., Vieira, J. and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103–119.