

# Expression and Function of Proteins during Development of the Basal Region in Rice Seedlings\*

Naoki Tanaka‡, Shigeyuki Mitsui§, Hiroya Nobori§, Koki Yanagi§, and Setsuko Komatsu‡¶

A differential display of proteins with a two-dimensional polyacrylamide gel electrophoresis approach was used to analyze protein expression changes during development of the basal region in rice seedlings (*Oryza sativa* L. cv. Nipponbare). The proteins were detected as 700 Coomassie Brilliant Blue-stained spots with pI values from around 3.5 to 9.0. A proteome reference map was established for the basal region of two-week-old seedlings. The basal region proteome map was used to analyze quantitative variations in the tissue during development from 2-, 4-, 6-, 8-, and 10-week-old seedlings. During development, 31 proteins were up-regulated, and 30 proteins were down-regulated compared with the 2-week-old basal region proteome map. The main functions of these proteins were primary metabolism and protein synthesis or maintenance. Calreticulin precursor, enolase, and voltage-dependent anion channel were identified among the up- and down-regulated proteins. The twin spots of calreticulin precursor and enolase with different pI values are possibly due to post-translational modifications such as phosphorylation. In addition, seven proteins showed developmental stage-specific expression. All of the developmentally regulated proteins of the basal region were clustered by the S-system, a differential equation that fit to time course of cluster and analyzed for cluster relationships. Proteins with unknown functions were tentatively assigned to functional groups based on cluster relationships. Basal region development proteome data will be valuable for resolving questions in functional genomics. In addition, cluster analysis of the basal region proteome during development will be useful for the assessment of functional proteins. *Molecular & Cellular Proteomics* 4: 796–808, 2005.

Rice (*Oryza sativa* L.) is an excellent model plant among the monocot crop species based on the fact that it has a relatively small genome of about 440 Mb (1) and is suitable for efficient genetic analysis and transformation. Draft rice genome se-

quences of the major subtypes *japonica* (2) and *indica* (3) have been published. The complete map-based genome sequences of chromosomes 1 (4) and 4 (5) of the Nipponbare cultivar have been reported and it is expected that the complete genome of *japonica* will be decoded and made available to the public (International Rice Genome Consortium, Japan).

In the postgenomic era, proteomics is becoming increasingly important because proteins are directly related to function (6). Because proteins are involved in most processes of living cells, a detailed understanding of the proteome is critical to the study of cells and organisms at the molecular level. Furthermore nucleotide sequences provide limited information about the protein complement of the genome, in large measure due to post-transcriptional regulation, which results in a lack of correlation between transcript levels and protein abundance. With the availability of genome sequences, proteomics is playing an increasingly important role in genome annotation and has recently been applied toward the understanding of plant development, disease resistance, photosynthesis, and other aspects of the plant proteome (7, 8).

The last decade has seen considerable progress in developing the field of rice proteomics with several studies describing the separation and characterization of proteins from various tissues or subcellular components (9, 10). Several studies have dealt with the construction of proteomes for complex samples from rice, such as leaf, embryo, endosperm, root, stem, shoot, anther, and callus (11–19), and subcellular components such as Golgi bodies, mitochondria, plasma membrane, vacuolar membrane, Golgi membrane, and chloroplasts have also been studied (20–22). These reports have focused on mapping the proteins and constructing a data base of expressed proteins of various tissues and subcellular components but have not addressed the changes in protein expression over time that are critical to understanding development in complex organisms.

Differential proteome analyses, which systematically study changes in the proteome during growth and development and under diverse environmental stimuli, are the next challenge in proteomics. Studies of environmental stimuli have focused on searching for marker proteins in response to stimuli, such as treatment with the plant hormones jasmonic acid (23), brassinolide (24), or gibberellin (19, 25); growth under salt (26);

From the ‡Department of Molecular Genetics, National Institute of Agrobiological Sciences, Tsukuba 305-8602, Japan and the §Mitsubishi Space Software Co., Ltd., Tokyo 105-6137, Japan

Received, December 20, 2004, and in revised form, March 7, 2005  
Published, MCP Papers in Press, March 11, 2005, DOI 10.1074/mcp.M400211-MCP200

drought (27); ozone stress (28); and infection with blast fungus (29, 30) or *Rice yellow mottle virus* (31).

Time course-dependent protein expression analyses would also be important because quantitative measures of reproducibility were not reported nor were rigorous quantitative analyses applied to group proteins into expression classes. Recently, using the de-etiolated (greening) of maize chloroplast as a model system, a general protocol that can be used to generate high quality, reproducible data sets for comparative plant proteomics was developed (32). In this report, hierarchical and nonhierarchical statistical methods were used to analyze the expression patterns of 526 high quality, unique protein spots on two-dimensional (2D)<sup>1</sup> gels.

The basal region of rice seedlings, including the crown, is a functionally important region where many critical metabolic and regulatory activities take place that eventually control the height and robustness of the plant. In a previous study, we analyzed the leaf sheath proteome, which includes proteins of the basal region (16, 19). No clear function could be predicted for 20% of the proteins in the leaf sheath proteome, but a majority of proteins involved in central metabolic pathways and energy production were identified. Although the basal region is important for rice plant growth, there are as yet no quantitative measures of proteome data. In the present study, changes in basal region proteins at five time points after sowing were analyzed quantitatively. Clustering analysis of differentially accumulated proteins during development was also carried out to clarify relationships among the proteins.

#### EXPERIMENTAL PROCEDURES

**Plant Material**—Rice (*O. sativa* L. cv. Nipponbare) seedlings were grown in plastic seedling pots under white fluorescent light (600  $\mu\text{m m}^{-2} \text{s}^{-1}$ ; 12-h light period/day) at 25 °C and 70% relative humidity in a growth chamber (Sanyo, Osaka, Japan).

**Protein Extraction**—Two grams of basal regions of fresh seedlings of each developmental stages were homogenized with 4 ml of a lysis buffer (33) containing 8 M urea, 2% Nonidet P-40, 0.8% Ampholine (pH 3.5–10 and pH 5–8, Amersham Biosciences), 5% 2-mercaptoethanol, and 5% polyvinyl pyrrolidone using a glass mortar and pestle on ice. Homogenates were centrifuged twice at 15,000 rpm in a RA-50 JS rotor (Kubota, Tokyo, Japan) for 5 min each. The supernatants of 40  $\mu\text{l}$  (100  $\mu\text{g}$  of protein) were subjected to 2D-PAGE. Three independent experiments of protein extraction from the rice tissues and of 2D-PAGE analysis were performed.

**Two-dimensional Polyacrylamide Gel Electrophoresis**—Samples were separated by 2D-PAGE in the first dimension by IEF for low pI range (around pI 3.5–8.0) or IPG tube gels (Daiichi Kagaku, Tokyo, Japan) for high pI range (pI 6.0–10.0) and in the second dimension by SDS-PAGE. IEF tube gels (11-cm length and 0.3-cm diameter) were prepared by filling with IEF gel solution consisting of 8 M urea, 3.5% acrylamide, 2% Nonidet P-40, and 2% Ampholine (pH 3.5–10 and pH 5–8). Electrophoresis was carried out at 200 V for 30 min followed by 400 V for 16 h and 600 V for 1 h. For IPG electrophoresis, samples

were applied to the acidic side of gels. Electrophoresis using IPG tube gels of 11  $\times$  0.3 cm was carried out at 400 V for 1 h followed by 1,000 V for 16 h and 2,000 V for 1 h. After IEF or IPG, SDS-PAGE in the second dimension was performed using 15% polyacrylamide gels with 5% stacking gels. Gels were stained with Coomassie Brilliant Blue (CBB).

**Gel Image Analysis**—2D polyacrylamide gels using IEF (around pI 3.5–8.0) and IPG (pI 6.0–10.0) for first dimension were overlapped around pI 5.6. The proteins in the edges of each gel were placed carefully to overlap the corresponding proteins. 2D-PAGE images were formed and evaluated automatically, and the amount of protein in each spot was estimated with Image-Master 2D Elite software (Version 2.0; Amersham Biosciences). The amount of a protein spot was expressed as its volume, defined as the sum of the intensities of the pixels that form the spot. To correct for variability due to CBB staining and to reflect quantitative variations in the intensity of protein spots, spot volumes were normalized as a percentage of the total volume of all the spots present in a gel. The pI and molecular mass of each protein were determined using a 2D-PAGE marker (Bio-Rad).

**Cleveland Peptide Mapping**—Following separation by 2D-PAGE, gel pieces containing protein spots were excised, and the protein was electroeluted using an electrophoretic concentrator (Nihon-Eido, Tokyo, Japan) at 2 watts of constant power for 2 h. After electroelution, the protein solution was dialyzed against deionized water for 2 days and lyophilized. Proteins were redissolved in 20  $\mu\text{l}$  of SDS sample buffer (0.5 M Tris-HCl (pH 6.8), 10% glycerol, 2.5% SDS, 5% 2-mercaptoethanol) and loaded onto an SDS-polyacrylamide gel. The sample was overlaid with 20  $\mu\text{l}$  of a solution containing 10  $\mu\text{l}$  *Staphylococcus aureus* V8 protease (0.1  $\mu\text{g } \mu\text{l}^{-1}$ ; Pierce) and 10  $\mu\text{l}$  of SDS sample buffer. Electrophoresis was performed until the sample and protease were stacked in the stacking gel, interrupted for 30 min to allow the protein to digest (34). After electrophoresis, gels were stained with CBB.

**N-terminal and Internal Amino Acid Sequence Analyses**—Proteins were sequenced by electroblotting onto a PVDF membrane (Pall, Port Washington, NY) and detected by CBB staining. Stained proteins were excised from the PVDF membrane and directly subjected to Edman degradation on a gas-phase protein sequencer (Procise 494, Applied Biosystems, Foster City, CA).

**Matrix-assisted Laser Desorption-Ionization Time-of-flight Mass Spectrometry**—CBB-stained proteins were excised from gels, washed with 25% methanol, 7% acetic acid for 12 h at room temperature, and destained with 50 mM  $\text{NH}_4\text{HCO}_3$  in 50% methanol for 1 h at 40 °C. After drying under vacuum, gel spots were incubated in 50  $\mu\text{l}$  of a reduction solution (10 mM EDTA, 10 mM DTT, and 100 mM  $\text{NH}_4\text{HCO}_3$ ) at 60 °C for 1 h. The gel spots were dried under vacuum and incubated in 50  $\mu\text{l}$  of an alkylation solution (10 mM EDTA, 10 mM iodoacetamide, 100 mM  $\text{NH}_4\text{HCO}_3$ ) at room temperature for 30 min in the dark. After washing with water, the gel spots were minced, dried under vacuum, and digested in 50  $\mu\text{l}$  of 10 mM Tris-HCl (pH 8.0) containing 1  $\mu\text{M}$  trypsin (Sigma) at 37 °C for 10 h. Acetonitrile (100  $\mu\text{l}$ ) containing 0.1% trifluoroacetic acid was added to each gel piece and sonicated. Purification of the generated peptides was achieved using Zip-Tips (Millipore, Bedford, MA). The purified peptides (2  $\mu\text{l}$ ) were added directly to a 10 mg  $\text{ml}^{-1}$   $\alpha$ -cyano-4-hydroxycinnamic acid, 0.3% trifluoroacetic acid, 50% acetonitrile matrix and air-dried onto a plate for analysis using MALDI-TOF MS (Voyager-DE RP, Applied Biosystems). Matching of empirical peptide mass values with theoretical digests and sequence information obtained from the data base was performed using Mascot Version 2.0 software (Matrix Science Ltd., London, UK).

For MALDI-TOF analysis, four criteria were used to assign a positive match with a known protein. (i) The deviation between the experimental and theoretical peptide masses should be less than 50

<sup>1</sup> The abbreviations used are: 2D, two-dimensional; GA, gibberellin; CBB, Coomassie Brilliant Blue; VDAC, voltage-dependent anion channel; UPGMA, unweighted pair group method with arithmetic mean; EF, elongation factor; Os, *O. sativa*.

ppm. (ii) At least four different predicted peptide masses needed to match the observed masses for an identification to be considered valid. (iii) The coverage of protein sequences by the matching peptides must reach a minimum of 10%. (iv) The score that was obtained from the analysis with Mascot software indicates the probability of a true positive identification and must be at least 50.

**Clustering Analysis**—Clustering protein time course data for the estimation of interactions were analyzed using the clustering method unweighted pair group method with arithmetic mean (UPGMA).

$$\dot{X}_i = \alpha_i \prod_{j=1}^n X_j^{g_{ij}} - \beta_i \prod_{j=1}^n X_j^{h_{ij}}$$

$X_i$ : i-th protein expression

$\alpha_i$ : i-th protein production velocity coefficient

$g_{ij}$ : coefficient of i-th protein production velocity and j-th protein expression

$\beta_i$ : i-th protein degradation velocity coefficient

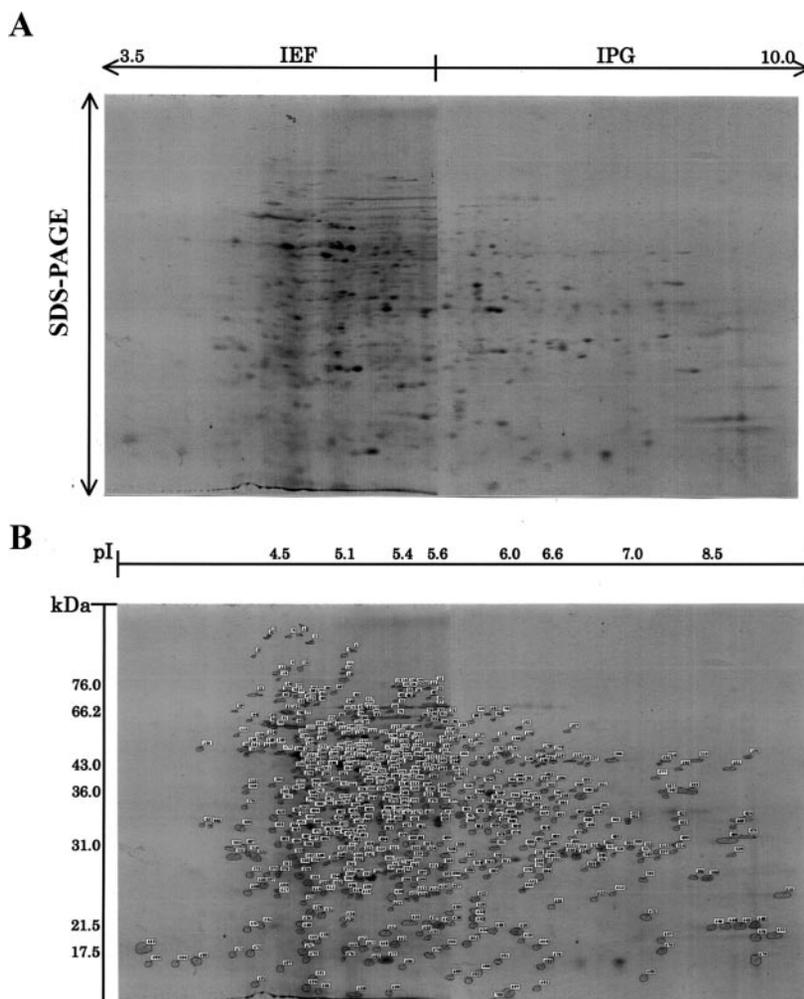
$h_i$ : coefficient of i-th protein degradation velocity and j-th protein expression

FIG. 1. **The formula of S-system.** The differential equation is described in detail under “Experimental Procedures.”

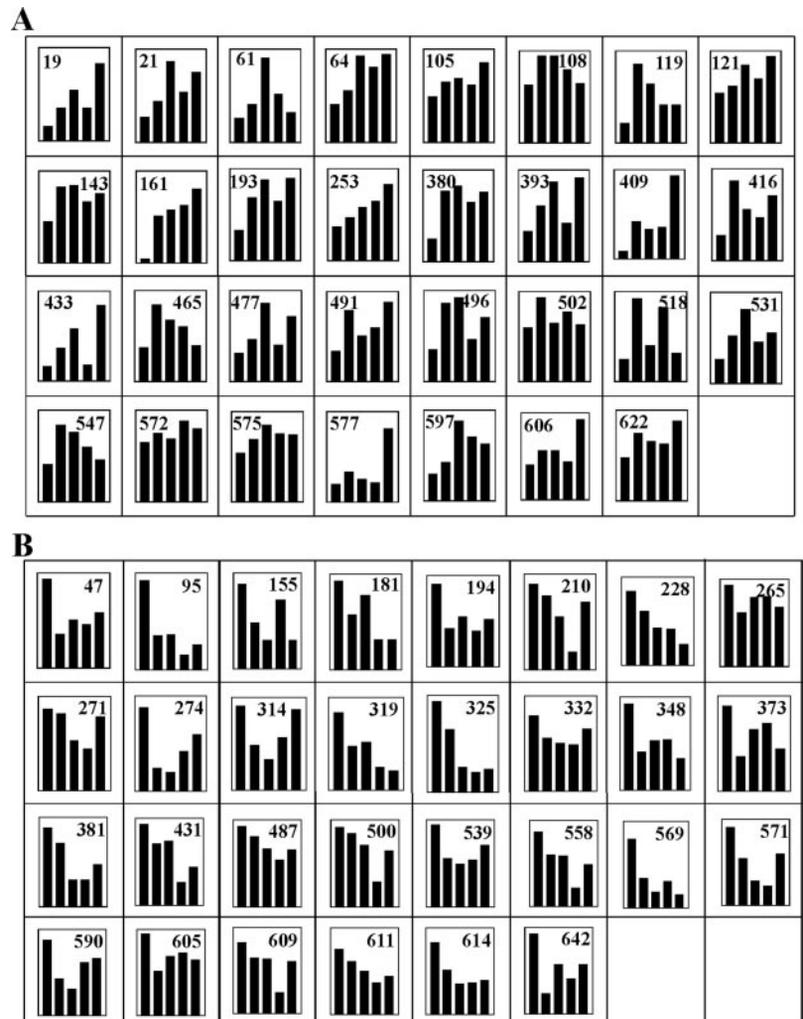
Clustering analysis was performed by two steps as follows. The first step was clustering. Time course-specified clustering analysis was generated in the following way. The time course was normalized with the initial value and then evaluated by the natural logarithm of that normalized value. Next the clustering process in two phases was performed. The first phase was a normal clustering using time course, and the second phase was using fluctuation of time course where fluctuation was evaluated at each time point as the difference between the current and its previous time course. Because the intersection in these two phases of clustering was considered, the cluster was always contained in the same group regardless of the clustering process (using time course or its fluctuation) if the group contained the cluster.

The second step was estimating interaction between clusters. The clusters that interacted were estimated by the representative time course that was calculated at each time point using the medium value. The medium value was calculated using the representative time course for each cluster evaluated in the first step. Mathematical gene interaction network optimization software (Minos) was developed to estimate cluster interaction. Minos utilized the S-system differential equation and estimated cluster interaction by a set of differential equation coefficients that simulates the time course. Fig. 1 shows the formula of the S-system. In the S-system differential equation,  $\alpha_i$  and  $\beta_i$  were positive values. Minos integrated the S-system differential equation and picked out the set of  $\alpha_i$ ,  $g_{ij}$ ,  $\beta_i$ , and  $h_{ij}$  that seemed to fit

FIG. 2. **2D-PAGE pattern of proteins from the basal region in rice seedlings.** Proteins were extracted from the basal region of 2-week-old rice seedlings, separated by 2D-PAGE, and detected by CBB staining (A). The positions of individual proteins on the gels were evaluated with automated Image-Master 2D Elite software (B). Seven hundred proteins were analyzed. In the first dimension, IEF (low pI range) and IPG (high pI range) were used and overlapped at around pI 5.6. The pI and relative molecular mass of each protein were determined using a 2D-PAGE marker (Bio-Rad).



**FIG. 3. Protein expression during development of the rice basal region.** Rice seedlings were grown for 2 (columns 1), 4 (columns 2), 6 (columns 3), 8 (columns 4), and 10 (columns 5) weeks after sowing. Proteins were extracted from the basal region in seedlings, separated by 2D-PAGE, and detected by CBB staining. Following scanning, the gel patterns were analyzed using the 2D Elite software, and the relative abundance ratio of proteins was determined. Compared with the abundance of the 2-week-old seedling sample, increased (A) or decreased (B) proteins after the time point are shown.



to the time course of clusters.  $g_{ij}$  represented the relation between expression of the  $j$ th cluster and the expression generation rate of the  $i$ th cluster. As for  $h_{ij}$ , if  $h_{ij}$  was negative or positive, the formula was increased or decreased, respectively. To estimate cluster interaction,  $g_{ij}$  and  $h_{ij}$  for simulating the time course were calculated.

## RESULTS

**Protein Identification by Two-dimensional Polyacrylamide Gel Electrophoresis**—Building a reference proteome map for the basal region in rice seedlings is a prerequisite for analyzing protein expression during development of the tissue and is constructed by separating the proteins according to their intrinsic charge (pI) and molecular weight. The 2D map of proteins was reproducibly observed in several independent experiments with the high and low pI ranges overlapping at around pI 5.6 (Fig. 2A). The 2D map of the basal region consisted of about 700 proteins identified using Image-Master 2D Elite software (Fig. 2B). The basal region contained a high abundance of proteins in the acidic pI range from around pI 4.5 to 6.0 (Fig. 2B). Using only the IEF (for pI range around pI 3.5–8.0) gel for first dimension separation, 352 protein spots were detected by 2D-PAGE and CBB staining (16),

whereas using the IEF and IPG (for high pI range pI 6.0–10.0) gels, 431 proteins were detected (19, 25). The IEF and IPG gels with gradient SDS-PAGE can resolve proteins over a wide range (from around pI 3.5 to 10). The present study showed 700 protein spots in 2D-PAGE using IEF and IPG gels in the first dimension.

**Identification of Proteins Expressed Differentially during Development**—To identify proteins that are regulated during development, 2D maps of tissue from seedlings grown for 2, 4, 6, 8, and 10 weeks after sowing were constructed (Fig. 2 and data not shown). Observations of triplicate 2D gels of the ~700 proteins detected in the basal region showed that 31 were increased and 30 were decreased relative to expression in 2-week-old seedlings during the observed 8 weeks of development with 10% increasing or decreasing as determined by 2D Elite software, respectively (Fig. 3). These regulated proteins were identified by MALDI-TOF MS and/or protein sequencing (Tables I and II). The N-terminal sequences of about two-thirds of the proteins could not be sequenced, likely due to a blocking group at the N terminus. The percentage of N-terminally blocked proteins encountered in the basal

TABLE I  
Increased proteins during development of basal region in rice seedlings

No. <sup>a</sup>	MM <sup>b</sup>	pI <sup>b</sup>	Identification <sup>c</sup>	Amino acid sequences <sup>d</sup>	Homologous protein	Percentage <sup>e</sup>	Accession no. <sup>f</sup>
	<i>kDa</i>					%	
<b>Metabolism</b>							
119	60.8	4.9	MS	(N-blocked)	Fatty acid $\alpha$ -oxidase	98	AF229813
121	60.5	5.1	MS	(N-blocked)	$\alpha$ -Amylase	49	X6419
253	50.4	5.0	Ed	(N-blocked)	UTP-glucose-1-phosphate uridylyltransferase	100	Q9SDX3
				I-IHTFNQSQY			
380	40.8	6.2	MS	(N-blocked)	Hydrogenase nickel incorporation protein	51	P26409
572	28.3	5.3	MS	(N-blocked)	Nitrate reductase	20	P36858
<b>Energy</b>							
193	54.1	5.1	MS	(N-blocked)	Enolase	114	AY335488
416	38.4	5.4	Ed	(N-blocked)	Malate dehydrogenase	90	Q9SML8
				I-KNITXLTRL			
433	37.4	4.5	Ed	(N-blocked)	Fructokinase	100	P37829
				I-VAFLTQGDAN			
				I-FMFYRNPSAD			
465	35.2	5.2	MS	(N-blocked)	NADH-ubiquinone oxidoreductase	78	AF124786
502	32.5	7.2	Ed	N-VGPGLYPEIG	Voltage-dependent anion channel (OsVDAC1)	100	Y18104
547	30.2	4.5	Ed	(N-blocked)	F <sub>1</sub> -ATPase $\delta$ -subunit	100	D88375
				I-AEPTKN			
<b>Protein synthesis</b>							
105	62.8	4.8	MS	(N-blocked)	Elongation factor Tu	89	AF303468
143	58.0	4.6	Ed	N-SAKEIAFD	FKBP-type peptidyl-prolyl cis-trans isomerase	100	Q8E7Q0
393	39.5	4.5	MS	(N-blocked)	Translation initiation factor 5A	46	AF094773
491	33.2	4.3	Ed	N-AVTFTDLHTA	Elongation factor 1 $\beta$ '	100	P29545
				I-SGKSSVLLDV			
496	32.8	4.4	Ed	N-AVTFTDLHTA	Elongation factor 1 $\beta$ '	100	P29545
				I-SGKSSVLLDV			
518	31.5	4.0	Ed	N-TAAEIIKK	30 S ribosomal protein S17e	75	P54026
606	25.6	4.9	Ed	I-VNIFKDDVVI	Probable BTF3b factor protein	100	A86314
<b>Protein destination and storage</b>							
61	69.2	4.3	Ed	N-EKVVGIDLGT	Heat shock protein 70	100	AY081331
477	34.2	4.9	MS	(N-blocked)	Heat shock protein 78, mitochondria	31	P33416
<b>Cell structure</b>							
409	38.6	7.0	MS	(N-blocked)	$\beta$ -Expansin (EXPB9)	36	AF261277
<b>Signal transduction</b>							
21	79.0	5.7	MS	(N-blocked)	Benzothiadiazole-induced MAP <sup>g</sup> kinase 2	56	AY524973
161	56.0	4.0	Ed	N-EVFFQEKFEED	Calreticulin precursor	100	Q9SP22
				I-FSNKDKTLVL			
575	28.1	4.0	Ed	(N-blocked)	SKP1/ASK-1-like	100	AY316293
				I-SQTIRHMIED			
<b>Defense</b>							
577	28.1	4.9	Ed	N-AVLEVVGL	Antiviral protein 2 precursor	100	Q40772
				I-VTVMTENT			
622	24.5	6.1	Ed	N-AYVDAAGEEV	Abscisic acid stress-inducible protein	100	AF039573
				I-KQHKHKQHLG			
<b>Unclear function</b>							
19	80.0	5.5	MS	(N-blocked)	<i>O. sativa</i> genomic DNA	77	AP005201
64	69.2	5.1	MS	(N-blocked)	ORC4 (origin recognition complex)	67	AB110983
108	61.9	4.9	MS	(N-blocked)	NAC6 <sup>h</sup>	93	AF254558
531	31.0	3.9	MS	(N-blocked)	Nucleoprotein	64	P31609
597	26.5	4.2	Ed	N-MLVYQDLLTG	Translationally controlled tumor protein homolog	100	P50906
				I-GGGDDEGVDD			

<sup>a</sup> The numbers refer to the spot numbers as given in Fig. 1.

<sup>b</sup> Molecular mass (MM) and pI are from the gel in Fig. 1.

<sup>c</sup> Methods of protein identification: MS and Edman degradation (Ed).

<sup>d</sup> N-terminal (N-) and internal (I-) amino acid sequences as determined by Edman degradation.

<sup>e</sup> The values indicate the sequence coverage for MS and the homology for the identity protein sequences for Edman degradation.

<sup>f</sup> Accession numbers in NCBI data base.

<sup>g</sup> Mitogen-activated protein.

<sup>h</sup> N-acetyl-cysteine 6.

TABLE II  
Decreased proteins during development of basal region in rice seedlings

No. <sup>a</sup>	MM <sup>b</sup>	pI <sup>b</sup>	Identification <sup>c</sup>	Amino acid sequences <sup>d</sup>	Homologous protein	Percentage <sup>e</sup>	Accession no. <sup>f</sup>
		<i>kDa</i>				%	
<b>Metabolism</b>							
228	52.0	5.3	MS	(N-blocked)	ADP-ribose pyrophosphatase	64	Q01976
265	49.1	5.5	MS	(N-blocked)	Cytosolic 6-phosphogluconate dehydrogenase	68	AF486280
271	48.5	5.9	MS	(N-blocked)	S-Adenosylmethionine synthetase	73	AJ296743
373	41.1	5.5	Ed	(N-blocked)	RGP1 (reverse glycosylating protein)	100	AJ011078
614	25.1	4.6	MS	(N-blocked)	Methyl-CpG-binding protein	78	P51608
<b>Energy</b>							
194	53.9	5.0	MS	(N-blocked)	Enolase	114	AY335488
274	48.3	6.2	MS	(N-blocked)	Succinate dehydrogenase iron protein	49	AB017428
500	32.6	7.5	Ed	(N-blocked)	Voltage-dependent anion channel (OsVDAC2)	100	AJ251562
539	30.5	5.5	MS		Adenylosuccinate synthase	29	Q97D87
590	27.3	5.8	Ed	N-GVNEDKYTSD	Glyceraldehyde-3-phosphate dehydrogenase	100	P08735
<b>Transcription</b>							
210	52.9	8.1	MS	(N-blocked)	Transcription factor MADS21	44	AY177693
<b>Protein synthesis</b>							
319	44.8	5.6	MS	(N-blocked)	60 S ribosomal protein L19	68	P84098
348	42.8	5.4	MS	(N-blocked)	60 S ribosomal protein L19	68	P84098
381	40.7	4.9	MS	(N-blocked)	50 S ribosomal Protein L33	60	Q8KCPO
558	29.4	4.5	MS	(N-blocked)	60 S ribosomal protein L8	42	O13672
569	28.7	5.5	MS	(N-blocked)	Tyrosyl-tRNA synthetase	26	O29482
<b>Protein destination and storage</b>							
95	63.9	4.4	Ed	(N-blocked)	Heat shock protein	100	M36115
181	54.7	4.8	Ed	(N-blocked)	Heat shock protein	100	M36115
325	44.5	4.4	Ed	N-TAXERAKRTL	<i>Equus caballus</i> heat shock protein	90	AF411802
332	44.0	4.3	Ed	N-STAQTTIEID	Heat shock protein 70.2	100	S48024
<b>Cell structure</b>							
609	25.4	5.0	Ed	I-VRTGDYRQLF	$\alpha$ -Tubulin	100	X91807
<b>Signal transduction</b>							
47	73.3	4.5	Ed	N-SDPLFYEPFD	Calreticulin family	100	CAD40786
155	56.9	5.5	Ed	N-EVFFQEKF	Calreticulin precursor	100	Q9SP22
314	45.0	5.4	MS	(N-blocked)	Signal transduction 54-kDa protein	64	Q57565
487	33.3	4.6	Ed	(N-blocked)	Auxin-responsive protein IAA9	70	Q38827
<b>Defense</b>							
605	25.7	6.5	Ed	(N-blocked)	Chitinase III-like protein	100	BAD31057
642	21.2	5.7	Ed	(N-blocked)	Superoxide dismutase	100	O65768
<b>Secondary metabolism</b>							
571	28.6	5.6	Ed	N-AKNYPVXAE	Putative ascorbate peroxidase	90	AAO17000
611	25.4	4.7	Ed	(N-blocked)	<i>Hordeum vulgare</i> ascorbate peroxidase	80	AF411227
<b>Unclear function</b>							
431	37.6	5.2	MS	(N-blocked)	Abnormal long morphology protein 1	65	ALMI-SCHPO

<sup>a</sup> The numbers refer to the spot numbers as given in Fig. 1.

<sup>b</sup> Molecular mass (MM) and pI are from the gel in Fig. 1.

<sup>c</sup> Methods of protein identification: MS and Edman degradation (Ed).

<sup>d</sup> N-terminal (N-) and internal (I-) amino acid sequences as determined by Edman degradation.

<sup>e</sup> The values indicate the sequence coverage for MS and the homology for the identity protein sequences for Edman degradation, respectively.

<sup>f</sup> Accession numbers in NCBI data base.

region was 64%. Of the proteins found to increase in expression during development, 65% (20 of 31) and 63% (19 of 31) of proteins showing decreased expression were N-terminally

blocked. This result was consistent with a previous study that found that of 134 rice proteins, 79 (59%) were found to have blocked N termini (15).

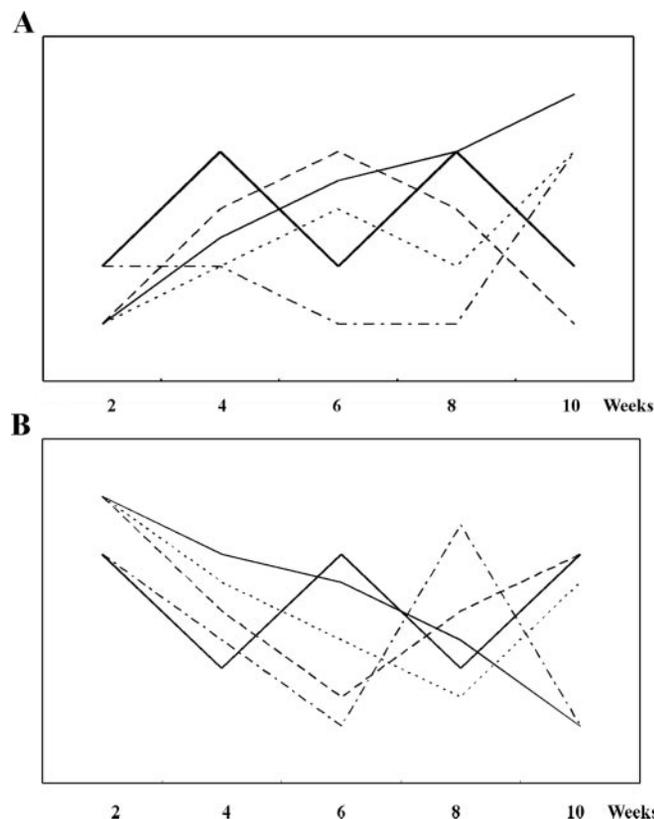


FIG. 4. The representative patterns of protein accumulation during development of the rice basal region. The representative protein expression patterns during developmental stage of increased (A) and decreased (B) proteins are shown from Fig. 3. Details of the expression patterns i–v were described in text. A, expression patterns i (thin line), ii (dashed line), iii (dotted line), iv (dashed and dotted line), and v (thick line). B, expression patterns i (thin line), ii (dashed line), iii (dotted line), iv (dashed and dotted line), and v (thick line).

The developmentally regulated basal region proteins had different time courses of expression. The representative trend lines are shown in Fig. 4. Five different time course expression patterns could be discerned for proteins with increased expression (Fig. 4A): (i) a continuous increase (Numbers 161 and 253), (ii) an increase from 2 to 6 weeks with a decrease from 8 and 10 weeks (Numbers 61, 108, 119, 465, 547, 575, and 597), (iii) an increase from 2 to 6 weeks, a decrease at 8 weeks, and an increase at 10 weeks (Numbers 19, 21, 64, 105, 121, 143, 193, 380, 393, 433, 477, 496, 531, and 606), (iv) no increase until 4 weeks, a decrease from 6 to 8 weeks, and an increase at 10 weeks (Numbers 409, 416, 491, 577, and 622), and (v) an alternating increase and decrease at successive time points (Numbers 502, 518, and 572).

Proteins that decreased during development had five different expression patterns (Fig. 4B): (i) a continuous decrease (Numbers 228 and 319), (ii) a decrease from 2 to 6 weeks with an increase from 8 to 10 weeks (Numbers 274, 314, 500, 539, and 590), (iii) a decrease for the first 8 weeks with an increase in the final week (Numbers 210, 271, 325, 332, 381, 487, 558, 571, 609, 611, and 614), (iv) a decrease through the first few

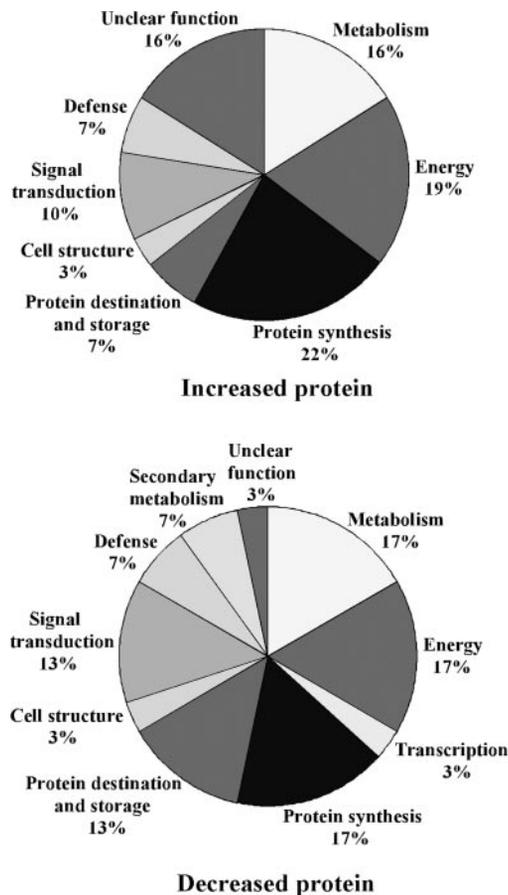


FIG. 5. Assignment of the identified proteins to functional categories. After assessment of each protein using the classifications described by Bevan *et al.* (1998), the total number and the percentage of proteins in each category were determined. The distribution of proteins in each functional category is shown for proteins that increased and decreased during development of basal region in rice seedlings based on the total number of proteins identified.

weeks with a sudden increase at 6 or 8 weeks followed by a decrease at 10 weeks (Numbers 155, 181, 265, 348, 373, 605, and 642), and (v) an alternating decrease and increase at successive time points (Numbers 47, 95, 194, 431, 569, and 642).

Abundant proteins were identified, and their functions were determined using the classifications by Bevan *et al.* (35). The developmentally regulated proteins were functionally involved with primary metabolism, transcription, protein synthesis, transport or targeting and storage, signal transduction, energy, cell structure, defense, secondary metabolism, and functions that have not been defined (Fig. 5 and Tables I and II). Elongation factors (elongation factor Tu, Number 105; elongation factor 1 $\beta$ ', Numbers 491 and 496) increased, and ribosomal proteins (60 S ribosomal protein L19, Numbers 319 and 348; 50 S ribosomal protein L33, Number 381; 60 S ribosomal protein L8, Number 558) decreased during basal region development. The cell structure protein  $\beta$ -expansin (Number 409) increased, whereas  $\alpha$ -tubulin (Number 609)

TABLE III  
Differentially regulated proteins during development of basal region in rice seedlings

No. <sup>a</sup>	MM <sup>b</sup>	pI <sup>b</sup>	Identification <sup>c</sup>	Amino acid sequences <sup>d</sup>	Homologous protein	Regulation <sup>e</sup>	Percentage <sup>f</sup>	Accession no. <sup>g</sup>
		<i>kDa</i>						
		%						
155	56.9	5.5	Ed	N-EVFFQEKF	Calreticulin precursor	Down	100	Q9SP22
161	56.0	4.0	Ed	N-EVFFQEKFED I-FSNKDKTLVL	Calreticulin precursor	Up	100	Q9SP22
193	54.1	5.1	MS	(N-blocked)	Enolase	Up	114	AY335488
194	53.9	5.0	MS	(N-blocked)	Enolase	Down	114	AY335488
500	32.6	7.5	Ed	(N-blocked)	Voltage-dependent anion channel	Down	100	AJ251562
502	32.5	7.2	Ed	I-VQYLHPHAGL N-VGPGLYPEIG I-LQYSHDYAGV	Voltage-dependent anion channel	Up	100	Y18104

<sup>a</sup> The numbers refer to the spot numbers as given in Fig. 1.

<sup>b</sup> Molecular mass (MM) and pI are from the gel in Fig. 1.

<sup>c</sup> Methods of protein identification: MS and Edman degradation (Ed).

<sup>d</sup> N-terminal (N-) and internal (I-) amino acid sequences as determined by Edman degradation.

<sup>e</sup> Direction of change in protein level during development of the basal region: Up, increased; Down, decreased.

<sup>f</sup> The values indicate the sequence coverage for MS and the homology for the identity protein sequences for Edman degradation, respectively.

<sup>g</sup> Accession numbers in NCBI data base.

decreased. The heat shock proteins, which have a range of functions including cell targeting and storage, both increased and decreased.

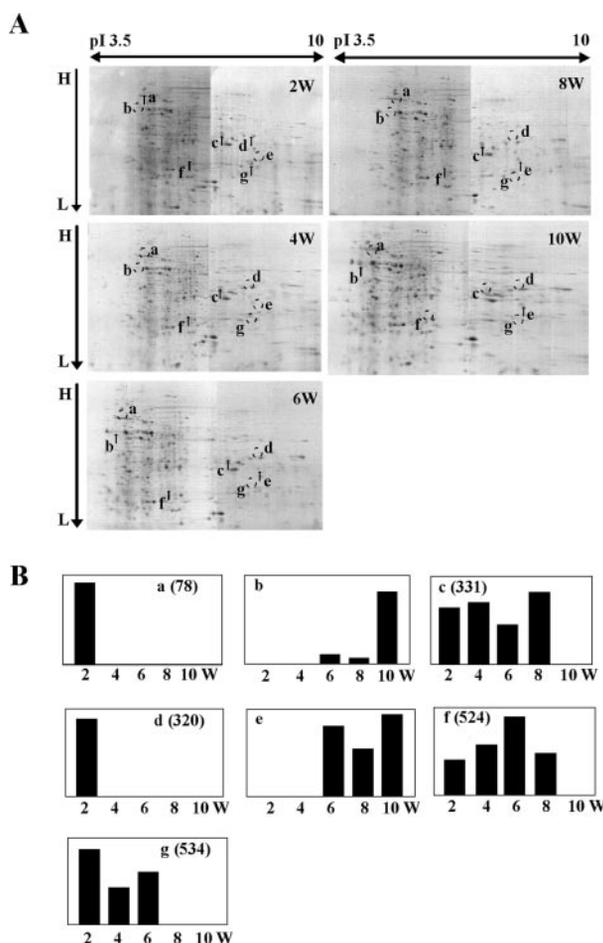
Several proteins were identified as the same protein having the same molecular weights but different pI values that were expressed differentially during development (Table III). Calreticulin precursor with pI 5.5 (Number 155) and pI 4.0 (Number 161) decreased and increased, respectively, during development of the basal region. Two other sets having different pI spots were enolase, pI 5.1 (Number 193) and pI 5.0 (Number 194), increased and decreased, respectively, and voltage-dependent anion channel (VDAC) protein, pI 7.5 (Number 500) and pI 7.2 (Number 502), decreased and increased, respectively. In a previous report, twin protein spots identified as calreticulin had isoelectric points (pI values of 4.0 and 4.3) and expression levels that were changed by gibberellin (GA<sub>3</sub>) treatment of the leaf sheath (25). Expression of the pI 4.0 protein was down-regulated, whereas the pI 4.3 protein was up-regulated when the leaf sheath was treated with GA<sub>3</sub>.

**Identification of Basal Region Developmentally Repressed and de Novo Proteins**—Five proteins that were completely repressed and two that appeared *de novo* during development were also identified (Fig. 6A), and the relative amounts were estimated (Fig. 6B). Of the five repressed proteins, two were detected only at week 2 (Numbers 78 and 320), two were absent at 10 weeks (Numbers 331 and 524), and one was absent from 8 weeks onward (Number 534). These proteins were heat shock protein 82, an as yet unidentified protein, putative GTP-binding protein, cellulose synthase-like, and RSSG8 mRNA (Table IV). The two *de novo* proteins detected from 6 weeks on were calreticulin precursor (Fig. 6B, spot b) and  $\beta$ -tubulin (spot e).

**Cluster Analysis of Basal Region Proteins**—Proteins that were regulated during development of the basal region but

whose functions were unknown were also tentatively identified by cluster analysis (Tables I and II). Because the time course data of five points (2, 4, 6, 8, and 10 weeks) has very few constraints with 61 regulated proteins, clustering protein time course data for the estimation of the interaction was used. Using this analysis, proteins that are regulated by other proteins or proteins that regulate another protein are classified into the same cluster. When proteins are classified into the same cluster, they are assumed to have the same regulation as other members of the cluster. The 31 developmentally up-regulated and 30 down-regulated proteins were analyzed by the UPGMA clustering method, and 32 clusters were found. Interactions of these clusters were estimated by the S-system (Fig. 7).

The unknown protein represented by spot Number 19, which increased during development, was identified as *O. sativa* genomic DNA (Table I), indicating that the protein was expressed from an ORF in the published genomic DNA sequence. Protein Number 19 negatively regulated Cluster 29 (Number 228, ADP-ribose pyrophosphatase) but was regulated both negatively and positively by Cluster 30 (Number 181, Rubisco-binding protein  $\beta$ -subunit; and Number 319, 60 S ribosomal protein L19) and by Clusters 13 (Number 577, antiviral protein 2 precursor) and 15 (Number 433, fructokinase), respectively (Fig. 5). Protein Number 64, ORC4 (origin recognition complex), belonged to Cluster 10 with enolase (Number 193), UTP-glucose-1-phosphate uridylyltransferase (Number 253), and elongation factor 1 $\beta$ ' (Number 496). Cluster 10 was regulated by Cluster 17 (Number 569, tyrosyl-tRNA synthase). Protein Number 108, NAC6, was part of Cluster 4, which did not interact with any other clusters. Protein Number 531, a nucleoprotein constituent of Cluster 12, was negatively regulated by Cluster 16 (Number 155, calreticulin precursor). Protein Number 597 of Cluster 11 translationally controlled tumor protein homolog, positively regulated Cluster 27, neg-



**FIG. 6. Dramatically regulated proteins during development of the rice basal region.** Proteins extracted from the basal region of 2-, 4-, 6-, 8-, and 10-week-old seedlings were separated by 2D-PAGE (A) as described in Fig. 2. H, high molecular weight; L, low molecular weight. During development of the basal region, proteins disappeared (spots a, c, d, f, and g; shown as arrows in 2D-PAGE map) or proteins that were not detected in the 2-week-old sample appeared (spots b and e; shown as dotted circles). The proteins detected in the 2-week-old sample are indicated by spot numbers in parentheses, whereas the proteins that did not appear in the 2-week-old sample (spots b and e) are indicated by the spot letter only. Compared with the abundance of the 2-week-old seedling sample, regulated proteins after the time point are shown (B).

actively regulated Clusters 6 and 31, and was negatively regulated by Cluster 17. Among the down-regulated proteins of unknown function, Number 431, abnormal long morphology protein, was a member of Cluster 32 along with transcription factor MADS21 (Number 210), voltage-dependent anion channel (Number 500), 60 S ribosomal protein L8 (Number 558), and  $\alpha$ -tubulin (Number 609). Cluster 32 regulated Clusters 18 and 21 positively and was negatively regulated by Cluster 21.

#### DISCUSSION

**The Basal Region and Leaf Sheath Proteome**—Proteins from leaf sheaths, including the basal region, of seedlings

have been analyzed because the tissue was found to be important in the regulation of the height of rice plants (16, 19, 25). The detected protein spots in the basal region were more abundant than in leaf sheath because the leaf sheath proteins form a subset of the total basal region proteome. Because the same weights of leaf sheath and basal regions were extracted and the leaf sheath contained more water than the basal region, extracts from the leaf sheath contained less protein mass than the basal region extracts. Few of the developmentally regulated proteins corresponded with proteins from the leaf sheath proteome data base, suggesting that the developmentally regulated proteins had low or non-existent expression levels in the leaf sheath.

In this study, the protein from the basal region was extracted with a lysis buffer according to the method of O'Farrell (33). The lysis buffer has been well established for soluble protein extraction from rice tissues (9). Therefore we thought that the lysis buffer would be suitable for protein extraction from the basal region.

**Function of Developmentally Regulated Proteins**—The proteins that are regulated during development of the basal region were classified according to function. This is a helpful tool for visualizing the general distribution of functions served by the proteins of a given tissue. The largest portion of protein function was dedicated to central metabolic pathways, energy production, and protein synthesis and maintenance (Fig. 5 and Tables I and II).

Cell elongation and expansion were important steps for development of the basal region. Expansin (Number 409) was identified because of its ability to restore long term extension to cell walls that were isolated from growing plant tissues. The mechanism of expansin action appears to involve disruption of the hydrogen bonds between cellulose microfibrils and cross-linking glycans (36). This function is supported by the observation that expansin increased almost continuously during development of the basal region (Fig. 3A). However,  $\alpha$ -tubulin (Number 609), a cytoskeletal component, decreased during development (Fig. 3B), and cellulose synthase-like protein (Number 524) was completely repressed by 10 weeks in seedlings (Fig. 6B). These proteins were thought to increase during development because of their importance in cell growth. These results suggest that cell growth and the proteins that regulate cell structure are complex mechanisms, that the presence or absence of individual proteins in a given tissue cannot be surmised based on putative functions, and that much more work is required to fully understand the role of regulatory factors in tissue and organ formation.

Protein synthesis regulation is also an important factor during development. Elongation factors EF-Tu (Number 105) and EF-1 $\beta$ ' (Numbers 491 and 406) increased during development (Table I), whereas ribosomal proteins decreased (Table II). These contradictory results of the intuitive role of protein-synthesizing components indicate that further research will be needed specifically on the interactions and downstream ef-

TABLE IV  
Dramatically regulated proteins during development of the rice basal region

No. <sup>a</sup>	MM <sup>b</sup>	pI <sup>b</sup>	Identification <sup>c</sup>	Amino acid sequences <sup>d</sup>	Homologous protein	Percentage <sup>e</sup>	Accession no. <sup>f</sup>
	<i>kDa</i>					%	
a (78)	66.8	4.4	MS	(N-blocked)	Heat shock protein 82	71	Z11920
b			Ed	N-EVFFQEKFED	Calreticulin precursor	100	Q9SP22
c (331)	44.1	6.2	Ed	N-DDVLALTEST	Putative GTP-binding protein	100	Q09654
d (320)	44.8	7.0	MS	(N-blocked)	Not hit		
e			MS	(N-blocked)	$\beta$ -Tubulin 7	90	L10634
f (524)	31.5	5.6	MS	(N-blocked)	Cellulose synthase-like (CSLC2)	37	AF435650
g (534)	31.0	7.0	MS	(N-blocked)	Guichao2 RSSG8 mRNA	90	AY034141

<sup>a</sup> Spots a–g are shown in Fig. 2. The numbers in parentheses refer to the spot numbers as given in Fig. 1.

<sup>b</sup> Molecular mass (MM) and pI are from the gel in Fig. 1.

<sup>c</sup> Methods of protein identification: MS and Edman degradation (Ed).

<sup>d</sup> N-terminal (N-) and internal (I-) amino acid sequences as determined by Edman degradation.

<sup>e</sup> The values indicate the sequence coverage for MS and the homology for the identity protein sequences for Edman degradation, respectively.

<sup>f</sup> Accession numbers in NCBI data base.

fects of the regulated proteins in protein synthesis during development.

Changes in protein levels alone do not imply changes in function as many proteins are known to be regulated by post-translational modifications. It is possible that post-translational modifications, such as phosphorylation, may affect in-gel migration of a protein such that a protein classified as repressed may be present elsewhere on the gel or that a putative *de novo* protein may be found in a new place in the array. As shown by the present study, it is possible to precisely identify the modulation of calreticulin precursor during development of the basal region (Table III). Calreticulin precursor showed an expected molecular mass of 56 kDa and was represented by twin spots on 2D-PAGE (Numbers 155 and 161) having pI values of 5.5 and 4.0, respectively (Fig. 3). Furthermore this protein was detected as a developmentally specific protein and was located between Numbers 155 and 161 on the gel (Fig. 6B, spot b). Calreticulin shifted from Number 155 to spot b and Number 161, or from a high pI spot to low pI spots. In a previous study, when leaf sheaths were elongated by GA<sub>3</sub> treatment, calreticulin shifted from pI 4.0 (low pI) to pI 4.3 (high pI) (25). The calreticulin shifts observed during development of the basal region and GA-stimulated elongation of leaf sheath, however, were in opposite directions. A previous study has shown the involvement of *in vivo* phosphorylation activity of calreticulin (pI 4.3) in long term rice cultured suspension cells but not in short term cultured cells (37). These findings suggest that a transient state of calreticulin plays an important role in signal transduction cascades by affecting Ca<sup>2+</sup> homeostasis during developmental regulation. Enolase has also been identified as twin spots (Table III). Several genes for enolase have been found in public data bases (NCBI or DDBJ). It is indicated that the twin spots are different members of the same gene family. Enolase is a ubiquitous enzyme that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate, the only dehydration

step in the glycolytic pathway. Expression of the enolase gene in tomato was significantly higher in root than in green tissue and was higher in ripened pink fruit than in green fruit. In *Arabidopsis*, enolase gene expression was constant in different tissues (38). In the present study, the enolases having pI 5.1 (Number 193) and pI 5.0 (Number 194) increased and decreased, respectively, during development (Fig. 3). Because glycolysis plays a role in energy metabolism, differential expression of enolase isoforms would regulate the metabolism during development of the basal region when the tissue requires energy. In *Arabidopsis*, enolase was localized to the cytosolic face of the outer mitochondria membrane (39), suggesting that the enolase increased during development (Number 193) is attached to mitochondria and is functional. VDAC is a mitochondrial outer membrane ion channel that allows the diffusion of negatively charged solutes like succinate, malate, and ATP in the fully open state. Expression of three of the four identified genes of the rice *vdac* family (*Osvdac1*, *Osvdac2*, and *Osvdac3*) were investigated (35). *Osvdac1* and *Osvdac3* were differentially expressed, whereas *Osvdac2* was constitutively expressed in all tissue. Based on N-terminal amino acid sequences, the VDAC having a pI of 7.2 (Number 502) corresponded to OsVDAC1 and increased during development. The VDAC having a pI of 7.5 (Number 500) corresponded to OsVADC2 and decreased during development. The differential expression caused by developmental regulation of rice VDAC in the basal region is most probably coordinated by differential regulation of the genes. It is suggested that differentially accumulated isoform proteins play an important role during development of the basal region.

Seven proteins had developmentally specific expression (Fig. 6 and Table IV). The functions of the proteins were: chaperone, signal transduction, cell structure, and unknown. HSP82 accumulated only in the 2-week-old sample, suggesting that HSP82 plays a role in protein folding and maintenance in an early developmental stage. A cytoskeletal component,

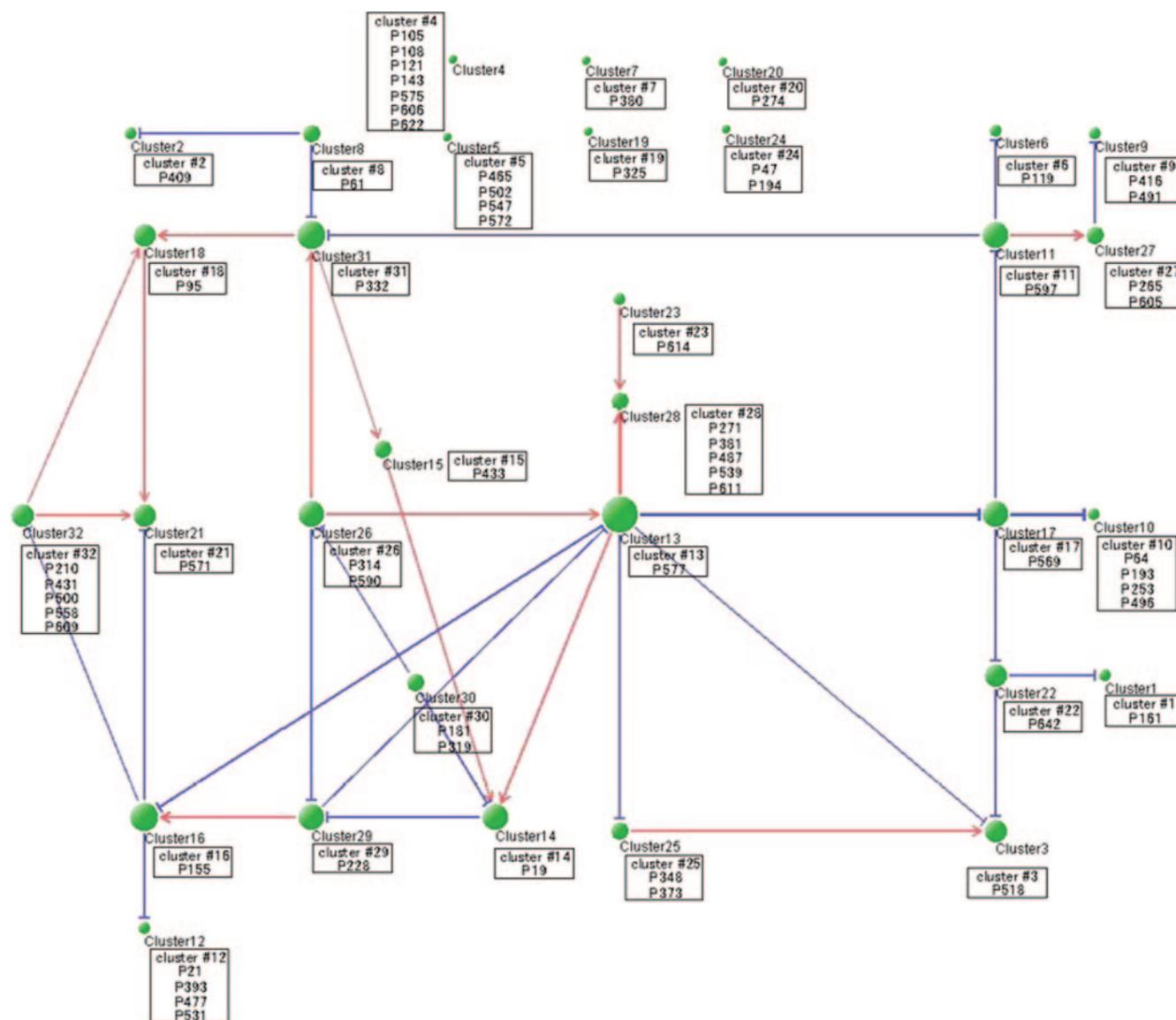


FIG. 7. Clustering analysis of the regulated proteins during development of the basal region in rice seedlings. The 30 proteins that increased and the 31 proteins that decreased during development were clustered by the UPGMA method. The interactions of these clusters were analyzed by the S-system. Arrows and T-bars show up-regulation and down-regulation, respectively. The size of each cluster as shown by circles represents the number of interactions.

$\beta$ -tubulin 7, increased during development. As discussed above, cytoskeletal and cell structure component  $\alpha$ -tubulin and cellulose synthase-like protein were differentially regulated. Because only those regulated proteins that underwent major changes during development were analyzed, proteins that are subject to more subtle regulation will need to be studied to understand in detail the mechanisms of basal region development.

*Estimation of Proteins of Unknown Function by Cluster Analysis*—Although a growing number of comparative proteomic studies have been reported in plant systems, the grouping of proteins into expression classes has generally been qualitative, and rigorous quantitative measures have

been lacking. In the present study, a clustering approach was used to determine relationships in protein expressions (Fig. 7). Using clustering analysis, identified proteins with an unknown function can be related to other proteins whose functions have been determined. A protein that was found to increase during development of the basal region was identified only as *O. sativa* genomic DNA. This protein (Number 19) was up-regulated by Cluster 13 (Number 577, antifungal protein 2) and Cluster 15 (Number 433, fructokinase) and was down-regulated by Cluster 30 (Number 181, Rubisco-binding protein  $\beta$  subunit and 60 S ribosomal protein L19). On the other hand, Number 19 protein negatively regulated Cluster 29 (Number 228, ADP-ribose pyrophosphatase). These results

suggest that Number 19 protein is related to pathogen defense, sugar metabolism, and protein maintenance. Abnormal long morphology protein (Number 431 in Cluster 32) and nucleoprotein (Number 531 in Cluster 12) were regulated negatively by calreticulin precursor (Number 155 in Cluster 16), suggesting that these proteins were regulated under  $Ca^{2+}$  signaling. ORC4 (Number 64 in Cluster 10) and translationally controlled tumor protein homolog (Number 597 in Cluster 11) were negatively regulated by tyrosyl-tRNA synthase (Number 569 in Cluster 17) and may be directly related to protein synthesis. Because the other functionally unknown protein NAC6 (Number 108 in Cluster 4) did not interact with other clusters, it is not possible to extrapolate its function.

In conclusion, the technique of differential display of proteins with 2D-PAGE presents a powerful approach to the study of complex patterns of protein expression under different developmental stages, and cluster interaction analyses based on the S-system will be invaluable for the identification of protein functions. This approach will be applied to resolving the interactions for proteins and will lead to identifying the roles for proteins involved in rice plants development.

**Acknowledgments**—We thank Dr. K. Higo for critical suggestions and discussions. We also thank Dr. G. Yang for discussions and Dr. S. Shen and Amersham Biosciences for technical support.

\* This work was supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences and by a grant from the Japanese Ministry of Agriculture, Forestry, and Fisheries Rice Genome Project (Japan). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Fax: 81-298-38-7464; E-mail: skomatsu@affrc.go.jp.

#### REFERENCES

- Devos, M. K., and Gale, D. M. (2000) Genome relationships: the grass model in current research. *Plant Cell* **12**, 637–646
- Goff, S. A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B. M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.-I., Chen, L., Cooper, B., Park, S., Wood, T. C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldregde, G., Scholl, T., Miller, R. M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., and Briggs, S. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100
- Yu, J., Hu, S., Wang, J., Wong, G. K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L., and Yang, H. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92
- Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y., Itonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J., Kim, H.-L., Eun, M.-Y., Yano, M., Jiang, J., and Gjobori, T. (2002) The genome sequence and structure of rice chromosome 1. *Nature* **420**, 312–316
- Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J., Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y., Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L. S., Yu, Z., Fan, D., Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J., Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R., Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J., Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q., Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S., Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J., Li, J., Hong, G., Xue, Y., and Han, B. (2002) Sequence and analysis of rice chromosome 4. *Nature* **420**, 316–320
- Pandy, A., and Mann, M. (2000) Proteomics to study genes and genomes. *Nature* **405**, 837–846
- Ressignol, M. (2001) Analysis of the plant proteome. *Curr. Opin. Biotechnol.* **12**, 131–134
- van Wijk, K. J. (2001) Challenges and prospects of plant proteomics. *Plant Physiol.* **126**, 501–508
- Rakwal, R., and Agrawal, G. K. (2003) Rice proteomics: current status and future perspectives. *Electrophoresis* **24**, 3378–3389
- Komatsu, S., and Tanaka, N. (2005) Rice proteome analysis: a step toward functional analysis of the rice genome. *Proteomics* **5**, 938–949
- 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**, 935–942
- 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
- Komatsu, S., Muhammad, A., and Rakwal, R. (1999) Separation and characterization of proteins from green and etiolated shoots of rice (*Oryza sativa* L.): towards a rice proteome. *Electrophoresis* **20**, 630–636
- Komatsu, S., Rakwal, R., and Li, Z. (1999) Separation and characterization of proteins of rice (*Oryza sativa* L.) suspension cultured cells. *Plant Cell Tissue Organ Cult.* **55**, 183–192
- Tsugita, A., Kawakami, T., Uchiyama, Y., Kamo, M., Miyatake, N., and Nozu, Y. (1994) Separation and characterization of rice proteins. *Electrophoresis* **15**, 708–720
- Shen, S., Matsubae, M., Takao, T., Tanaka, N., and Komatsu, S. (2002) A proteomic analysis of leaf sheath from rice. *J. Biochem.* **132**, 613–620
- Imin, N., Kerim, T., Weinman, J. J., and Rolfe, B. G. (2001) Characterization of rice anther proteins expressed at the young microspore stage. *Proteomics* **1**, 1149–1161
- Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L., Deciu, C., Haynes, P. A., Hays, L., Schieltz, D., Ulaszek, R., Wei, J., Wolters, D., and Yates, J. R. (2002) Proteomic survey of metabolic pathway in rice. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11969–11974
- Tanaka, N., Konishi, H., Khan, M. M. K., and Komatsu, S. (2004) Proteome analysis of rice tissues by two-dimensional electrophoresis: an approach to the investigation of gibberellin regulated proteins. *Mol. Genet. Genomics* **270**, 485–496
- Mikami, S., Hori, H., and Mitsui, T. (2001) Separation of distinct compartments of rice Golgi complex by sucrose density gradient centrifugation. *Plant Sci.* **161**, 665–675
- Heazlewood, J. L., Howell, K. A., Whelan, J., and Millar, A. H. (2003) Towards an analysis of the rice mitochondrial proteome. *Plant Physiol.* **132**, 230–242

22. Tanaka, N., Fujita, M., Handa, H., Murayama, S., Uemura, M., Kawamura, Y., Mitsui, T., Mikami, S., Tozawa, Y., Yoshinaga, T., and Komatsu, S. (2004) Proteomics of the rice cell: systematic identification of the protein population in subcellular compartments. *Mol. Genet. Genomics* **271**, 566–576
23. Rakwal, R., and Komatsu, S. (2000) Role of jasmonate in the rice (*Oryza sativa* L.) self-defense mechanism using proteome analysis. *Electrophoresis* **21**, 2492–2500
24. Konishi, H., and Komatsu, S. (2004) A proteomics approach to investigating promotive effects brassinolide on lamina inclination and root growth in rice seedlings. *Biol. Pharm. Bull.* **26**, 401–408
25. Shen, S., Sharma, A., and Komatsu, S. (2003) Characterization of proteins responsive to gibberellin in the leaf-sheath of rice (*Oryza sativa* L.) seedling using proteome analysis. *Biol. Pharm. Bull.* **26**, 129–136
26. Abbasi, F. M., and Komatsu, S. (2004) A proteomic approach to analyze salt-responsive proteins in rice leaf sheath. *Proteomics* **4**, 2072–2081
27. Salekdeh, G. H., Siopongco, J., Wade, L. J., Ghareyazie, B., and Bennett, J. (2002) Proteomic analysis of rice leaves during drought stress and recovery. *Proteomics* **2**, 1131–1145
28. Agrawal, G. K., Rakwal, R., Yonekura, M., Kubo, A., and Saji, H. (2002) Proteome analysis of differentially displayed proteins as a tool for investigating ozone stress in rice (*Oryza sativa* L.) seedlings. *Proteomics* **2**, 947–959
29. Konishi, H., Ishiguro, K., and Komatsu, S. (2001) A proteomics approach towards understanding blast fungus infection of rice grown under different levels of nitrogen fertilization. *Proteomics* **1**, 1162–1171
30. Kim, S. T., Cho, K. S., Yu, S., Kim, S. G., Hong, J. C., Han, C.-d., Bae, D. W., Nam, M. H., and Kang, K. Y. (2003) Proteomic analysis of differentially expressed proteins induced by rice blast fungus and elicitor in suspension-cultured cells. *Proteomics* **3**, 2368–2378
31. Ventelon-Debout, M., Delalande, F., Brizard, J.-P., Diemer, H., Van Dorselaer, A., and Brugidou, C. (2004) Proteome analysis of cultivar-specific deregulation of *Oryza sativa indica* and *O. sativa japonica* cellular suspensions undergoing Rice yellow mottle virus infection. *Proteomics* **4**, 216–225
32. Lonhosky, P. M., Zhang, X., Honavar, V. G., Dobbs, D. L., Fu, A., and Rodermeil, S. R. (2004) A proteomic analysis of maize chloroplast biogenesis. *Plant Physiol.* **134**, 560–574
33. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021
34. Cleveland, D. W., Fisher, S. G., Kirschner, M. W., and Laemmli, U. K. (1977) Peptide mapping by limited proteolysis in sodium dodecyl sulphate and analysis by gel electrophoresis. *J. Biol. Chem.* **252**, 1102–1106
35. Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C., Bergkamp, R., Dirkse, W., van Staveren, M., Stiekema, W., Drost, L., Ridley, P., Hudson, S. A., Patel, K., Murphy, G., Piffanelli, P., Wedler, H., Wedler, E., Wambutt, R., Weitzenegger, T., Pohl, T. M., Terryn, N., Gielen, J., Villarroel, R., De Clerck, R., Van Montagu, M., Lecharry, A., Auborg, S., Gy, I., Kreis, M., Lao, N., Kavanagh, T., Hempel, S., Kotter, P., Entian, K. D., Rieger, M., Schaeffer, M., Funk, B., Mueller-Auer, S., Silvey, M., James, R., Montfort, A., Pons, A., Puigdomenech, P., Douka, A., Voukelatou, E., Milioni, D., Hatzopoulos, P., Piravandi, E., Obermaier, B., Hilbert, H., Dusterhofs, A., Moores, T., Jones, J. D. G., Eneva, T., Palme, K., Benes, V., Rechman, S., Ansoerge, W., Cooke, R., Berger, C., Delseny, M., Voet, M., Volckaert, G., Mewes, H. W., Klosterman, S., Schueller, C., and Chalwatzis, N. (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* **391**, 485–488
36. Li, Y., Jones, L., and McQueen-Mason, S. (2003) Expansins and cell growth. *Curr. Opin. Plant Biol.* **6**, 603–610
37. Komatsu, S., Masuda, T., and Abe, K. (1996) Phosphorylation of a protein (pp56) is related to the regeneration of rice cultured suspension cells. *Plant Cell Physiol.* **37**, 748–753
38. Van Der Straeten, D., Rodrigues-Pousasa, R. A., Goodman, H. M., and Van Montagu, M. (1991) Plant enolase: gene structure, expression, and evolution. *Plant Cell* **3**, 719–735
39. Giege, P., Heazlewood, J. L., Roessner-Tunali, U., Millar, A. H., Fernie, A. R., Leaver, C. J., and Sweetlove, L. J. (2003) Enzymes of glycolysis are functionally associated with the mitochondrion in Arabidopsis cells. *Plant Cell* **15**, 2140–2151
40. Deleted in proof