JARQ 25, 223–229 (1992)

Proteins Encoded by Genes Linked With a Semidwarfing Gene in Rice

Hisashi HIRANO*, Atsushi NAKAMURA, Fumio KIKUCHI** and Setsuko KOMATSU***

* Department of Molecular Biology, National Institute of Agrobiological Resources (Tsukuba, Ibaraki, 305 Japan)
** Institute of Plant Breeding, University of Tsukuba (Tsukuba, Ibaraki, 305 Japan)

Abstract

Two-dimensional gel electrophoresis (2D-PAGE) patterns of seed embryos were compared between a long culm cultivar, Norin 29 and its semidwarf near-isogenic line, SC-TN1, in rice (*Oryza sativa L.*), and thereby two proteins apparently related to semidwarfism were identified.
The proteins found in Norin 29 and SC-TN1 were designated as semidwarfism-related protein (SRP)-1 and SRP-2, respectively. The crossing tests showed that SRP-1 and SRP-2 were controlled by codominant alleles at a single locus, and that this locus was linked with the semidwarfing gene *sd-1* locus. The partial amino acid sequences determined were highly homologous with those of storage protein glutelin α subunits in the seed endosperms. In the seed endosperms of Norin 29 and SC-TN1, a total of 8 glutelin α subunits were identified by 2D-PAGE. The amino acid sequences of SRP-1 and SRP-2 were identical with the glutelin α5a-and α5b-subunits, respectively, which existed in both endosperms and embryos unlike the other glutelin α subunits.

Discipline: Biotechnology
Additional key words: amino acid sequence, gel electrophoresis, glutelin, seed embryo protein, seed endosperm protein

Introduction

A breakthrough in rice breeding was attained with the development of semidwarf cultivars characterized by lodging resistance, nitrogen responsiveness and erect leaves. The success of the "Green Revolution" is directly related to the intensive use of those semidwarfs. To date, a great number of semidwarf cultivars have been developed for practical use, including Taichung Native 1, IR 8, Ai-jio-nan-te, Shiranui through a conventional cross-breeding, and Calrose 76, Reimei through a mutation breeding. It is now apparent that the semidwarfing genes carried by those semidwarf cultivars with economic importance are present at the same locus as the cv. Dee-geo-woo-gen gene, *sd-1* locus, inspite of their different origins[1,3,8].

If the *sd-1* gene could be cloned, it would be made available to develop semidwarf transgenic rice plants. One of the widely used and reliable techniques to clone the *sd-1* gene might be the following steps based on protein analysis. First, a protein which may contribute to semidwarfism is identified; second, the partial amino acid sequence of the protein is determined; and finally, a DNA encoding the protein is cloned using an oligonucleotide probe synthesized on the basis of the amino acid sequence. If the direct product of the *sd-1* gene cannot be identified, while a protein produced by any other gene linked with the *sd-1* gene can be identified, the *sd-1* gene might be cloned by "gene walking" from that gene encoding the protein.

Semidwarfism near-isogenic line

A semidwarf near-isogenic line, SC–TN1, was produced by seven backcrosses, using a tall Japanese

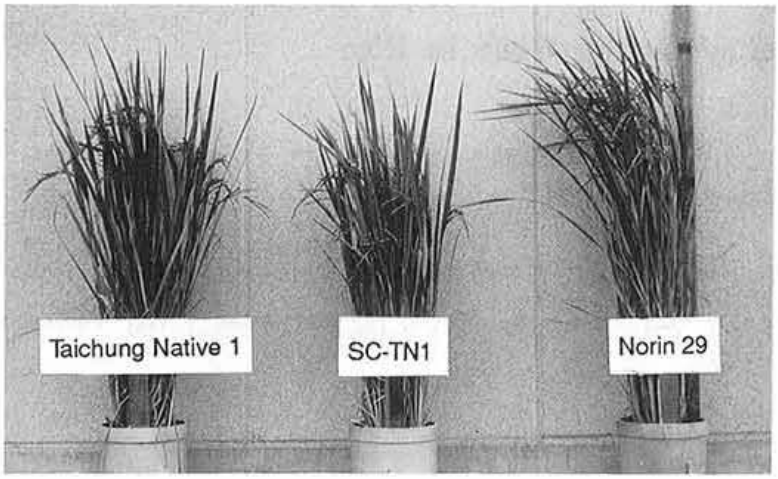

Plate 1. Taichung Native 1, Norin 29 and its semidwarf near-isogenic line SC-TN1

cultivar Norin 29 as the recurrent parent and a semidwarf cultivar Taichung Native 1 as a donor parent[8,9]. Taichung Native 1 is a progeny of Tsai-yuan-chung × Dee-geo-woo-gen and has the short-stature gene, i.e. *sd-1* derived from Dee-geo-woo-gen. A crossing test confirmed that the *sd-1* gene from Taichung Native 1 had been transferred to the near-isogenic line. The phenotypes of the line such as heading date, panicle length, grain length and seed fertility were almost identical with those of Norin 29, except that the near-isogenic line was short-statured[8] (Plate 1).

It was presumed through a comparison of the electrophoretic patterns between Norin 29 and its semidwarf near-isogenic line that the changes in protein composition related to semidwarfism, while no other relevant traits could be easily identified.

Identification of proteins apparently related to semidwarfism

The two-dimensional gel electrophoresis (2D-PAGE) patterns of the rice embryos were compared between Norin 29 and its semidwarf near-isogenic line, SC-TN1[9]. Plate 2 shows the 2D-PAGE patterns of the embryo proteins which were detected by silver staining in Norin 29 and SC-TN1. A large number of polypeptide spots (>500) were identified in each pattern. The electrophoretic pattern of all of the polypeptide spots except two was identical in Norin 29 and SC-TN1. Only one major difference was observed in two polypeptide spots. Since these polypeptides seemed to be related with semidwarfism, they were tentatively designated as semidwarfism-related proteins (SRP)-1 and SRP-2. The presence of SRP-1 was intensively detected in Norin 29, but only weakly in SC-TN1, while the presence of SRP-2 was clearly identified in SC-TN1, but weakly in Norin 29 (Plate 2). SRP-1 and SRP-2 have the same molecular mass 34kD, but different isoelectric points, 7.4 and 7.7, respectively.

Genetic analysis of SRP-1 and SRP-2

The cross of Norin 29 (SRP-1) and SC-TN1 (SRP-2) yielded an F_1 progeny having both SRP-1 and SRP-2 (SRP-1/SRP-2) (Table 1). The F_2 progeny segregated into three types, having SRP-1, SRP-1/SRP-2 and SRP-2, giving an acceptable fit to a 1:2:1 ratio (Table 1). These results suggest that a single locus control SRP-1 and SRP-2 with co-dominant alleles. The gene symbol *Srp* was tentatively given to this locus with alleles Srp^1 and Srp^2, respectively[9]. SRP-1/SRP-2 was controlled by the heterozygote of Srp^1 and Srp^2 [9].

In the early stage of the study, the *Srp* locus was believed to correspond to the *sd-1* locus. However, later the *Srp* locus was found to be linked with the

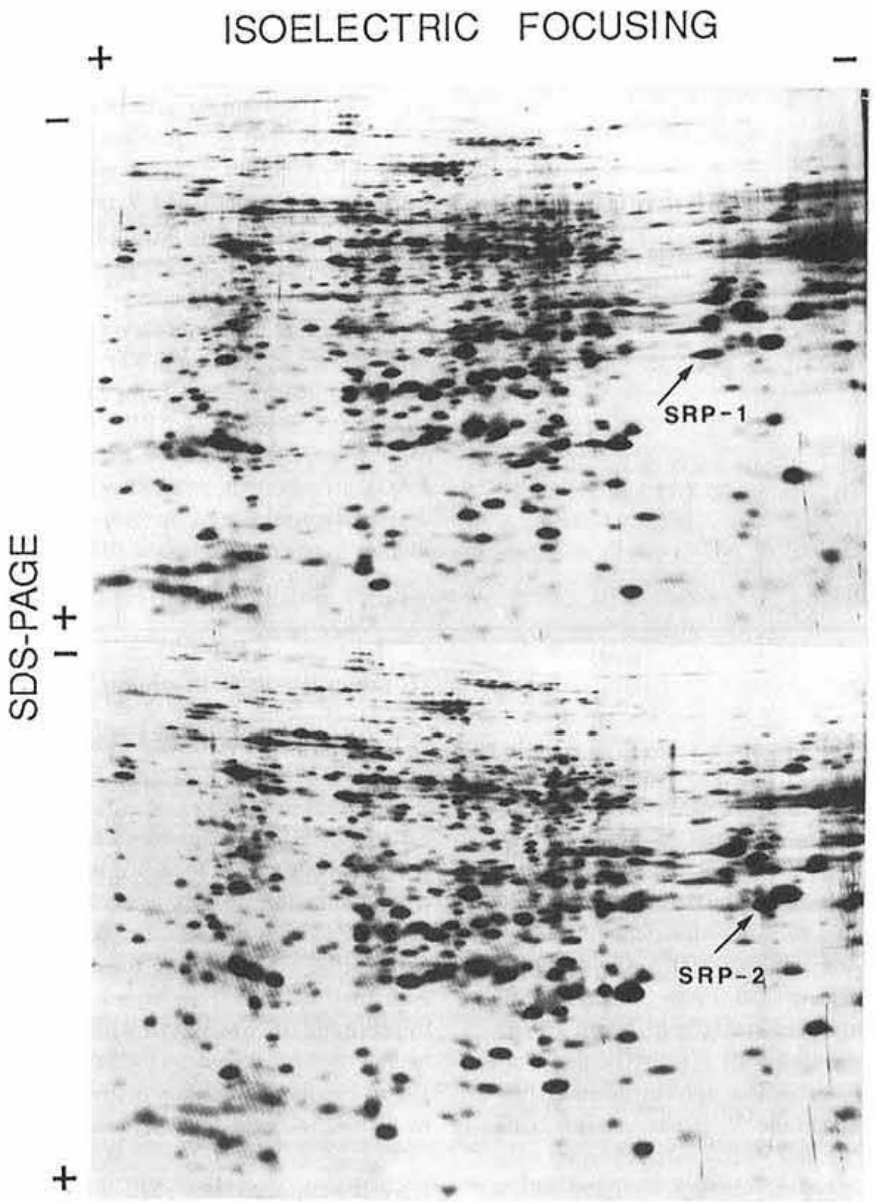

Plate 2. 2D-PAGE pattern of rice embryo proteins of Norin 29 and SC-TN1

sd-1 locus through a linkage analysis[10]. In the F_2 population, a small number of semidwarf plants with SRP-1 and tall plants with SRP-2 were observed. These plants were recognized as recombinants. The F_2 population was separated into 90 tall plants with SRP-1 or SRP-1/SRP-2, 11 tall plants with SRP-2, 9 semidwarf plants with SRP-1 or SRP-1/SRP-2 and 26 semidwarf plants with SRP-2 (Table 2). Thus, it was obvious that the *Srp* locus is linked with the *sd-1* locus in chromosome 1. The recombination value of the *sd-1* and *Srp* loci calculated with the maximum likelihood method was $15.5 \pm 3.4\%$.

Amino acid sequence analysis

From the results obtained with a linkage analysis,

Table 1. Segregation of SRP-1 and SRP-2 proteins in rice embryo

Population	Number of plants			Total	χ^2 (1:2:1)	P value
	SRP-1	SRP-1/ SRP-2	SRP-2			
Norin 29/ Norin 29	20	0	0	20		
SC-TN1/ SC-TN1	0	0	20	20		
SC-TN1/ Norin 29 F_1	0	20	0	20		
SC-TN1/ Norin 29 F_2	15	27	13	55	0.164	0.90–0.95

Table 2. Segregation of culm length in F_2 population of a cross between SC-TN1 and Norin 29

Population	Number of plants			χ^2 (3:1)	P value
	Tall	Semi-dwarf	Total		
SC-TN1/ Norin 29 F_2	355	121	476	0.045	0.75–0.90

it was expected that the *sd-1* gene would be cloned by "gene walking" from the *Srp* gene coding for SRP. To clone a DNA encoding SRP, the amino acid sequences of SRP-1 and SRP-2 were analyzed. If the sequence of SRP could be partially determined, it would be possible to synthesize oligonucleotides corresponding to the sequence, which would be used for the cloning of the gene encoding SRP as probes.

Proteins from the seed embryos of Norin 29 and SC-TN1 were separated with 2D-PAGE and electroblotted from the gel onto a polyvinylidene difluoride (PVDF) membrane[4,5]. The amino acid sequences of SRP-1 and SRP-2 electroblotted onto the PVDF membrane were directly analyzed with a gas-phase sequencer. However, no amino acid sequence data were obtained, suggesting that the N-termini of SRP-1 and SRP-2 be blocked.

A novel technique of on-membrane deblocking for the protein having a blocking group at the N-terminus[6,7] was applied to the sequence determination of SRP-1 and SRP-2. In the first step of this technique, the protein is separated by PAGE, electroblotted onto the PVDF membrane and the N-terminal amino acid having a blocking group is released with pyroglutamyl peptidase or acrylamino acid-releasing enzyme to determine the amino acid sequence. SRP-1 and SRP-2 could be sequenced after on-membrane pyroglutamyl peptidase digestion. The N-terminal amino acid of SRP-1 and SRP-2 was presumed to be pyroglutamic acid which was formed by cyclization of glutamine or glutamic acid. The sequences of 16 and 9 residues from the N-termini of SRP-1 and SRP-2, respectively, were determined by the gas-phase protein sequencer[5] (Fig. 1).

On the other hand, the embryo proteins were separated by 2D-PAGE, and SRP-1 and SRP-2 were separately electroeluted from the gels. After electroelution, SRP-1 and SRP-2 were digested with *Staphylococcus aureus* V8 protease on the SDS-PAGE gel according to the Cleveland method[2]. The digests were separated by SDS-PAGE and electroblotted from the gel onto the PVDF membrane[5] and the internal amino acid sequences were determined (Fig. 1).

Homology search in amino acid sequences

The structural homology search using the SEQDB data base (Protein Research Foundation, Osaka) showed that the amino acid sequences of SRP-1 and SRP-2 were highly homologous with that deduced from the nucleotide sequence of type I cDNA encoding glutelin[12], which is a major storage protein in the seed endosperms. In the regions of SRP-1 and SRP-2 examined, the sequences were identical with that deduced from type I cDNA, except that the 60th residue was valine, while the corresponding residue in SRP-1 and SRP-2 was alanine, and the 122nd residue was glutamine, while it was histidine in SRP-2[6].

Identification of the endosperm glutelin subunits

In the seed endosperms of Norin 29 and SC-TN1, a total series of eight glutelin subunits was detected by 2D-PAGE (Plate 3). The structural differences of the eight glutelin subunits were analyzed to determine which glutelin subunits were identical or most homologous with SRP-1 and SRP-2[6].

All these glutelin subunits were found to have a blocked N-terminus. Therefore, after 2D-PAGE, the glutelin subunits which were electroblotted onto the

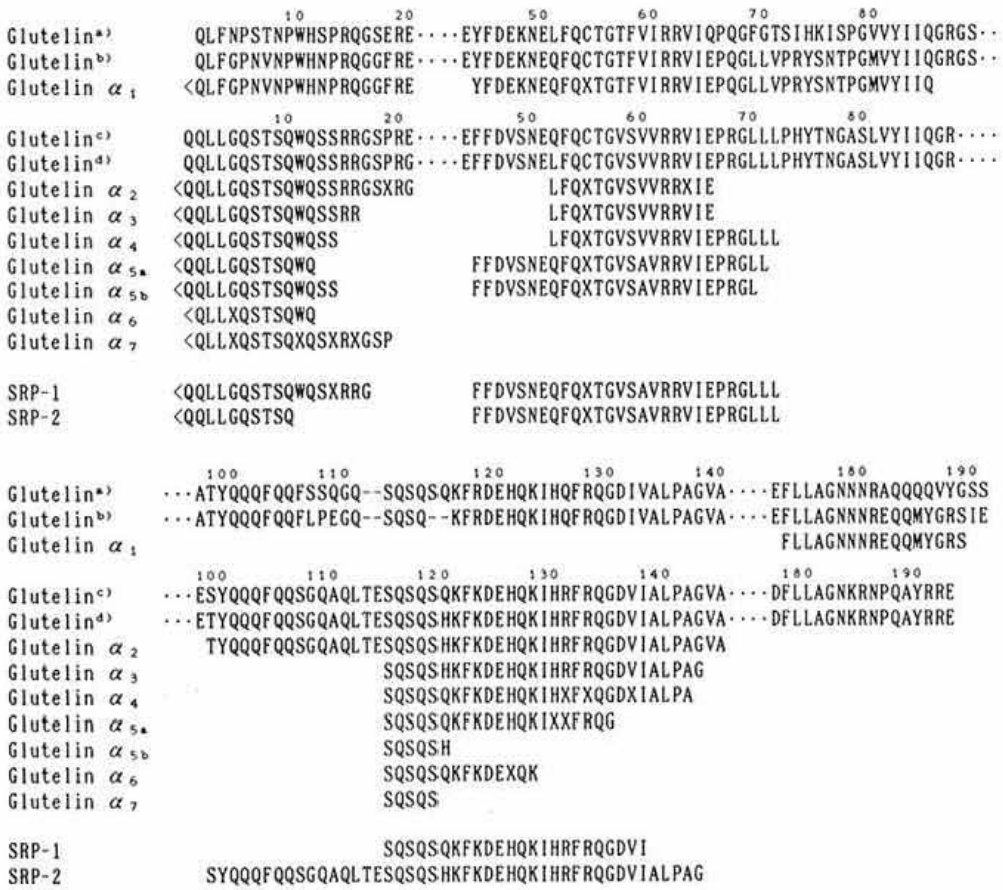

Fig. 1. Amino acid sequences of the seed glutelin α-subunits
The sequences deduced from nucleotide sequences. The different residues between a) and b), and between c) and d) are indicated in dots.
< Q represents pyroglutamic acid.
a), b), c), d) refer to 13), 11), 12) and 12), respectively.

PVDF membrane were digested with pyroglutamyl peptidase to remove the blocked N-terminal amino acid. After removal of the blocked amino acid, the amino acid sequences in the N-terminal regions of these protein subunits could be determined (Fig. 3). The glutelin fragments obtained after *in situ* protease digestion on SDS-PAGE gel were electroblotted onto the PVDF membrane and the internal amino acid sequences of the subunits were determined.

The results of these experiments revealed that the amino acid sequences of SRP-1 and SRP-2 were most homologous with those of the glutelin α_{5a}- and α_{5b}-subunits.

As described above, SRP-1 and SRP-2 have isoelectric points, 7.4 and 7.7, respectively, but with an identical molecular mass, 32kD. The glutelin α_{5a}- and α_{5b}-subunits showed the same isoelectric points and molecular mass as SRP-1 and SRP-2, respectively. From these results, it is concluded that SRP-1 and SRP-2 are identical with glutelin α_{5a}- and α_{5b}-subunits, respectively.

Complementary DNA encoding the glutelin subunits

The amino acid sequences of the glutelin α-subunits were compared with those deduced from the nucleotide sequences of DNAs encoding the

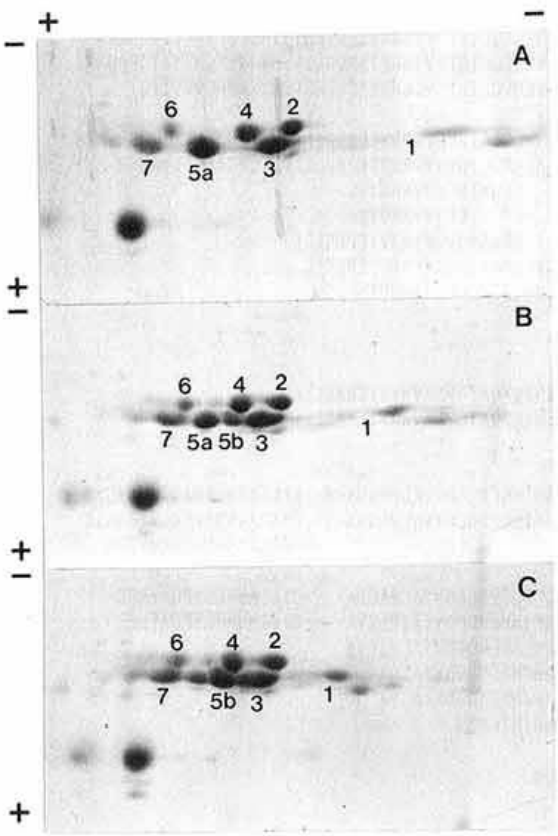

Plate 3. 2D-PAGE patterns of the seed endosperm glutelin α-subunits
A: Norin 29, B: A+C mixture,
C: SC-TN1

glutelin subunits. This comparison showed that the sequence of glutelin α_1-subunit was highly homologous with those deduced from the nucleotide sequence of cDNA[11]. The amino acid sequences of the glutelin α_2-, α_3- and α_4-subunits were found to be identical with that deduced from the nucleotide sequences of type II cDNA[12]. The sequences of glutelin α_{5a}-subunit (SRP-1) and α_{5b}-subunit (SRP-2) were homologous with that deduced from the sequence of type I cDNA[12]. No genes encoding glutelin α_6- and α_7-subunits have been reported so far. In these subunits, the N-terminal glutamine that is present in the other subunits was deleted. The α_6- and α_7-subunits seem to be encoded by the gene that is structurally similar to the type I and II cDNAs, but not identical.

Glutelin species in the embryos

As described earlier, it is confirmed that the proteins identical with the endosperm glutelin α-subunits exist in the embryos. The question is: Are the glutelin α_{5a}- and α_{5b}-subunits only or all of the glutelin α-subunits expressed in the embryos?

After the embryo proteins were separated by 2D-PAGE, an analysis was made to identify the amino acid sequences of the proteins which showed a similar molecular mass and isoelectric point to those of the endosperm glutelin α-subunits. However, no glutelin α-subunits except for the α_{5a}- and α_{5b}-subunits could be identified in the embryos of Norin 29 and SC-TN1. This result suggests that the mechanism regulating the expression of the genes encoding the glutelin subunits be probably different between the embryos and endosperms. The glutelin is presumed to act as a source of nitrogen and amino acids for germination in the endosperms. The glutelin α_{5a}- and α_{5b}-subunits in the embryos are likely to have a special function which would be different from the storage proteins.

Conclusion

The *Srp* locus is linked with the *sd-1* locus with a recombination value of 15.5%. The size of the rice genome is estimated to be approximately 10^6 kbp. Accordingly, the value of 15.5% roughly corresponds to 500–1,000 kbp. Since the cDNAs and genes encoding the glutelin have been cloned, it is expected that the *sd-1* gene could be isolated in future by applying the "gene walking" technique.

This study was performed in collaboration with Mr. H. Kajiwara, Dr. F. Sakiyama and Dr. S. Tsunasawa.

References

1) Chang, T. T. et al. (1965): Genetic analysis of plant height, maturing and other quantitative traits in the cross Peta × I-geo-tze. *J. Agr. Assoc. China News Serv.*, **39**, 1–8.
2) Cleveland, D. W. et al. (1977): Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.*, **252**, 1102–1106.
3) Hargrove, T. H., Coffman, W. R. & Cabanilla, V. L.

(1980): Ancestry of improved cultivars as parents in Asian rice breeding programs. *Crop Sci.*, **20**, 721–727.
4) Hirano, H. (1989): Microsequence analysis of winged bean seed proteins electroblotted from two-dimensional gel. *J. Protein Chem.*, **8**, 115–130.
5) Hirano, H. & Watanabe, T. (1990): Microsequencing of proteins electrotransferred onto immobilizing matrices from polyacrylamide gel electrophoresis: application to an insoluble protein. *Electrophoresis*, **11**, 573–580.
6) Hirano, H. et al. (1991): Structural homology between semidwarfism-related proteins and glutelin seed protein in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* (In press).
7) Hirano, H. et al. (1991): Microsequence analysis of N-blocked proteins electroblotted onto an immobilizing matrix from polyacrylamide gels. *Anal. Sci.*, **7** (suppl.), 941–944.
8) Kikuchi, F. et al. (1985): Genetic analysis of semidwarfism in high-yielding rice varieties in Japan. *Bull. Nat. Inst. Agr. Sci.*, Ser. D **36**, 125–145.
9) Nakamura, A., Hirano, H. & Kikuchi, F. (1991): Identification and genetic analysis of semidwarfism-related proteins in rice (*Oryza sativa* L.). *Theor. Appl. Genet.*, **81**, 376–380.
10) Nakamura, A., Hirano, H. & Kikuchi, F. (1991): A protein gene *Srp*(t) linked with a semidwarfing gene *sd-1* in rice (*Oryza sativa* L.). *Jpn. J. Breed.*, **41**, 517–521.
11) Masumura, T. et al. (1989): Nucleotide sequence of a cDNA encoding a major rice glutelin. *Plant Mol. Biol.*, **12**, 723–725.
12) Takaiwa, F., Kikuchi, S. & Oono, K.(1987): A rice glutelin family — A major type of glutelin mRNAs can be divided into two classes. *Mol. Gen. Genet.*, **208**, 15–22.
13) Takaiwa, F., Kikuchi, S. & Oono, K. (1989): The complete nucleotide sequence of new type cDNA coding for rice storage protein glutelin. *Nucl. Acids Res.*, **17**, 3289.

(Received for publication, Oct. 16, 1991)