J. Biochem. **132**, 613–620 (2002)

A Proteomic Analysis of Leaf Sheaths from Rice[1]

Shihua Shen,*[2] Masami Matsubae,† Toshifumi Takao,† Naoki Tanaka,* and Setsuko Komatsu*[3]

Department of Molecular Biology, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602; and †Institute for Protein Research, Osaka University, Yamadaoka 3-2, Suita, Osaka 565-0871

Received December 26, 2001; accepted July 27, 2002

The proteins extracted from the leaf sheaths of rice seedlings were separated by 2-D PAGE, and analyzed by Edman sequencing and mass spectrometry, followed by database searching. Image analysis revealed 352 protein spots on 2-D PAGE after staining with Coomassie Brilliant Blue. The amino acid sequences of 44 of 84 proteins were determined; for 31 of these proteins, a clear function could be assigned, whereas for 12 proteins, no function could be assigned. Forty proteins did not yield amino acid sequence information, because they were N-terminally blocked, or the obtained sequences were too short and/or did not give unambiguous results. Fifty-nine proteins were analyzed by mass spectrometry; all of these proteins were identified by matching to the protein database. The amino acid sequences of 19 of 27 proteins analyzed by mass spectrometry were similar to the results of Edman sequencing. These results suggest that 2-D PAGE combined with Edman sequencing and mass spectrometry analysis can be effectively used to identify plant proteins.

Key words: leaf sheath, mass spectrometry, protein sequencer, rice, 2-D PAGE.

Rice is one of the most important crops in Eastern Asia. A vast number of rice cultures as well as wild types of rice are grown worldwide, and their genetic and molecular make-up has been under active investigation. Rice (*Oryza sativa* L.) is a model plant for monocotyledons, especially for members of the grass family. Numerous attributes, such as small genome size, diploid nature, transformability, and establishment of genetic and molecular resources, make it a tractable organism for plant biologists. With an estimated genome size of 430 million base pairs (*1*), it is feasible to obtain the complete genome sequence of rice using current technologies. The functions of proteins depend considerably on post-translational modification and protein-protein interaction, processes that cannot be deduced from nucleic acid data. Therefore, efficient approaches for identifying proteins, for determining protein expression in different tissues and under different conditions, for identifying post-translational modification of proteins in response to different stimuli, and for characterizing protein interaction will be critical for understanding biological processes in the functional genomics and proteomics era (*2*). Proteomics will provide more fundamental insights into organismal development and homeostasis than can be provided by the genome sequence. Furthermore, proteomics has already become an important tool for drug discovery, and the analysis of yeast and *Escherichia coli* protein expression patterns (*3*).

The first large-scale plant proteomic work published was on rice (*4*). During the following years, proteomics emerged in various areas of biology, for example, Arabidopsis (*5*), maize (*6*), and pea (*7*). Furthermore, the important progress achieved in protein separation and identification makes objectives beyond the simple cataloguing of proteins a realistic aim. Present proteomics research aims at both identifying new proteins in relation to their function and ultimately at unraveling how their expression is controlled within regulatory networks (*8*). On the rice plant, there have been some quite classical works dealing with the construction of proteomes from complex origins, such as the leaf, embryo, endosperm, root, stem, shoot and callus proteome (*4, 9–12*). Proteomics studies to date have mainly focused on the changes in genome expression that are triggered by environmental factors. Examples of descriptive proteomes include the global comparison of green and etiolated rice shoots (*10*), and analysis of rice leaves and stems as to the effects of jasmonic acid treatment as a model for defense-associated responses (*13*). On the other hand, there has been no report about the leaf sheath proteome in spite of the importance for the elongation of rice. Briefly, an attempt was also made to determine the physiological significance of the proteins from rice leaf sheaths thus identified from their internal amino acid sequences. In the present study, a combination of 2-D PAGE, Edman sequencing, MALDI-TOF MS and MS/MS allowed unique analysis of the rice leaf sheath proteome.

[1]This work was supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences and by a grant for rice genome project PR-1201, MAFF, Japan.

[2]S. Shen is a Japan Science and Technology Agency Fellow from the Institute of Botany, Chinese Academy of Sciences, P.R. China.

[3]To whom correspondence should be addressed. Tel: +81-298-38-7446, Fax: +81-298-38-7408, E-mail: skomatsu@.affrc.go.jp

Abbreviations: MALDI-TOF MS, matrix-assisted laser desorption ionization time of flight mass spectrometry; CBB, Coomassie Brilliant Blue; TCA, trichloroacetic acid; GA, gibberellin; PVDF, polyvinylidene difluoride; ER, endoplasmic reticulum; RuBisCO, ribulose-1,5-bisphosphate carboxylase; SSU, small subunit; LSU, large subunit; SOD, superoxide dismutase; BBTI, Bowman Birk trypsin inhibitor; BiP, luminal binding protein.

© 2002 by The Japanese Biochemical Society.

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

MATERIALS AND METHODS

Chemicals—Urea was obtained from ICN (Aurora, OH, USA), and Ampholines from Amersham Pharmacia Biotech (Uppsala, Sweden). Acrylamide and 2-D SDS-PAGE standards were purchased from Bio-Rad (Richmond, CA, USA). *Staphylococcus aureus* V8 protease was obtained from Pierce (Rockford, IL, USA), and trypsin and α-cyano-4-hydroxycinnamic acid from Sigma (St. Louis, MO, USA).

Plant Material—Rice (*Oryza sativa* L. cv. Nipponbare) seedlings were grown under white fluorescent light (6,000 lux, 12 h light period/day) at 25°C. Three-cm-long leaf sheath segments were cut from 2-week-old rice seedlings.

Protein Extraction and 2-D PAGE—A portion (500 mg) of leaf sheaths was homogenized in 1 ml of a homogenization buffer containing 20 mM Tris-HCl (pH 7.5), 250 mM sucrose, 10 mM EGTA, 1 mM PMSF, 1 mM DTT, and 1% Triton X-100. The homogenate was centrifuged at 15,000 rpm for 5 min and then for 10 min in a TMA-4 rotor (Tomy, Tokyo), and the supernatant was placed in a new tube. The proteins were precipitated with 10% trichloroacetic acid (TCA) on ice for 20 min, and then centrifuged at 15,000 rpm for 10 min, and the pellet was suspended in 250 μl of a lysis buffer (*14*). The supernatant (50 μl) was subjected to 2-D PAGE (*14*). IEF was carried out in a glass tube, and SDS-PAGE in the second dimension was performed in a 17% separation gel. The isoelectric point and molecular mass of each protein were calibrated using 2-D PAGE standards.

Image Analysis—The CBB-stained gel was scanned using a flatbed scanner, and the data were analyzed using Image Master 2D Elite software (Amersham Pharmacia Biotech).

N-Terminal Amino Acid Sequence Analysis—After 2-D PAGE, proteins were electroblotted onto a PVDF membrane (Fluorotrans; Pall BioSupport Division, Port Washington, NY, USA) using a semi-dry transfer blotter (Nippon Eido, Tokyo), and visualized by CBB staining. The proteins were excised from the PVDF membrane and analyzed with a gas-phase protein sequencer (Model 492; PE Applied Biosystems, Foster City, CA, USA).

Internal Amino Acid Sequence Analysis—The proteins were separated by 2-D PAGE and stained with CBB. Gel pieces containing the protein spots were removed, and the proteins were electro-eluted from the gel pieces using an electrophoretic concentrator (ISCO, Lincoln, CA, USA) at 200 V for 2 h. After electro-elution, each protein solution was dialyzed overnight against pure water. The proteins were dissolved in 20 μl of an SDS sample buffer (pH 6.8) and then applied to a sample well in an SDS-PAGE gel. Twenty microliters of a solution containing 10 μl of *Staphylococcus aureus* V8 protease (0.1 μg/μl) in deionized water and 10 μl of an SDS sample buffer (pH 6.8) was overlaid on the sample solution. Electrophoresis was performed until the sample and protease were stacked in the upper gel, and interrupted for 30 min for digestion of the protein (*15*). Electrophoresis was then continued, and the separated digests were electroblotted onto a PVDF membrane and subjected to amino acid sequencing as above. The amino acid sequences were compared with protein sequences in a Swiss-Prot database using the FASTA sequence alignment program (*16*).

MALDI-TOF MS and ESI MS/MS—The CBB-stained protein spots were excised from a gel, washed with 25% (v/v) methanol and 7% (v/v) acetic acid for 12 h at room temperature, and then destained with 50 mM NH_4HCO_3 in 50% (v/v) methanol for 1 h at 40°C. Proteins were reduced with 10 mM DTT in 100 mM NH_4HCO_3 for 1 h at 60°C and then incubated with 40 mM iodoacetamide in 100 mM NH_4HCO_3 for 30 min at room temperature. The gel pieces were minced and allowed to dry, and then rehydrated in 100 mM NH_4HCO_3 with 1 pmol trypsin at 37°C overnight. The digested peptides were extracted from the gel slices with 0.1% TFA in 50% (v/v) acetonitrile/water three times. The peptide solution, thus obtained, was dried up and reconstituted with 30 μl of 0.1% TFA in 5% acetonitrile/water, and then desalted with ZipTip C18™ pipette tips (Millipore, Bedford, MA, USA). Matrix-assisted laser desorption ionization (MALDI)-MS was performed using a Voyager Elite XL time-of-flight mass spectrometer (Applied Biosystems, Framingham, MA, USA). The above peptide solution was mixed with the matrix solution, the supernatant of a 50% acetonitrile solution saturated with α-cyano-4-hydroxycinnamic acid, and then air dried on the flat surface of a stainless steel plate. Calibrations were carried out using a standard peptide mixture (*17*). The mass spectra were subjected to a sequence database search with Mascot software (Matrix Science, London, UK), and evaluated based on the obtained candidate sequence with MS-Match, a software aid for protein identification by MS (http://www/protein.osaka-u.ac.jp/organic).

Electrospray ionization (ESI)–MS/MS was carried out with a hybrid quadrupole orthogonal acceleration tandem mass spectrometer (Q-TOF) (Micromass, Manchester, UK). Solutions (1 μl) containing peptides were loaded into a borosilicate nanoflow tip (Micromass), and then set in an ESI source. Sodium iodide was used to calibrate the instrument. MS/MS data were processed with a maximum entropy data enhancement program, MaxEnt 3™ (Micromass). The resultant spectra were interpreted with SeqMS, a software aid for *de novo* sequencing by MS/MS (http://www/protein.osaka-u.ac.jp/organic). The sequence tags obtained were also used for the homology search in the database with FASTA.

RESULTS

Isolation of Proteins by 2-D PAGE—Since O'Farrell (*14*) introduced 2-D PAGE, it has been known as the most effective, as well as one of the simplest, methods of separating proteins directly from cells. Due to its high-resolving power, the technique has been employed to study alterations in cellular protein expression in response to various stimuli (*6*), and to monitor global changes that occur in the proteins at different developmental stages, or in tissues and organisms (*5, 11, 19*). In this study, proteins were extracted from rice leaf sheaths and separated by 2-D PAGE. A large number of protein spots were observed on a 2-D PAGE gel stained with CBB (Fig. 1A). Since it was often difficult to obtain sequenceable amounts of root proteins by the conventional protein extraction method (*4*), the proteins were concentrated by TCA precipitation prior to separation by 2-D PAGE in this study. This simple method was effective for recovering many proteins in sequenceable amounts. In a 2-D PAGE pattern, the proteins often showed slightly differ-

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

ent electrophoretic mobility according to the gel, and thus could not be easily compared without standardization (*4*). The protein pattern on a gel was calibrated according to the molecular masses and isoelectric points of the proteins. In conjunction with automated gel scanning and computer-aided image analysis, 352 protein spots were detected and numbered on the 2-D PAGE map (Fig. 1B). These image data may facilitate comparison of the protein patterns obtained with different experimental systems.

Analysis of Proteins by Edman Sequencing—To obtain sequence tags by Edman sequencing following separation by 2-D PAGE, the proteins were electroblotted onto a PVDF membrane and detected by CBB staining. Eighty-four protein spots for leaf sheaths were subjected to N-terminal sequencing (Table I). Among them, the amino acid sequences of 44 proteins (44/84 = 52.4%) were determined.

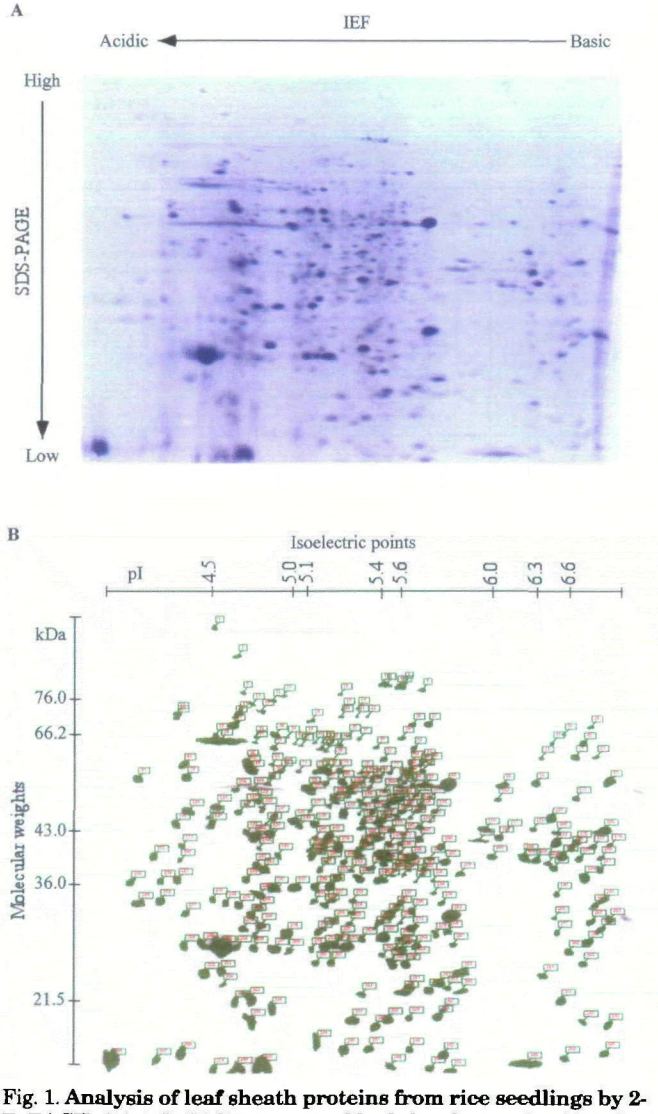

Fig. 1. **Analysis of leaf sheath proteins from rice seedlings by 2-D PAGE.** (A) 2-D PAGE pattern of leaf sheath proteins. Proteins were extracted from leaf sheaths, separated by 2-D PAGE, and detected by CBB staining. (B) Image analysis of leaf sheath proteins. The synthetic map was constructed using Image Master 2D Elite software. The protein image revealed 352 protein spots. Right to left: isoelectric focusing for the first dimension. Top to bottom: SDS-PAGE for the second dimension.

Four proteins (LS226, LS261, LS286, and LS316) (4/84 = 4.7%) were present in very low amounts and the amino acid peak materials were not sequenced. Forty proteins (40/84 = 47.6%) could not be directly sequenced, probably because they were N-terminally blocked. The data suggest that the N-termini of rice proteins were blocked at a higher frequency than observed for animal or microbial proteins (*19*). Blockage is common for plant proteins, and is caused by acetylation or the addition of a pyrrolidone carboxylate group to the N-terminal amino acid residue (*20, 21*). This problem could be overcome by digesting proteins with proteases and sequencing the resulting internal peptides. In this study, using the Cleveland peptide mapping method (*15*), internal amino acid sequencing was performed for 12 proteins (LS083, LS093, LS095, LS129, LS208, LS212, LS237, LS261, LS278, LS286, LS290, and LS291) (Table I).

The amino acid sequences determined here were compared with those of known proteins of plant and organisms in various public databases using FASTA. Twenty-nine proteins were found to be identical to ones already reported. Eighteen proteins were homologous to proteins from rice plants. Thirty-one proteins were homologous to ones from other plants, microbial proteins, or animals (Table I). These proteins are the lumenal binding protein (BiP) (LS030 and LS034), calreticulin (LS079 and LS083), tublin alpha-1 chain (LS089), ribulose-1,5-bisphosphate carboxylase (RuBisCO) activase (LS129 and LS162), phosphoribulo kinase (LS159), Bowman Birk trypsin inhibitor (BBTI) (LS261), G-box binding factor 4 (LS309), superoxide dismutase (SOD) (Cu/Zn) (LS310, LS317, LS327, LS328, LS329, and LS332), glycine cleavage system H protein precursor (LS340), RuBisCO small subunit (SSU) (LS346), RuBisCO large subunit (LSU) (LS083 and LS095), Elongation factor (LS208), Photosystem II oxygen-evolving complex (LS237) and ascorbate peroxidase (LS278), calcium-binding protein (LS008), Major capsid protein L1 (LS321), Thaumatin-like protein (LS326), peroxiredoxin (LS335), and Sal T (LS347, LS349, and LS351). Spots LS030 and LS034 have the same N-terminal sequences and are highly homologous to BiP from the endoplasmic reticulum (ER) protein in rice; spots LS079 and LS083 have the same N-terminal sequences and are identified as calreticulin from ER protein in rice. However, the isoelectric points of these proteins were different. These differences may be due to post-translational modifications or modifications in preparation processes. Edman sequencing is relatively time consuming, while protein identification by peptide mass fingerprinting allows high throughput analysis of samples (*22*).

Identification of Proteins by MS—Protein spots resolved by 2-D PAGE were excised from the gel and digested in-gel with trypsin. The resultant tryptic peptides of each protein spot were measured by MALDI-TOF MS. Based on a set of the observed molecular masses, a sequence database search (NCBInr: http://www.ncbi.nlm.nih.gov and DDBJ: http://www.ddbj.nig.ac.jp) using Mascot was carried out for protein identification. To further confirm the identification, several tryptic peptides of the above proteins were analyzed by ESI-MS/MS, yielding MS/MS fragment ions which were efficient for acquiring the sequence tags of proteins. Figure 2A shows the MALDI spectrum of a tryptic digest of LS261. The mass values listed in the spectrum were matched to rice BBTI (CAB88208) with a sequence coverage of 43%. Five peptides in the digest were further sub-

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

TABLE I. ***Identification of rice leaf sheath proteins with a protein sequencer.***

Spot No.	kDa	pI	Sequence	Homologous protein	Homology (%)	Accession
LS004	80.5	5.4	N-Blocked			
LS005	80.5	5.5	N-Blocked			
LS006	80.5	5.6	N-Blocked			
LS008	80	5.7	N-VAASKLQ	Calcium-binding protein 1 (*Dictyostelium discoideum*)	85.7	P42529
LS014	74	5.3	N-Blocked			
LS023	72	4.2	N-Blocked			
LS030	67	4.8	N-KEETKKLGTV	Lumenal binding protein (*Oryza sativa*)	100	AAB63469
LS034	66	4.9	N-KEETKKLGTV	Lumenal binding protein (*Oryza sativa*)	100	AAB63469
LS046	63	5.2	N-Blocked			
LS056	61	5.2	N-Blocked			
LS077	57	5.5	N-Blocked			
LS079	56	4.3	N-EVFFQEKFED	Calreticulin (*Oryza sativa*)	100	BAA88900
LS083*	56	4	N-EVFFQEKFED	Calreticulin (*Oryza sativa*)	100	BAA88900
			I-AKAGEDDDDL			
			I-AKKPEDWDDK			
			I-GYDDIPKEIP			
			I-FSNKDKTLVL			
LS089	55	4.8	N-MREIIS	Tubulin alpha-1 chain (*Oryza sativa*)	100	P28752
LS093	53.5	5.8	N-Blocked			
			I-TGEIKGHYLN	RuBisCO large subunit	100	P12089
			I-TLGFVDLLR			
LS095	53	5.7	N-Blocked			
			I-RDKLNKY	RuBisCO large subunit	100	P12089
			I-MTLGFVDLLR			
LS096	53	6.1	N-Blocked			
LS108	47	5.3	N-Blocked			
LS129	46	4.6	N-KELDEGKQTD	RuBisCO activase (*Oryza sativa*)	100	P93431
			I-GKQTDQDRWK			
LS147	45	4.5	N-Blocked			
LS159	43	4.8	N-SVDKPVVIGL	Phosphoribulokinase	100	PS0276
LS162	42	4.7	N-KELDEGK	RuBisCO activase (*Oryza sativa*)	100	P93431
LS208	38	5.5	N-Blocked			
			I-LANKGXAXL	Elongation factor	85.7	P42480
LS212	38	4.7	N-Blocked			
			I-VAFLTQGXAA	Unknown		
LS226	37	4.6	N-IGVXYG	Unknown		
LS237	36	4.9	N-Blocked	Photosystem II oxygen-evolving complex	100	A60731
			I-KDGIDYAAFT			
			I-VKGTGTANQ			
LS239	35	4.8	N-LLXTAELXPK	Unknown		
LS247	34	5.6	N-Blocked			
LS261*	32	5.8	N-Blocked			
			I-VEAPGD	Bowman Birk trypsin inhibitor (*Oryza sativa*)	100	AJ277468
LS266	31	5.5	N-Blocked			
LS278	29	5	N-Blocked			
			I-VTGGPEVPFH	Ascorbate peroxidase	100	BAB17666
LS286*	28.3	4.4	N-QVPL			
			I-YARLSDKDD	Unknown		
LS290*	28	4.5	N-Blocked			
			I-YARLSDKDD	Unknown		
LS291*	27.8	4.6	N-Blocked			
			I-YARLSDKDD	Unknown		
LS303	26.8	5.4	N-Blocked			
LS304	26.4	5.7	N-Blocked			
LS305	26	5.5	N-Blocked			
LS306	25.7	5.2	N-Blocked			
LS307	25.5	5.3	N-Blocked			
LS308	25.2	6.7	N-Blocked			
LS309*	25	5.6	N-QVEVXXLAA	G-box binding factor 4 (*Arabidopsis thaliana*)	66.7	P42777
LS310*	24.8	5.5	N-VKAVA	SOD (Cu/Zn) 2 (*Oryza sativa*)	100	P28757
LS311	24.5	4.5	N-Blocked			
LS312	24.2	4.4	N-Blocked			
LS313	24	5.8	N-Blocked			
LS314	23.9	6.3	N-Blocked			
LS315	23.7	5.7	N-Blocked			
LS316	23.3	4.5	N-SGAGAG	Unknown		
LS317*	23.1	5.6	N-VKAVVVLGSS	SOD (Cu/Zn) 1 (*Oryza sativa*)	100	P28756
LS318	22.9	5.8	N-Blocked			
LS319*	22.6	5.7	N-AGKDNPLGTL	Unknown		
LS320*	22.5	5.8	N-AIXQTIPGTD	Unknown		
LS321*	22.3	5.7	N-AKSGIPLDIV	Major capsid protein L1 (*Bovine papilloirus* type 4)	80	P08341
LS322*	22	5.3	N-AGGGEGDGG	C97454 rice callus cDNA clone (*Oryza sativa*)	100	C97454

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

TABLE I. continued

Spot No.	kDa	pI	Sequence	Homologous protein	Homology (%)	Accession
LS323*	21.5	4.7	N-AAPYVKET	Unknown		
LS324	21.5	6.4	N-Blocked			
LS325	21.1	4.6	N-Blocked			
LS326*	20.8	4.9	N-ATFTITN	Thaumatin-like protein precursor (*Oryza sativa*)	100	P31110
LS327*	20.6	5.8	N-VKAVAVLA	SOD (Cu/Zn) 2 (*Oryza sativa*)	100	P28757
LS328*	20.3	5.7	N-VKAVAVL	SOD (Cu/Zn) 2 (*Oryza sativa*)	100	P28757
LS329*	20.1	5.7	N-VKAVAVLSN	SOD (Cu/Zn) 2 (*Oryza sativa*)	100	P28757
LS330	20	5.5	N-Blocked			
LS331*	19.6	4.7	N-ATVMILMQP	Unknown		
LS332*	19.4	5.3	N-ATKKAVAVLK	SOD (Cu/Zn) chloroplast precursor (*Oryza sativa*)	100	P93407
LS333	19.3	6.7	N-Blocked			
LS334	19.1	6.9	N-Blocked			
LS335*	18.8	5.5	N-APVAVGDTLP	Peroxiredoxin (*Oryza sativa*)	100	AF203879
LS336	18	5.2	N-Blocked			
LS337	17.6	5.4	N-Blocked			
LS338	17.2	5.7	N-Blocked			
LS339	17	5.3	N-Blocked			
LS340*	16.8	4.3	N-STVLDGLKYS	Glycine cleavage system H protein precursor (*Mesembryanthemum crystallinum*)	90	P93255
LS341	16.8	6.9	N-Blocked			
LS342	16.7	5.7	N-Blocked			
LS343	16.6	6.4	N-Blocked			
LS344*	16.5	3.9	N-SGLDFVDLDN	Unknown		
LS345*	16.3	4.2	N-PYTFVD	Unknown		
LS346*	16.1	6.2	N-XQVWPIEGIK	RuBisCO small subunit (*Oryza sativa*)	100	P18566
LS347*	16	4.8	N-TLVKIGPWGG	Salt gene product (*Oryza sativa*)	100	OSAF001395
LS348	15.9	5.6	N-Blocked			
LS349*	15.8	4.6	N-TLVKIGPWGG	Salt gene product (*Oryza sativa*)	100	OSAF001395
LS350	15.6	5.7	N-Blocked			
LS351*	15.5	4.5	N-TLVKIGPWGG	Salt gene product (*Oryza sativa*)	100	OSAF001395
LS352	15.3	5.8	N-Blocked			

RuBisCO: ribulose-1,5-bisphosphate carboxylase. SOD: superoxide dismutase. N: N-terminal amino acid sequence. I: internal amino acid sequence. *Data were compared with those obtained on mass spectrometry (Table II).

jected to ESI-MS/MS, and partial sequencing was performed on the remaining peptides, leading to identification of an additional peptide of BBTI from rice (Fig. 2, B and C). Fragmentation of the doubly charged precursor ion at the *m/z* ratio of 702.4 yielded an ion series, from which the sequence was determined to be CCDSIVQLPQR (Fig. 2B), with cysteine (Fig. 2C) being alkylated with a carbamidomethyl group (see "MATERIALS AND METHODS") (*23*).

As a result, 59 protein spots were identified, based on the observed mass values listed in Table II or/and those of the fragment ions obtained on ESI-MS/MS (LS261, LS290, and LS311). Most of the leaf sheath proteins (53/59 = 89.8%) could be matched with the reported amino acid se-quences of rice proteins. The remaining 6 proteins occasionally matched the sequences reported for other plants as homologous proteins, spots LS303, LS319 and LS325 matched proteins from *Arabidopsis thaliana*; LS341 one from *Tortula ruralis*; and LS350 one from *Yougia iaponica*. Among those, 8 functional proteins have been reported, calreticulin (LS083), 2Cys-peroxiredoxin ABS1 precursor (LS311), SOD (Cu/Zn) 1 (LS317), 60S ribosomal protein L5 (LS319), SOD (Cu/Zn) 2 (LS328), SOD (Cu/Zn) chloroplast precursor (LS332), Nucleoside diphosphate kinase I (NDKI) (LS342), and RuBisCO SSU A (LS346). The functions of the other 46 proteins remain unknown.

Comparison of Results Obtained Using Protein Sequences and Peptide Mass Fingerprints—Fifty-nine proteins of leaf sheaths from rice seedlings were analyzed by Edman sequencing and MS. For Edman sequencing, most of the proteins were N-terminally blocked. Using MS, all proteins were identified by matching with proteins from rice and other species (Tables I and II). Similar proteins are spot LS083, homologous to calreticulin, a calcium-binding protein located in the ER (*24*); spot LS261, matching BBTI; spot LS317, identified as a SOD (Cu/Zn) 1, a cytoplasmic protein that destroys radicals that are normally produced within cells and are toxic to biological systems (*25*); spot LS322, found to be a C97454 rice callus cDNA clone; spot LS332, identified as a chloroplast SOD (Cu/Zn) (*26*); and spot LS346, matching the RuBisCO SSU from chloroplast protein, which is catalyzed in the first step of the Calvin cycle (*27*).

The N-terminal sequences of three proteins, spots LS327, LS328, and LS329, were identical to each other, and to the sequence of SOD (Cu/Zn) 2 (AC: P28757) from rice (*28*). Using MALDI-TOF MS, these proteins were found to belong to a protein family of SOD; spots LS327 and LS328 were homologous to SOD (Cu/Zn) 2; and spot LS329 was homologous to SOD [EC 1.15.1.1] (Fe) (AC: JG0179). Spots LS347, LS349, and LS351 had the same N-terminal sequences and were homologous to salt stress protein (*29*). On MS analysis, the three proteins matched salt-induced proteins. Two proteins (LS347 and LS351) matched the mannose-binding rice lectin (AC: BAA25369), which is a salt-induced protein, and another (LS349) matched a salt gene-product protein (AC: AAB53810).

Twenty-seven proteins were identified with both methods, 19 of which (19/27 = 70%) exhibited homology to proteins in rice databases (Tables I and II). The N-terminal sequences are important because the N-termini of mature proteins that often differ from those predicted from the open-reading frames of DNA sequences are shown.

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

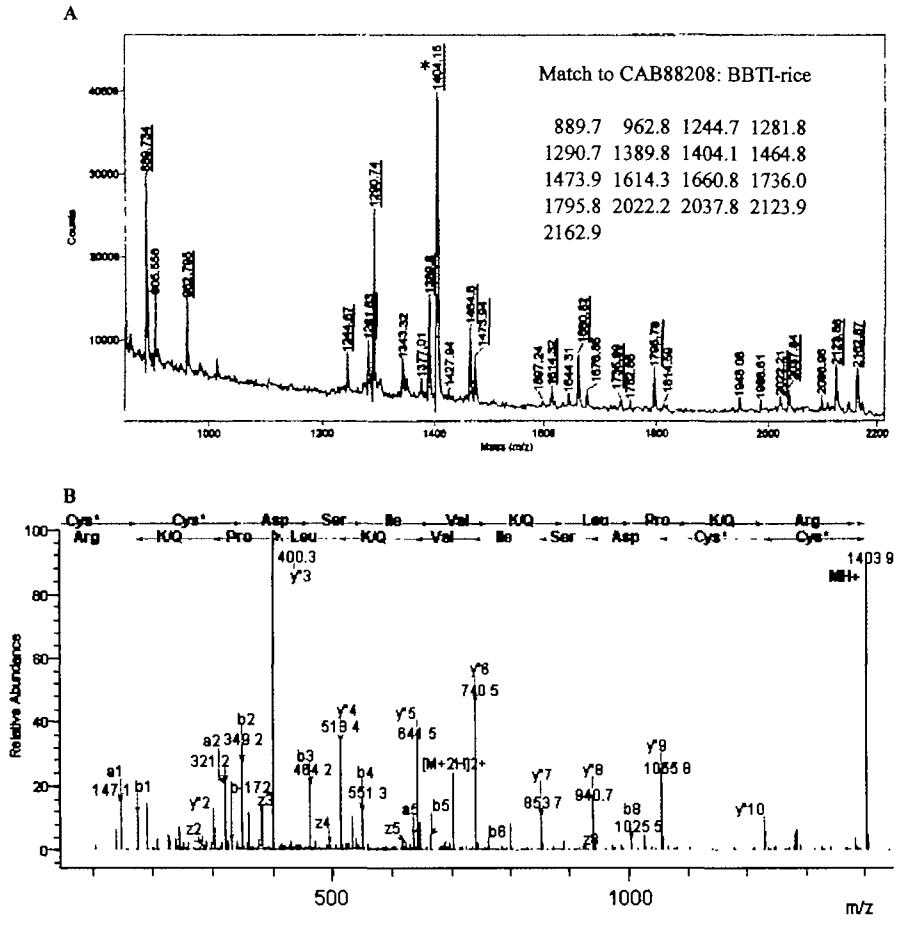

Fig. 2. **Identification of protein LS261 by mass spectrometry.** (A) The MALDI-TOF MS spectrum of peptides generated by tryptic digestion of protein LS261 matched BBTI from rice (underlining). (B) The MS/MS spectrum of the peptide recovered after in-gel tryptic digestion. Determination of the sequence tag, CCDSIVQLPQR, is shown. (C) Amino acid sequence of BBTI. The tryptic peptides observed in (A) are underlined, and the sequences obtained by ESI-MS/MS are in boldface.

MALDI-TOF MS analysis not only indicates the likely identity of proteins but can also confirm whether the protein species analyzed are post-translationally modified species or not (*2, 18*). Sometimes, the results of peptide mass fin-gerprinting need to be further confirmed by sequence tags. However, with the combination of these two techniques, it is possible to identify the corresponding genes.

DISCUSSION

Regulation of rice plant length is important for lodging resistance and then for increasing the grain yield. Since the leaf sheath is the tissue responsible for the elongation of rice plants, the plant length could be altered if the tissue growth is regulated. However, the mechanism of leaf sheath elongation remains largely unknown. Previously, rice proteome data-files for rice leaves, endosperms, embryos, roots, shoots and suspension cultured cells were constructed (*4, 9–12*). Their availability means that a leaf sheath proteome data-file can be compared with other rice tissue proteome data-files, and then leaf sheath–specific functional proteins can be identified. In the leaf sheath proteome data-file, RuBisCO activase (Table I; LS129 and LS162) was detected as a functional protein. RuBisCO activase regulates Ca^{2+}-dependent protein kinase and is the gibberellin (GA)-binding phosphoprotein (*30*). Over-expression of the RuBisCO activase gene in transgenic rice plants promoted seedling growth, on the other hand, the antisense transgenic rice plants showed dramatically reduced growth. A GA response mutant of rice, *slender rice1*, shows the GA signal controls the stem growth rate (*31, 32*). The results suggest that the GA signal transduction is one of the essential processes for rice stem growth. The profile shows calreticulin (Table I; LS079 and LS083) is expressed in the leaf sheath. Calreticulin also regulates Ca^{2+}-dependent protein kinase and regulates the regeneration of rice cultured suspension cells (*24*), suggesting that the Ca^{2+} signal is important for stem growth and development.

The MALDI-TOF MS list includes many hypothetical proteins (Table II). Determination of the functions of these proteins is important because the proteins might be related to the GA signal or Ca^{2+} signal pathways, and the knowledge of leaf sheath elongation mechanisms could progress.

A proteomics approach based on 2-D PAGE, Edman

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

TABLE II. **Identification of rice leaf sheath proteins by MALDI-TOF MS.**

Spot No.	Peptide sequences matched (Da)	Homologous protein	Seq. coverage (%)	Accession
LS069	659.6 704.2 726.1 729.3 824.6 827.6 889.6 916.8 1059.9 1776.6 1798.3 2729.4	Nitrilate-like protein	29	AB027054
LS073	659 7 684.5 729.3 824.7 889.6 916.8 1270.6 1776.4 1798.7 2729.7	Nitrilate-like protein	25	AB027054
LS083*	1025.5 1068.6 1090.9 1189.3 1504.6 1770.3 1909.4 1942.7 2091.7 2455.7 2521.2 2680.2 3661.1	Calreticulin	40	BAA88900
LS099	662.1 744.9 776.5 1112.13 1867.7 2578.2 2601.3 2712.2	Endo-1,3-beta-glucanase	24	AF443600
LS162	609.2 683.7 686.1 733.1 741.4 815.6 842.6 859.7 887.4	OsGAZox1	13	AB059416
LS212	625.4 656.3 694.6 804.8 848.8 860.2 971.2 1116.4 1179.4 1201.7 1591.9 1778.9 2069.9 2440.3 2760.9	Nrt2	26	AB008519
LS261*	889.7 962.8 1244.7 1281.8 1290.7 1389.8 1404.1 1464.8 1473.9 1614.3 1660.8 1736 1795.8 2022.2 2037.8 2123.9 2162.9	Bowman Birk trypsin inhibitor	43	CBA88208
LS286*	1297.9 1447.1 1474.4 3016.3	Hypothetical protein	25	BAA81756
LS290*	920.6 1505.9 1534.8 1634.1 2015.1 2309.4	Hypothetical protein	26	AAG03098
LS291*	792.9 989.2 1298.3 1460.3 1672.5 3014.3	Chloroplast 30S ribosomal protein	26	P12146
LS303	1243.2 2102.4 2378.4 2830.7	Hypothetical protein	36	CAC09513
LS305	1548.1 1620.6 1814.4 2046 2186.8 2794	Hypothetical protein	38	BAA90380
LS306	1642.7 1519 2448.2	Membrane-associated salt-inducible protein	18	BAA95872
LS307	1230.1 1691.1 1879.3 3215 3621.7	Similar to water stress inducible protein	37	BAA84793
LS308	1819.4 1876.6 2785.9	Hypothetical protein	27	AAG13553
LS309*	861.4 894.7 1122.7 1278.6 1392 1498.8 1602.1 1773.2 1843 1894.4 2077.7 2123	GSH-dependent dehydroascorbate reductase 1	72.8	BAA90672
LS310*	1505.4 1962.2 2531.3 2564.5 3185.3	Probable SOD	45	T04312
LS311	923.6 985.6 1063.6 1486.4 2273.2 2510	2Cys-peroxiredoxin ABS1 precursor	62	Q96468
LS312	1231.3 1480.3 1790.2 1844.9 2846.6	Probable submergence induced protein 2	37	T02787
LS313	1230.4 1730 2511.2 2649.4 2776.5	Histone H1	32	S59589
LS314	1763.5 2576.1 3427.2	Hypothetical protein	24	BAA95856
LS315	1597.9 2154.3 4468.6 4645.9	Hypothetical protein	47	BAB03435
LS316	614.0 625.2 636.4 647.6 676.6 705.5 770.9 969.8 1179.4 1257.9 1342.5 2058.1	Zinc finger transcription factor	25	AF332876
LS317*	1140.2 1427.8 1716.2 2044.3 3595	SOD (Cu/Zn) 1	64	P28756
LS318	1060 1134.5 1284 1791.9 2497.7 2775.6 2888.8 3313	Hypothetical protein	46	BAA89589
LS319*	1394.9 1444 1551.2 1706 2887.7	60S ribosomal protein L5	84	CAA79041
LS320*	1602.4 1958.8 2163.6	Hypothetical protein	25	BAA96163
LS321*	1572.4 1748.4 2057.7 2163.4 2450.9 2579	WSI18 protein	32	S52642
LS322*	1113.3 1340.2 1506.6 1917.6 1961.8 2163.7 2465.9 2574.5	C97454 rice callus cDNA clone	62	C97454
LS323*	1451.3 1686.3 1876.7 2177.7	Hypothetical protein	27	AAG12490
LS324	2402 2475.7 2777.8	Hypothetical protein	35	BAA99382
LS325	1403.5 2123.6 2259.2 2570.6	Hypothetical protein	33	T01608
LS326*	959.4 1382.1 1967.2 2093.3 2095.4 2248.7	Rice cDNA, partial sequence	62	D48816
LS327*	1764.1 1864.2 1891.6 2023.6	SOD (Cu/Zn) 2	26	P28757
LS328*	1716.8 2044.9 2759.5	SOD (Cu/Zn) 2	25	P28757
LS329*	1487.7 2468.3 3598.3	SOD [EC 1.15.1.1] (Fe)	12	JG0179
LS331*	1855.7 3232.4 3618.8	Hypothetical protein	23	BAA83581
LS332*	2866.8 3570.4 2564.5	SOD (Cu/Zn), chloroplast precursor	30	P93407
LS333	1214.9 1792 2573.6	Resistance gene analog PIC23	28	AAF21363
LS334	1695.9 2740.5 3321.5	Hypothetical protein	28	AAG13511
LS335*	1061 1996.4 2532.3 3577.7	Hypothetical protein	49	BAA96208
LS336	1650 1813.4 2274.9	Hemoglobin	30	AAF44664
LS338	1215.4 2164.8 2291.4 2678.4	Early light-induced protein, low molecular weight	39	T02036
LS339	1158.3 1759.7 1892.1 2135.4	Hypothetical protein	33	BAA96575
LS340*	1135.5 1696.3 2165.1 3183.5	Hypothetical protein	54	BAA90379
LS341	1011.5 1992.3 2274.7	Ribosomal protein S16	38.7	AAD23965
LS342	1246.2 1709.6 2670.3 3169.1	Nucleoside diphosphate kinase I	42	Q07661
LS343	1134.7 1209.1 1273.3 1792.3 2185.9 2385.6	Ubiquitin/ribosomal protein S27a fusion protein	54	S40240
LS344*	1545.1 3321.2	Hypothetical protein	32	BAB17094
LS345*	1135.4 2741.2 2773	Hypothetical protein	52	BAA96632
LS346*	1277.5 1475.9 3225.3	RuBisCO small subunit	29	P18566
LS347*	1407.8 1455.5 1535.8 1838.2 1992.2 3356.1 3034.2 2191.7 3117.6 3709.4	Mannose-binding rice lectin	96.6	BAA25369
LS348	1134.7 1695.4 1963.3 2311 2384.3	Ubiquitin/ribosomal protein CEP52	54	S33633
LS349*	1455.9 1650.4	Salt gene product	31	AAB53810
LS350	1650 1423.8 2290.7 2531	Ribosomal protein L20	40	X07676
LS351*	1455.9 1536 1993.1 3034.9 3118.1 3356.7	Mannose-binding rice lectin	96.6	BAA25369

GSH: glutathion. RuBisCO: ribulose-1,5-bisphosphate carboxylase. CEP: chondrocyte expressed protein. SOD: superoxide dismutase. *Data were compared with those obtained on Edman sequencing (Table I).

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012

sequencing and MS has enabled us to identify a number of leaf sheath proteins from rice seedlings. Three hundred fifty-two protein spots were detected on CBB staining. Eighty-four spots were analyzed by Edman sequencing; the amino acid sequences of 44 of these proteins were determined and found to be homologous to those of known proteins. Using MS, 56 protein spots were analyzed, and the amino acid sequences of these proteins matched ones in the protein database. Twenty-seven proteins were identified with both methods, 19 of which exhibited homology to proteins in the rice database. These results suggest that MS is a powerful tool for high-throughput identification of proteins, such as those obtained from a 2D-PAGE gel.

REFERENCES

1. Sasaki, T. (1998) The rice genome project in Japan. *Proc. Natl. Acad. Sci. USA* **95**, 2027–2028
2. Li, J. and Assmann, S.M. (2000) Mass spectrometry. An essential tool in proteome analysis. *Plant Physiol.* **123**, 807–809
3. Koc, E.C., Burkhart, W., Blackburn, K., Moseley, A., Koc, H., and Spremulli, L.L. (2000) A proteomics approach to the identification of mammalian mitochondrial small subunit ribosomal proteins. *J. Biol. Chem.* **275**, 32585–32591
4. Komatsu, S., Kajiwara, H., and Hirano, H. (1993) A rice protein library: a data-file of rice proteins separated by two-dimensional electrophoresis. *Theor. Appl. Genet.* **86**, 953–942
5. Kamo, M., Kawakami, T., Miyatake, N., and Tsugita, A. (1995) Separation and characterization of *Arabidopsis thaliana* protein by two-dimensional gel electrophoresis. *Electrophoresis* **16**, 423–430
6. Chang, W.W., Huang, L., Shen, M., Webster, C., Burlingame, A.L., and Roberts, J.K. (2000) Patterns of protein synthesis and tolerance of anoxia in root tips of maize seedlings acclimated to a low-oxygen environment, and identification of proteins by mass spectrometry. *Plant Physiol.* **122**, 295–318
7. Peltier, J.B., Frison, G., Kalume, D.E., Roepstorff, P., Nilsson, F., Adamska, I., and van Wijk, K.J. (2000) Proteomics of the chloroplast: Systematic identification and targeting analysis of luminal and peripheral thylakoid proteins. *Plant Cell* **12**, 319–342
8. Rossignol, M. (2001) Analysis of the plant proteome. *Curr. Opin. Biotechnol.* **12**, 131–134
9. Zhong, B., Karibe, H., Komatsu, S., Ichimura, H., Nagamura, Y., Sasaki, T., and Hirano, H. (1997) Screening of rice genes from a cDNA catalog based on the sequence data-file of proteins separated by two-dimensional electrophoresis. *Breed. Sci.* **47**, 245–251
10. Komatsu, S., Muhammad, A., and Rakwal, R. (1999) Separation and characterization of proteins from green and etiolated shoots of rice: Towards a rice proteome. *Electrophoresis* **20**, 630–636
11. Komatsu, S., Rakwal, R., and Li, Z. (1999) Separation and characterization of proteins in rice suspension cultured cells. *Plant Cell Tissue Organ Culture* **55**, 183–192
12. Tsugita, A., Kawakami, T., Uchiyama, Y., Kamo, M., Miyatake, N., and Nozu, Y. (1994) Separation and characterization of rice proteins. *Electrophoresis* **15**, 708–720
13. Rakwal, R. and Komatsu, S. (2000) Role of jasmonate in the rice self-defense mechanism using proteome analysis. *Electrophoresis* **21**, 2492–2500
14. O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021
15. Cleveland, D.W., Fischer, S.G., Kirschner, M.W., and Laemmli, U.K. (1977) Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel. *J. Biol. Chem.* **252**, 1102–1106
16. Pearson, W.R. and Lipman, D.J. (1998) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448
17. Jensen, O.N., Wilm, M., Shevchenko, A., and Mann, M. (1999) *2-D Proteome Analysis Protocols*, pp. 513–530, Humana Press, Totowa, NJ
18. Natera, S.H., Guerreiro, N., and Djordjevic, M.A. (2000) Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. *Mol. Plant Microbe Interact.* **13**, 995–1009
19. Tsugita, A., Kamo, M., Kawakami, T., and Ohki, Y. (1996) Two-dimensional electrophoresis of plant proteins and standardization of gel patterns. *Electrophoresis* **17**, 855–865
20. Takahashi, N., Takahashi, Y., and Putnam, F.W. (1985) Complete amino acid sequence of human hemopexin, the hemebinding protein of serum. *Proc. Natl. Acad. Sci. USA* **82**, 73–77
21. Wellner, D., Paneerselvam, C., and Horecker, B.L. (1990) Sequencing of peptides and proteins with blocked N-terminal amino acids: N-acetylserine or N-acetylthreonine. *Proc. Natl. Acad. Sci. USA* **87**, 1947–1949
22. Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., and Mann, M. (1996) Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. USA* **93**, 14440–14445
23. Boja, E.S. and Fales, H.M. (2001) Overalkylation of a protein digest with iodoacetamide. *Anal. Chem.* **73**, 3576–3582
24. Li, Z. and Komatsu, S. (2000) Molecular cloning and characterization of calreticulin, a calcium-binding protein involved in the regeneration of rice cultured suspension cells. *Eur. J. Biochem.* **267**, 737–745
25. Sakamoto, A., Okumura, T., Kaminaka, H., and Tanaka, K. (1995) Molecular cloning of the gene (SodCc1) that encodes a cytosolic copper/zinc-superoxide dismutase from rice. *Plant Physiol.* **107**, 651–652
26. Kaminaka, H., Morita, S., Yokoi, H., Masumura, T., and Tanaka, K. (1997) Molecular cloning and characterization of a cDNA for plastidic copper/zinc-superoxide dismutase in rice. *Plant Cell Physiol.* **38**, 65–69
27. Matsuoka, M., Kano-Murakami, Y., Tanaka, Y., Ozeki, Y., and Yamamoto, N. (1988) Classification and nucleotide sequence of cDNA encoding the small subunit of ribulose-1.5-bisphosphate carboxylase from rice. *Plant Cell Physiol.* **29**, 1015–1022
28. Sakamoto, A., Ohsuga, H., and Tanaka, K. (1992) Nucleotide sequences of two cDNA clones encoding different Cu/Zn-superoxide dismutases expressed in developing rice seed. *Plant Mol. Biol.* **19**, 323–327
29. Claes, B., Dekeyser, R., Villarroel, R., Van den Bulcke, M., Bauw, G., Van Montagu, M., and Caplan, A. (1990) Characterization of rice gene showing organ-specific expression in response to salt stress and drought. *Plant Cell* **2**, 19–27
30. Sharma, A. and Komatsu, S. (2002) Involvement of a Ca^{2+}-dependent protein kinase component downstream to the gibberellin-binding phosphoprotein, RuBisCO activase, in rice. *Biochem. Biophys. Res. Commun.* **290**, 690–695
31. Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M., and Yamaguchi, J. (2001) *Slender rice*, a constitutive response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene *GAI/RGA/RHT/D8*. *Plant Cell* **13**, 999–1010
32. Itoh, H., Ueguchi-Tanaka, M., Sato, Y., Ashikari, M., and Matsuoka, M. (2002) The gibberellin signaling pathway is regulated by the appearance and disappearance of SLENDER RICE1 in nuclei. *Plant Cell* **14**, 57–70

Downloaded from http://jb.oxfordjournals.org/ at Islamic Azad University on September 29, 2012