Genetic analysis of cysteine-poor prolamin polypeptides reduced in the endosperm of the rice esp1 mutant

Ushijima, Tomokazu
Institute of Genetic Resources, Faculty of Agriculture, Kyushu University

Matsusaka, Hiroaki
Institute of Genetic Resources, Faculty of Agriculture, Kyushu University

Jikuya, Hiroyuki
Bio-Architecture Center, Kyushu University

Ogawa, Masahiro
Department of Life Science, Yamaguchi Prefectural University

他

出版情報：Plant Science. 181 (2), pp.125-131, 2011-08. Elsevier
バージョン：

Title:

Genetic analysis of cysteine-poor prolamin polypeptides reduced in the endosperm of the rice *esp1* mutant

Running title:

Genetic analysis of cysteine-poor prolamins in the rice *esp1* mutant

Authors:

Tomokazu Ushijima†[1], Hiroaki Matsusaka†[1], Hiroyuki Jikuya[2*], Masahiro Ogawa[3], Hikaru Satoh[1], and Toshihiro Kumamaru[1]

Address:

[1]Institute of Genetic Resources, Faculty of Agriculture, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan

[2]Bio-Architecture Center, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan

[3]Department of Life Science, Yamaguchi Prefectural University, Sakurabatake, Yamaguchi 753-8502, Japan

*Present address: Shimadzu Corp., 1, Nishinokyo-Kuwabara-cho, Nakagyo-ku, Kyoto 604-8511, Japan

† These authors contributed equally to this work.

Corresponding Author: Toshihiro Kumamaru

E-mail: kumamaru@agr.kyushu-u.ac.jp

Tel & FAX: 81-92-642-3057

Abstract

The *esp1* mutant CM21 specifically exhibits reduced levels of cysteine-poor (CysP) prolamin bands with pIs of 6.65, 6.95, 7.10, and 7.35 in rice seed. Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis demonstrated that the bands with pIs 6.65, 6.95, and 7.35 are encoded by different structural genes. These results suggest that the *Esp1* locus encodes a regulatory factor involved in the synthesis and/or accumulation of CysP prolamin molecules. Isoelectric focusing (IEF) analysis of CysP prolamins in chromosome substitution lines showed that structural genes for bands with pI values of 6.95, 7.10, and 7.35, which are reduced in *esp1* mutant lines, are located as a gene cluster in the 44.2 cM region on chromosome 5.

Keywords: MALDI-TOF MS, mutant, prolamin, rice, starchy endosperm, storage protein

1. **Introduction**

Prolamin is an alcohol-soluble protein that predominantly accumulates in protein bodies in cereal endosperm. Classical and molecular genetic studies have shown that most of the prolamins in *Triticeae* are encoded by multigene families located at complex loci on the homoeologous group 1 chromosomes [1]. The locations of loci that control homologous groups of proteins are highly conserved in three *Triticeae* species [1]. Gene clusters encoding various forms of zein (a class of maize prolamin) have been identified in some chromosomal regions [2-4]. However, loci for the gene family that encodes prolamins in rice have not yet been identified.

Cereal prolamins are classified into two types: cysteine-poor (CysP) and cysteine-rich (CysR). CysP prolamins contain few cysteine residues, whereas CysR prolamins contain relatively more cysteine residues [5]. CysP prolamins are easily extracted by alcohol solution in the absence of reducing agents, whereas CysR prolamins are insoluble in alcohol solution in the absence of reducing agents. All classes of prolamin in rice are stored in protein body type I (PBI) as intracisternal protein granules in endoplasmic reticulum (ER) lumen [6, 7]. PBI is a spherical structure with a diameter of 1–2 μm and exhibits polysomes attached to its membranes [7, 8].

Prolamins in rice are composed of 10-kDa, 13-kDa, and 16-kDa polypeptide species [6, 9]. The prolamin fraction of the *japonica* rice variety, Kinmaze, consists of 10-kDa, 13-kDa (indicated as 13b in Ogawa et al. [6]), 14-kDa (indicated as 13a in Ogawa et al. [6]), and 16-kDa prolamins. The 10-kDa, 14-kDa, and 16-kDa prolamins are CysR species, while the 13-kDa prolamin is a CysP class prolamin and is the major prolamin found in rice [6]. Isoelectric focusing (IEF) of rice prolamins revealed high

micro-heterogeneity in 13-kDa proteins belonging to the CysP class, while the 10-kDa and 16-kDa prolamins belonging to the CysR class exhibit low micro-heterogeneity[10]. With respect to the maize prolamin zein, charge heterogeneities detected by IEF, revealed differences in amino acid composition among various zein components [11]. *In vitro* translation products of polysomes bound to PBs exhibit the same heterogeneity as native zein [12], and individual zein bands resolved by IEF correspond to the products of a system of structural genes [4]. These findings suggest that identification of the structural genes corresponding to CysP prolamin polypeptides in rice is viable.

Several gene loci control the level of prolamin accumulation during endosperm development in rice [13, 14]. Among these loci, the *Esp1* locus might not be a structural prolamin gene, but may rather exert a regulatory effect upon prolamin polypeptide accumulation [13]. The presence of a recessive allele at this locus strongly reduces CysP prolamin accumulation, without quantitatively modifying the CysR prolamin components. Although mutations in *Esp1* result in reduced levels of 13-kDa prolamin polypeptides, the morphological characteristics of prolamin-containing protein bodies are almost the same as those in wild-type [14, 15].

We hypothesized that the *Esp1* locus encodes a regulatory factor for the expression of prolamin genes to the subsequent accumulation of prolamin molecules in PBI. To better understand the genetic mechanisms responsible for prolamin polypeptide regulation, we examined our hypothesis through the genetic analysis of CysP prolamin polypeptides reduced in the *esp1* mutant. The IEF bands of CysP prolamins exhibiting reduced level in the *esp1* mutant were genetically analyzed to ascertain whether the *Esp1* gene encodes a regulatory factor that controls the accumulation of CysP prolamin polypeptides and to identify the gene family that corresponds to CysP prolamins.

2. Materials and Methods

2.1. Plant materials

Experiments were performed using the *esp1* mutant line, CM21 and EM711 derived from Kinmaze and Taichung65, respectively, chromosome segment substitute lines (CSSLs) and rice varieties: Kinmaze, Taichung65, Nipponbare, and Kasalath. The *esp1* mutation was induced by treatment with *N*-methyl-*N*-nitrosourea, resulting in seeds exhibiting low amounts of storage protein and reduced prolamin content [13-16] and stocked in Kyushu University in National Bioresource Project. CSSLs were developed from backcrosses of Nipponbare and Kasalath distributed by the Rice Genome Resource Center, the National Institute of Agrobiological Sciences, Japan.

2.2. Storage protein extraction

Total storage proteins were extracted from mature seeds using an extraction buffer consisting of 4% (w/v) SDS, 4 M urea, and 5% (v/v) 2-mercaptoethanol (2-ME) in 0.125 M Tris-HCl, with pH 6.8 (SDS-PAGE sample buffer). Total prolamin fraction was extracted from mature seeds using 1ml of 60% (v/v) 1-propanol containing 5% (v/v) 2-ME per single seed. CysP and CysR prolamin fractions were extracted from mature seeds using 1 ml of 60% (v/v) 1-propanol and 1 ml of 60% (v/v) 1-propanol containing 5% (v/v) 2-ME from residue extracted by 60% 1-propanol. Extracted prolamin fractions from single seed were dried by centrifugation evaporator and were dissolved using 50 µl of SDS-PAGE sample buffer or 10 µl of IEF sample solution (0.1% (w/v) SDS, 8.5 M urea, 5.0% (v/v) 2-ME, and 0.8% (w/v) Nonidet P-40).

2.3. SDS-PAGE

SDS-PAGE gels were prepared, and electrophoresis was performed according to the methods of Laemmli [17]. After electrophoresis, the gels were stained with a staining solution (0.1% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) methanol, and 7% (v/v) acetic acid), and subsequently destained with a solution of 25% (v/v) methanol and 10% (v/v) acetic acid.

2.4. IEF electrophoresis

IEF was performed using a Multiphor II Electrophoresis unit (GE Healthcare Biosciences). Horizontal slabs of IEF gels were prepared and subjected to electrophoresis according to the methods described by Brinegar and Peterson [18]. Gels contained 4% (w/v) acrylamide, 0.2% (w/v) methylene-bis-acrylamide, 6 M urea, 2% (w/v) Nonidet P-40, 2% (w/v) Ampholine (1% pH 3.5–10, 1% pH 6.0–8.0; GE Healthcare Biosciences), 0.01% (w/v) ammonium persulfate, 0.0005% (w/v) Riboflavin, and 0.0525% (v/v) TEMED. 8 μl of IEF sample solution in each prolamin fractions was applied to a filter paper (5 × 5 mm) at the anodic side. IEF was run for 190 min. (30 min. at 150 V, 60 min. at 400 V, 60 min. at 700 V, and 40 min. at 1000 V) at 0 ºC. After electrophoresis, the gels were fixed with 15% (v/v) trichloroacetic acid. Ampholine was removed from the gels using IEF washing solution (25% (v/v) ethanol and 10% (v/v) acetic acid). The gel was stained with IEF staining solution (0.1% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) ethanol, and 10% (v/v) acetic acid) and then destained with IEF destaining solution (25% (v/v) ethanol and 10% (v/v) acetic acid).

2.5. Two-dimensional (2-D) electrophoresis analysis

The bands of IEF (the first-dimensional electrophoresis), were excised from the gel and incubated on boiling water for 3 min. in 125 mM Tris-HCl, pH 6.8, containing 4% (w/v) SDS, and were then loaded onto SDS-PAGE for the second dimensional electrophoresis.

2.6. Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis

The 2-D gels were thoroughly washed with water to remove acetic acid, methanol, SDS, and salts. Excised gel spots along with 100 μl of 25 mM ammonium bicarbonate (pH 8.0) were placed in the tube and incubated for a minimum of 30 min. at 37 ºC. After the gel spots were washed by 100 μl of 50% acetonitrile, they were incubated in 50 mM ammonium bicarbonate for 30 min. at 37 ºC. The gel spots were dried in a speedvac, swelled using 10 μl of 50 mM ammonium bicarbonate (pH 8.0, containing 20 ng/μl trypsin), and incubated in a fridge for 30-60 min. After the trypsin solution was removed, 10 μl of 50mM ammonium bicarbonate (pH 8.0), was added to the gel spots. Spots were then incubated overnight at 30 ºC. After the gel spots were sonicated for 10 min., the supernatants (Sup1) were transferred to a fresh tube. To extract the peptides remaining in the gels, 20 μl of 50% acetonitrile and 0.1% TFA solution was added to the gel spots; gel spots were sonicated for 10 min., and the supernatants (Sup2) were mixed with Sup1. The mixture of supernatants was dried using a speedvac. To remove the ammonium bicarbonate from the dried supernatants, 10 μl of Milli-Q water was added. They were sonicated for 5 min., and then dried using a speedvac. The dried peptides were resuspended using 1.5 μl of the matrix solution and 10 mg matrix (α-cyano-4-hydroxycinnamic acid) in 1 ml of 50% acetonitrile and 0.1% TFA solution. After the peptide solutions were vortexed for 10

sec., 1.5 μl of the solution was placed on the MALDI sample plate. MALDI-TOF MS analysis was performed using the AXIMA-CFR plus mass spectrometer (Shimadzu Biotech, Manchester, U.K.).

2.7. RT-PCR

Total RNA was extracted by RNeasy plant mini kit (BioRad) from the developing seed at two weeks after flowering. First-strand cDNA was synthesized from 400 ng from total RNA with the PrimeScript RT reagent kit with gDNA Eraser (TaKaRa) and the cDNA was used as template, according to [19]. The specific primers for prolamin genes are listed in Supplemental Table 1. Polymerase chain reaction (PCR) was performed by 1 cycle at 94 °C for 2 min. and 25 cycles at 94 °C for 30 sec., 55 °C for 30 sec., and 72 °C for 40 sec. PCR products were electrophoresed on agarose gels and visualized by staining with ethidium bromide.

3. Results

3.1. Specification of the CysP prolamin polypeptides reducing in *esp1* mutation

Figure 1 shows the SDS-PAGE profile of CysP and CysR prolamin fractions in the *esp1* mutant line CM21. The 13-kDa polypeptide in the CysP prolamin fraction from *esp1* endosperm was considerably reduced compared to that of wild-type (WT) (Fig. 1C). However, the profile of polypeptides detected in the CysR prolamin fraction, was virtually the same for both CM21 and WT (Fig. 1D). These results indicate a reduction in the level of the 13-kDa CysP prolamin component, which is a characteristic

specific to CM21.

CysP prolamins from the endosperm of WT were separated by IEF into at least eight distinct bands with pIs of 6.60, 6.65, 6.85, 6.95, 7.10, 7.35, 7.40, and 8.00 (Fig. 2). CysP prolamin bands with pIs of 6.60, 6.85, 7.40, and 8.00 were virtually the same for both CM21 and WT, whereas bands with pIs of 6.65, 6.95, 7.10, and 7.35 in CM21 were significantly reduced relative to WT.

The composition of CysP prolamins in the F_2 seeds resulting from the cross between CM21 and Kinmaze was analyzed to determine whether the reduction of the four IEF bands in CM21, relative to WT, was due to the *esp1* mutation (Fig. 3). All F_2 seeds bisected and then divided into two portions. One portion was used to determine genotype of the F_2 seeds by SDS-PAGE analysis of the total protein extracted. CysP prolamins extracted from the other portion were used for IEF analysis. All *esp1* mutant F_2 seeds exhibiting reduced level of the 13-kDa polypeptide, also showed reduced levels of CysP prolamin bands with pIs of 6.65, 6.95, 7.10, and 7.35. In contrast, all wild-type F_2 seeds exhibiting normal levels of the 13-kDa polypeptide showed normal levels of all IEF bands. Wild-type and *esp1* mutants from the F_2 seeds of the cross were segregated in a ratio of 87:25, respectively. This indicates that the segregation mode is a close approximation to the expected ratio of 3:1. These results demonstrate that reduction of CysP prolamin bands with pIs of 6.65, 6.95, 7.10, and 7.35 in CM21 was due to the *esp1* mutation.

3.2. Identification of the genes for the CysP prolamins reducing by *esp1* mutation

To elucidate whether each prolamin band exhibiting reduced levels in *esp1*

mutants was encoded by independent structural genes, the mass spectrometry analysis of each reduced band in the *esp1* mutation was performed. CysP prolamin bands with pIs of 6.65, 6.95, 7.10, and 7.35 in wild-type were optimally separated by IEF gel electrophoresis [20] in the first dimension, and SDS-PAGE in the second dimension (Fig. 4). 2-D analysis of CysP prolamin demonstrated that all IEF CysP prolamin bands reduced in *esp1* mutants indicated a single 13-kDa band. The protein spots in the 2-D gel were excised and digested in-gel with trypsin, and were then analyzed by peptide mass fingerprinting (PMF) using MALDI-TOF MS. Homology search analyses for prolamin genes revealed the existence of 20 prolamin genes in the rice genome (Supplemental Table 2). The theoretical mass values of trypsin digested peptide fragment deduced from all prolamin sequences searched from database were shown in Supplemental Table 3.

The peptides with apparent mass values of 1195.38, 1713.35, 2096.80, 2779.31, and 4978.26 were detected by the MALDI-TOF MS spectrum of the bands with pI 6.65 (Table 1) and are corresponded to peptides with theoretical mass value of 1212.33, 1712.97, 2096.37, and 2778.10, and 4974.53 or 4988.55, respectively among the peptide fragments deduced from the Os05g0329100 or Os05g0330600 sequences (Table 1, Supplemental Table 3). As the apparent mass value of 4978.26 is closer to theoretical mass value of 4974.53 than 4988.55, it seems likely that pI 6.65 band may be Os05g0329100. The peptides with mass values of 1195.38, 1713.43, 2096.80, 2787.80, and 4975.21 were detected by the MALDI-TOF MS spectrum of the bands with pI 6.95, and are corresponded to peptides with theoretical mass value of 1212.33, 1712.97,

2096.37, 2787.17, and 4974.53, respectively, among the peptide fragments deduced from Os05g0328800 sequence (Table 1, Supplemental Table 3). It is likely that the peptides with apparent mass values of 1195.38 in the bands with pIs 6.65 and 6.95 (Table 1) resulted from the modification of the N-terminal peptide in the fragment with mass values of 1212.33 deduced from the Os05g0329100, Os05g0330600, or Os05g0328800 sequences (Supplemental Table 3). In the MALDI-TOF MS spectrum of the pI 7.10 band, peptides with apparent mass values of 1212.31, 1713.35, 2096.53, 2124.59, 2778.37, and 4977.06 were detected (Table1) and corresponded to peptides with theoretical mass value of 1212.33, 1712.97, 2096.37, 2124.43, 2778.16, and 4974.53 or 4988.55 among the peptide fragments deduced from the Os05g0329100 and/or Os05g0330600, and Os05g0329700 sequences (Table 1, Supplemental Table 3). Especially peptide with the theoretical mass value of 2124.43 is unique in Os05g0329700, but the peptide with the theoretical mass value of 2096.09 is not in Os05g0329700 (Supplemental Table 3). This suggests that the 2-D spot of the pI 7.10 band is a constituent of other prolamin molecules with Os05g0329700. In the MALDI-TOF MS spectrum of the pI 7.35 band, peptides with apparent mass values of 1212.35, 1713.35, 2096.44, and 4981.67 were detected (Table 1) and corresponded to peptides with theoretical mass value of 1212.33, 1712.97, and 2096. 37, and 4974.55 or 4988.55 among the peptide fragments deduced from the Os05g0330100 or the Os05g0330600 (Table 1, Supplemental Table 3). Though the apparent mass value of 4981.67 is the intermediate between the peptides with theoretical mass value of 4988.55 and 4974.55, it seems likely that the band with pI 7.35 may be Os05g0329600.

The theoretical pI and molecular mass of the protein deduced from the prolamin sequences on chromosome 5 were calculated. Though the theoretical molecular mass of all gene products are about 15 kDa, the theoretical pIs of all gene products were 8.22 or 8.23, which did not coincide to apparent values (Supplemental Table 4). It is reported that the discrepancy of apparent and theoretical pI values reflects the experimental methods including the effect of binding of ions and carrier ampholytes to the proteins, the structure and properties of the proteins such as number of disulfide bridges, ion binding, the degree and type of folding, the number of exposed ionized amino acid [21], and the post-translational modification [22]. Additionally, prolamins are insoluble in water and very hydrophobic, meaning that they contain low amount of charged amino acid residues than do soluble proteins. Such unique properties may influence the mobility in isoelectric focusing.

These results clearly demonstrate that the bands with pIs 6.65, 6.95 and a part of pI 7.10 that exhibit reduced levels in *esp1* mutants are encoded by different structural genes on chromosome 5, though it was unknown whether the gene corresponding to the bands with pIs of 7.10 is single. This suggests that the *Esp1* locus encodes a regulatory factor responsible for regulating the synthesis and/or accumulation of CysP prolamin molecules. Thus, it is proposed that the four CysP prolamin bands with pIs of 6.65, 6.95, 7.10, and 7.35, which exhibit a reduction in level in *esp1* mutants, should be designated as *Esp1*-regulated CysP prolamins.

3.3. Chromosome localization of the genes for CysP prolamin polypeptide reducing in

esp1 mutation

The genes corresponding to the bands with pIs of 6.65, 6.95, 7.10, and 7.35 were determined to be localized on chromosome 5 by MALDI-TOF MS analysis. Chromosome localization analysis of CysP prolamin genes was confirmed by using CSSLs. In Kasalath, three bands with pIs of 6.95, 7.10, and 7.35, which showed reduced levels in *esp1* mutants, were deleted. Furthermore, bands with pIs of 6.95, 7.10, and 7.35 were deleted in three subsequent CSSLs (Fig. 5A). The 44.2 cM region between 27.5 and 71.7 cM from the short arm end of chromosome 5 was derived from Kasalath in three CSSLs (Fig. 5B). These results reveal that the structural genes for the IEF bands with pIs of 6.65, 6.95, 7.10, and 7.35 that exhibited reduced levels in *esp1* mutants are located in the region of 44.2 cM on chromosome 5 as a gene cluster.

3.4. Identification and chromosome localization of the genes for the CysP prolamins showing normal level by *esp1* mutation

The same analysis was performed on the IEF bands of CysP prolamins in the *esp1* mutants that exhibited normal levels relative to wild-type. Among them, the genes for IEF bands with pIs of 6.60 and 6.85 were located in the region of 31.5 cM on chromosome 7 (Supplemental Fig. 1). 2-D analysis of CysP prolamin demonstrated that IEF CysP prolamin bands with pIs of 6.60 and 6.85 showing the normal level in *esp1* mutants indicated a single 13-kDa band (Supplemental Fig. 2). The peptides with apparent mass values of 2130.47, 2778.17, 3866.14 and 2143.81, 2422.14, 2794.75, 3853.31 were detected by the MALDI-TOF MS spectrum from the bands with pIs 6.60 and 6.85, respectively (Supplemental Table 3), and corresponded to peptide theoretical

mass value of 2130.39, 2777.16, 3861.33 and 2140.43 or 2154.45, 2421.78, 2794.20, 3853.31 among the peptide fragments deduced from the Os07g0219300 and Os07g0220050 or Os07g0219400 sequences, respectively (Supplemental Table 3, Supplemental Table 5). The genes corresponding to the bands with pIs of 7.40 and 8.00 could not be identified by MALDI-TOF MS analysis. These results demonstrate that the bands with pIs 6.60 and 6.85 that exhibit normal levels in *esp1* mutant are encoded by different structural genes and the genes are located in the region of 31.5 cM on chromosome 7 as a gene cluster.

The theoretical pI and molecular mass of the protein deduced from the prolamin sequences on chromosome 7 were calculated. Though the theoretical molecular mass of all gene products are about 15 kDa, the theoretical pIs of gene products were 7.02 or 8.23, which did not coincide to apparent values (Supplemental Table 4) same as CysP prolamins derived from the sequence on chromosome 5.

3.5. Expression analysis of the genes corresponding to the CysP prolamins by RT-PCR

In order to predict the function of the *Esp1* gene, the expression level of each prolamin gene in the *esp1* mutant was compared to the wild-type by the RT-PCR analysis. In the genes for Os05g0329100, Os05g0328800, Os05g0330600, which correspond to the CysP prolamin showing the reduced level in *esp1* mutant, the expression level decreased significantly compared with those of the wild-type, while the expression level of in the Os07g0219300 and Os07g0219400, which correspond to the CysP prolamin showing the normal level in *esp1* mutant, was not significant difference from the wild type (Fig. 6). The result such as the specific reduction of the prolamin genes in chromosome 5 suggests the possibilities that the protein encoded by *Esp1*gene

is transcriptional activator for the prolamin genes in chromosome 5 or that prolamin mRNAs are decaied according to mRNA quality control in the cytoplasm by the destruction after the translation on the prolamin biosynthesis process in *esp1* mutant.

4. Discussion

Although the IEF zein bands are known to represent the products of a system of structural genes in maize [4], the identities and locations of genes (indicated by IEF bands) that encode rice prolamin polypeptides were previously not known. Since the *esp1* mutation results in a reduced level of 13-kDa prolamin polypeptides and because the dose of the *Esp1* gene was not accompanied by a dosage effect on the 13-kDa polypeptide level, we hypothesized the *Esp1* gene possibly encodes a regulatory factor which regulates the synthesis and/or accumulation of prolamin [13]. The present study demonstrates that each prolamin IEF band, which shows a reduction in level in *esp1* mutant is encoded by structural genes that form a cluster, while the *Esp1* gene itself encodes a regulatory factor involved in the synthesis and/or the accumulation of CysP prolamins.

cDNAs encoding 13-kDa [23, 24], 10-kDa [25], and 16-kDa prolamins [26] have been isolated and characterized. The 10-kDa rice prolamin has a remarkably high content of methionine (20%) and cysteine (10%), which were estimated using the deduced polypeptide sequence. The genomic clones encoding 10-kDa and 13-kDa prolamins, have also been isolated and sequenced [27, 28]. It has been previously reported that 7–100 copies of the prolamin gene exist within the rice haploid genome [23, 25, 26, 28, 29]. Homology search analyses for prolamin genes revealed the existence of 20 prolamin genes in the rice genome (Supplemental Table 2), while the

expression of 21 genes among total 34 prolamin gene copies in rice genome are reported [30], though the detail reason of the difference is unknown. CysP prolamin genes were located on chromosomes 5 and 7; CysR prolamin genes were located on chromosomes 3, 6, 11, and 12. The gene clusters for *Esp1*-regulated CysP prolamins were located on chromosome 5 (Fig. 5).

With our finding that the *Esp1* gene regulates at least four CysP prolamin polypeptides in mind (Fig. 3), and considering that, in maize, the *O2* locus regulates the sulfur-poor 22-kDa zeins of the α-class [31], we expect that loss-of-function mutations at these loci may have similar effects upon the accumulation of CysP prolamins. *O2* gene encodes a protein with structural homologies to transcriptional activator of zein gene expression [32, 33]. Genes encoding a class of zein proteins form a gene cluster [2-4]. The *O2* locus is located on the short arm of chromosome 7, while the gene cluster encoding the 22-kDa class of zeins, which exhibit reduced levels in *o2* mutants, is mainly located on chromosome 4 [34]. Prolamin genes in rice, also form a gene cluster (Fig. 5, Supplemental Fig. 1), suggesting that duplication events of short chromosomal segments have occurred during evolution, and subsequently gave rise to clustered gene families.

The factor encoded by the *Esp1* gene has not yet been identified. Indeed, the presence of a transcription factor has been reported in a range of cereal prolamins. In rice, the transcription factors RISBZ1 [35], RITA1 [36], and REB [37] are able to bind to the GCN4 motif in a sequence-specific manner. One possible reason is that the *Esp1* gene product functions as an activator of transcription, as the genes for CysP prolamins which showing the reduced level in *esp1* mutant showed the decreased expression (Fig. 6). For the promoter motif of prolamin genes, all of CysP prolamin genes didn't possess the endosperm motif, while almost CysP prolamin genes possess it

[30]. But for GCN4 motif and TATA box within the promoter region in the CysP prolamin genes on the chromosomes 5 and 7 [30], the notable difference between the genes on both chromosomes was not detected. There are also other possibilities regarding the function of the product encoded by the *Esp1* gene. In maize, the *Floury1 (Fl1)* gene encodes a zein protein body membrane protein consisting of three predicted transmembrane domains and a C-terminal plant-specific domain of unknown function [38]. The maize *floury2* (*fl2*) mutation produces pleiotropic effects which include a soft, starchy endosperm, a reduction in the amount of 20 and 22-kDa family zeins, and changes in the arrangement of proteins within protein bodies [39, 40]. In *fl2* kernels, the 24-kDa α-zein contains a signal peptide which is normally removed during post-translational processing [41]. In any case, the Esp1 protein would function in the translation to or the accumulation of prolamins after the transcription of prolamin genes. It is expected that the genetic mechanism for the accumulation of the prolamin polypeptides into protein bodies will be elucidated by the analysis of *esp1* mutation.

When examined by electron microscopy, it was found the spherical ER-derived PBs in rice endosperm possess a high electron-dense core with a surrounding lamellar structure [6, 8, 42]. It was also demonstrated that *esp1* mutant was deficient in CysP 13-kDa polypeptides; while levels of the CysR 10-kDa prolamin were elevated in *Esp4* mutants, but depressed in *esp3* mutants [13, 14]. The spherical PB in *esp3* and *Esp4* mutants lacked the typical lamellar structure, while the spherical PB of *esp1* mutants had the same structural properties as wild-type [15]. This suggests that CysR 10-kDa prolamins are essential for the formation of the core in ER-derived PBs, while the function of CysP prolamins in the formation of ER-derived PBs remains unknown. It is reported that the knock-down of 13-kDa prolamins not only reduced the size of ER-derived PB, but also altered the rugged peripheral structure in rice [43] and that a sufficient amount of the cysteine poor 22-kDa zeins is necessary for maintenance of a

normal PB shape in maize[44]. PBs in the *o2* maize mutant genotype typically are smaller than their normal genotype counterparts, due to the profound decrease in levels of CysP 22-kDa α-zein [45]. Although PBs in *o2* mutant maize were morphologically similar to those of the normal genotype, the endosperm of *o2* maize at 14 days after flowering contains tubular arrays within the rough ER. These tubular arrays are tightly associated with developing PBs [46]. Yeast two-hybrid analysis of the various zein classes in maize showed strong interactions between the CysR γ-zeins and CysP α-zeins indicating an important role for the γ-zein in the binding and assembly of α-zein within the PB [47]. From the fact that the CysR prolamins are involved in inter-chain disulfide bonding [48] and that the 13-kDa prolamin is unable to bind to other prolamins through disulfide bonds, it is speculated that CysR prolamins constitute the core and the whole scaffold and CysP and CysR prolamins interact with each other. Further analyses will be required to elucidate the mechanisms regarding the role of the *Esp1*gene in prolamin regulation in PBs.

Acknowledgments

We sincerely thank Professor Yoichiro Aso, Kyushu University, for valuable suggestions. This work was supported in part by a grant-in-aid for scientific research from the Japanese Society for the Promotion of Science (16830009 and 21380008); the program for Promotion of Basic Research Activities for Innovative Biosciences from the Bio-oriented Technology Research Advanced Institution (BRAIN).

References

[1] P.R. Shewry, R. Casey, Seed proteins, in: P.R. Shewry, R. Casey, (Eds), Seed Proteins, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp.

1-10.

[2] C. Soave, R. Reggiani, N. Di Fonzo, F. Salamini, Clustering of genes for 20 Kd zein subunits in the short arm of maize chromosome 7, Genetics 97 (1981) 363-377.

[3] C. Soave, R. Reggiani, N. Di Fonzo, F. Salamini, Genes for zein subunits on maize chromosome 4, Biochem Genet 20 (1982) 1027-1038.

[4] C. Soave, N. Suman, A. Viotti, F. Salamini, Linkage relationships between regulatory and structural gene loci involved in zein synthesis in maize, Theor Appl Genet 52 (1978) 263-267.

[5] P.R. Shewry, A.S. Tatham, The characteristics, structure and evolutionary relationships of prolamins, in: P.R. Shewry, R. Casey, (Eds), Seed Protein, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 11-33.

[6] M. Ogawa, T. Kumamaru, H. Sato, N. Iwata, T. Omura, Z. Kasai, K. Tanaka, Purification of protein body-I of rice seed and its polypeptide composition, Plant Cell Physiol 28 (1987) 1517-1527.

[7] K. Tanaka, T. Sugimoto, M. Ogawa, Z. Kasai, Isolation and characterization of two types of protein bodies in the rice endosperm, Agric. Biol. Chem. 44 (1980) 1633-1639.

[8] D.B. Bechtel, B.O. Juliano, Formation of protein bodies in the starchy endosperm of rice (*Oryza sativa* L.) : a re-investigation, Ann Bot 45 (1980) 503-509.

[9] L.F. Shyur, T.N. Wen, C.S. Chen, Purification and characterization of rice prolamins, Bot Bull Acad Sin 35 (1994) 65-71.

[10] T. Hibino, K. Kidzu, T. Masumura, K. Ohtsuki, K. Tanaka, M. Kawabata, Amino acid composition of rice prolamin polypeptides, Agric Biol Chem 53 (1989) 513-518.

[11] E. Gianazza, V. Viglienghi, P.G. Righetti, F. Salamini, C. Soave, Amino acid composition of zein molecular components, Phytochemistry 16 (1977) 315-317.

[12] A. Viotti, E. Sala, P. Alberi, C. Soave, Heterogeneity of zein synthesis in vitro, Plant Sci Lett 13 (1978) 365-375.

[13] T. Kumamaru, H. Satoh, N. Iwata, T. Omura, M. Ogawa, Mutants for rice storage proteins. III. Genetic analysis of mutants for storage proteins of protein bodies in the starchy endosperm, Jpn. J. Get. 62 (1987) 333-339.

[14] T. Kumamaru, H. Satoh, N. Iwata, T. Omura, M. Ogawa, K. Tanaka, Mutants for rice storage proteins. 1. Screening of mutants for rice storage proteins of protein bodies in the starchy endosperm, Theor. Appl. Genet. 76 (1988) 11-16.

[15] M. Ogawa, T. Kumamaru, H. Satoh, T. Omura, T. Park, K. Shintaku, K. Baba, Mutants for rice storage proteins. 2. Isolation and characterization of protein bodies from rice mutants, Theor. Appl. Genet. 78 (1989) 305-310.

[16] M. Satoh-Cruz, A.J. Crofts, Y. Takemoto-Kuno, A. Sugino, H. Washida, N. Crofts, T.W. Okita, M. Ogawa, H. Satoh, T. Kumamaru, Protein disulfide isomerase like 1-1 participates in the maturation of proglutelin within the endoplasmic reticulum in rice endosperm, Plant Cell Physiol. 51 (2010) 1581-1593.

[17] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680-685.

[18] A.C. Brinegar, D.M. Peterson, Separation and characterization of oat globulin polypeptides, Arch Biochem Biophys 219 (1982) 71-79.

[19] T. Kumamaru, Y. Uemura, Y. Inoue, Y. Takemoto, S.U. Siddiqui, M. Ogawa, I. Hara-Nishimura, H. Satoh, Vacuolar processing enzyme plays an essential role in the crystalline structure of glutelin in rice seed, Plant Cell Physiol 51 (2010) 38-46.

[20] P.H. O'Farrell, High resolution two-dimensional electrophoresis of proteins, J Biol Chem 250 (1975) 4007-4021.

[21] A. Sillero, J.M. Ribeiro, Isoelectric points of proteins: theoretical determination, Analytical biochemistry 179 (1989) 319-325.

[22] B. Bjellqvist, B. Basse, E. Olsen, J.E. Celis, Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions, Electrophoresis 15 (1994) 529-539.

[23] T. Masumura, T. Hibino, K. Kidzu, N. Mitsukawa, K. Tanaka, S. Fujii, Cloning and characterization of a cDNA encoding a rice 13 kDa prolamin, Mol Gen Genet 221 (1990) 1-7.

[24] S. Sha, Y. Sugiyama, N. Mitsukawa, T. Masumura, K. Tanaka, Cloning and sequencing of a rice gene encoding the 13-kDa prolamin polypeptide, Biosci Biotechnol Biochem 60 (1996) 335-337.

[25] T. Masumura, D. Shibata, T. Hibino, T. Kato, K. Kawabe, G. Takeba, K. Tanaka, S. Fujii, cDNA cloning of an mRNA encoding a sulfur-rich 10 kDa prolamin polypeptide in rice seeds, Plant Mol Biol 12 (1989) 123-130.

[26] L.F. Shyur, T.N. Wen, C.S. Chen, cDNA cloning and gene expression of the major prolamins of rice, Plant Mol Biol 20 (1992) 323-326.

[27] G.H. Feng, L. Wen, J.K. Huang, B.S. Shorrosh, S. Muthukrishnan, G.R. Reeck, Nucleotide sequence of a cloned rice genomic DNA fragment that encodes a 10 kDa prolamin polypeptide, Nucleic Acids Res 18 (1990) 683.

[28] W.T. Kim, T.W. Okita, Structure, expression, and heterogeneity of the rice seed prolamines, Plant Physiol 88 (1988) 649-655.

[29] L.F. Shyur, C.S. Chen, Nucleotide sequence of two rice prolamin cDNAs, Nucleic Acids Res 18 (1990) 6683.

[30] J.H. Xu, J. Messing, Amplification of prolamin storage protein genes in different subfamilies of the Poaceae, Theor Appl Genet 119 (2009) 1397-1412.

[31] R. Kodrzycki, R.S. Boston, B.A. Larkins, The opaque-2 mutation of maize differentially reduces zein gene transcription, The Plant cell 1 (1989) 105-114.

[32] H. Hartings, M. Maddaloni, N. Lazzaroni, N. Di Fonzo, M. Motto, F. Salamini, R. Thompson, The O2 gene which regulates zein deposition in maize endosperm encodes a protein with structural homologies to transcriptional activators, Embo J 8 (1989) 2795-2801.

[33] R.J. Schmidt, F.A. Burr, M.J. Aukerman, B. Burr, Maize regulatory gene opaque-2 encodes a protein with a "leucine-zipper" motif that binds to zein DNA, Proc Natl Acad Sci U S A 87 (1990) 46-50.

[34] C. Soave, F. Salamini, Organization and regulation of zein genes in maize endosperm, Phil Trans R Soc Lond 304 (1984) 341-347.

[35] Y. Onodera, A. Suzuki, C.Y. Wu, H. Washida, F. Takaiwa, A rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif, J Biol Chem 276 (2001) 14139-14152.

[36] T. Izawa, R. Foster, M. Nakajima, K. Shimamoto, N.H. Chua, The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development, Plant Cell 6 (1994) 1277-1287.

[37] M. Nakase, N. Aoki, T. Matsuda, T. Adachi, Characterization of a novel rice

bZIP protein which binds to the α-globulin promoter, Plant Mol Biol 33 (1997) 513-522.

[38] D.R. Holding, M.S. Otegui, B. Li, R.B. Meeley, T. Dam, B.G. Hunter, R. Jung, B.A. Larkins, The maize *floury1* gene encodes a novel endoplasmic reticulum protein involved in zein protein body formation, Plant Cell 19 (2007) 2569-2582.

[39] C.R. Lending, B.A. Larkins, Effect of the *floury-2* locus on protein body formation during maize endosperm development, Protoplasma 171 (1992) 123-133.

[40] O.E. Nelson, E.T. Mertz, L.S. Bates, Second mutant gene affecting the amino acid pattern of maize endosperm proteins, Science 150 (1965) 1469-1470.

[41] C.E. Coleman, M.A. Lopes, J.W. Gillikin, R.S. Boston, B.A. Larkins, A defective signal peptide in the maize high-lysine mutant floury 2, Proc Natl Acad Sci U S A 92 (1995) 6828-6831.

[42] H. Yamagata, K. Tanaka, The site of synthesis and accumulation of storage proteins, Plant Cell Physiol. 27 (1986) 135-145.

[43] T. Kawakatsu, S. Hirose, H. Yasuda, F. Takaiwa, Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation, Plant Physiol (2010).

[44] Y. Wu, J. Messing, RNA interference-mediated change in protein body morphology and seed opacity through loss of different zein proteins, Plant Physiol 153 (2010) 337-347.

[45] K.B. Geetha, C.R. Lending, M.A. Lopes, J.C. Wallace, B.A. Larkins, *opaque-2 modifiers* increase ©-zein synthesis and alter its spatial distribution in maize endosperm, Plant Cell 3 (1991) 1207-1219.

[46] C.R. Lending, A novel tubular array associated with protein bodies in the rough endoplasmic reticulum of *opaque-2* maize, Protoplasma 195 (1996) 68-77.

[47] C.S. Kim, Y.M. Woo Ym, A.M. Clore, R.J. Burnett, N.P. Carneiro, B.A. Larkins, Zein protein interactions, rather than the asymmetric distribution of zein mRNAs on endoplasmic reticulum membranes, influence protein body formation in maize endosperm, Plant Cell 14 (2002) 655-672.

[48] Y. Kawagoe, K. Suzuki, M. Tasaki, H. Yasuda, K. Akagi, E. Katoh, N.K. Nishizawa, M. Ogawa, F. Takaiwa, The critical role of disulfide bond formation in protein sorting in the endosperm of rice, Plant Cell 17 (2005) 1141-1153.

Figure Legends

Fig. 1. SDS-PAGE analysis of prolamin polypeptides in wild-type and CM21 mutant seeds. (A) Total protein, (B) Total prolamin, (C) CysP prolamin, (D) CysR prolamin.

Fig. 2. IEF analysis of CysP prolamin polypeptides in wild type and CM21 mutant seeds.

Fig. 3. IEF analysis of CysP prolamins in F_2 seeds resulting from the cross between wild-type and CM21. + and - in F_2 seeds indicate the Kinmaze type and *esp1* mutant type by SDS-PAGE, respectively. Circles indicate the CysP prolamin bands reduced in the *esp1* mutant.

Fig. 4. Two-dimensional electrophoresis (IEF/SDS) analysis of CysP prolamin polypeptides in wild type showing the reduced level in the *esp1* mutant.

Fig. 5. Chromosome localization of genes encoding CysP prolamins showing reduced level in *esp1* mutants. 1: CSSL9918, 2: CSSL9919, 3: CSSL9920, 4: CSSL9937. (A) IEF analysis of CysP prolamins in seeds of CSSLs. Arrowheads indicate the CysP prolamin bands absent in CSSL9918, CSSL9919, and CSSL9920. White solid lines indicate the position of wild type Kinmaze-specific pI 7.10 CysP prolamin bands, while dotted white lines indicate the position of Kasalath-specific pI 7.12 CysP prolamin bands. (B) Graphical genotype of chromosome 5 of CSSLs. Open and closed boxes indicate the chromosomal segment derived from Nipponbare and Kasalath, respectively.

R1838 and R2289 are DNA markers on chromosome 5. Number in parentheses indicate the locus of DNA marker labeled at distance of cM from short arm end.

Fig. 6. RT-PCR analysis of prolamin gene expression.

RNAs from the developing seeds at 2 weeks after flowering were used for RT-PCR. CM21 and EM711 are *esp1* allelic lines. The original variety of CM21 and EM711 are Kinmaze and Taichung65, respectively.

Supplemental Fig. 1. Chromosome localization of genes encoding CysP prolamins showing non-reducing levels in *esp1* mutants. 1: CSSL99102, 2: CSSL99107, 3: CSSL9927. (A) IEF analysis of CysP prolamins in seeds of CSSLs. Arrowheads indicate the CysP prolamin bands deleted in CSSL99102 and CSSL99107. (B) Graphical genotype of chromosome 7 of CSSLs. Open and closed boxes indicate the chromosomal segment derived from Nipponbare and Kasalath, respectively. R565 and R1440 are DNA markers on chromosome 7. Number in parentheses indicates the the distance from short arm end.

Supplemental Fig. 2. Two-dimensional electrophoresis (IEF/SDS) analysis of CysP prolamin polypeptides in wild type showing same level in *esp1* mutant.

Legends to Tables

Table 1. Identification of the gene by MALDI-TOF MS analysis of the prolamin polypeptides in the wild-type corresponding to CysP prolamin polypeptides showing

reduced level in *esp1* mutant.

Theoretical mass value indicates the mass value of the theoretical trypsin-digested peptide fragment deduced from prolamin sequences of their DNA sequence in Data base. Theoretical mass values are evaluated by ExPASy Proteomics Server (http://expasy.org/tools/pi_tool.html). The Predicted Gene corresponding to each pI band was decided based on the combination of theoretical mass value. Whole of the theoretical mass value of the trypsin-digested peptide fragment deduced from the prolamin genes is shown in Supplemental Table 3.

Os05g0329100 and Os05g0329300 are the same sequence.

Supplemental Table 1. The sequence of specific primers for the prolamin genes for the RT-PCR.

Supplemental Table 2. Rice prolamin genes searched from the NCBI genomic database and rice annotation project database (RAP-DB).

Supplemental Table 3. Theoretical mass values of trypsin digested peptide fragment deduced from all prolamin sequences searched from the NCBI genomic database and rice annotation project database (RAP-DB).

Theoretical mass values that matched an apparent mass value of the trypsin digested peptide fragment detected by MALDI-TOF MS analysis of the bands with pIs, 6.65, 6.95, 7.10, and 7.35, reducing in *esp1* mutant in Table 1 are highlighted in red and the identified genes are indicated in yellow back ground. Theoretical mass values that matched an apparent mass value of the trypsin digested peptide fragment detected by

MALDI-TOF MS analysis of the bands with pIs, 6.6o and 7.85 showing the normal level in *esp1* mutant in Table 1 are highlighted in blue and the identified genes are indicated in green back ground.

*:Mass value of peptide fragment by trypsin digestion deduced from all prolamin sequences.

Supplemental Table 4. Theoretical pI and molecular mass of the protein deduced from prolamin sequence except for the signal peptides.

Supplemental Table 5. Identification of the gene by MALDI-TOF MS analysis of the prolamin polypeptides in the wild-type corresponding to CysP prolamin polypeptides showing same level in *esp1* mutant.

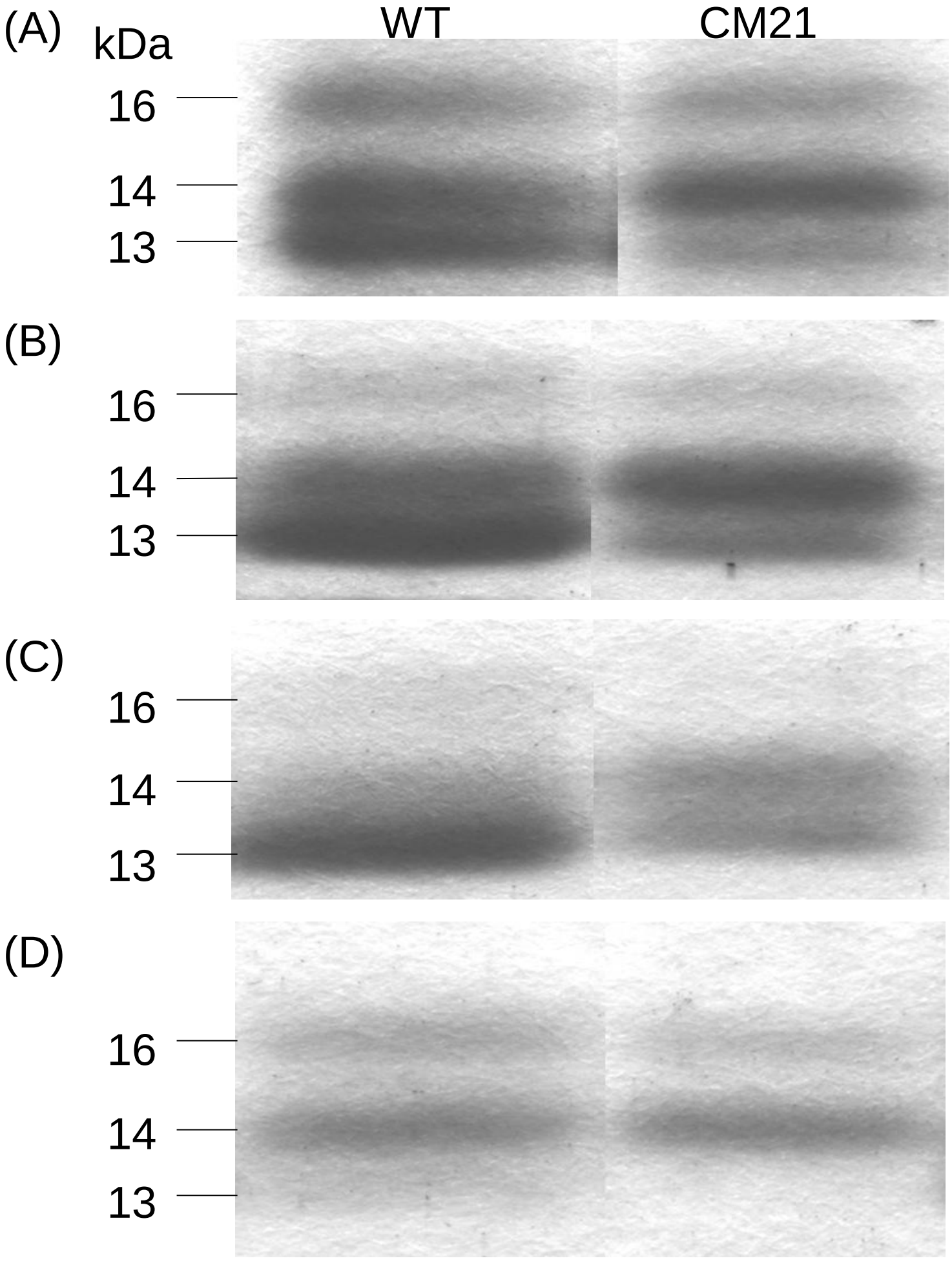


Fig. 1

Ushijima et al

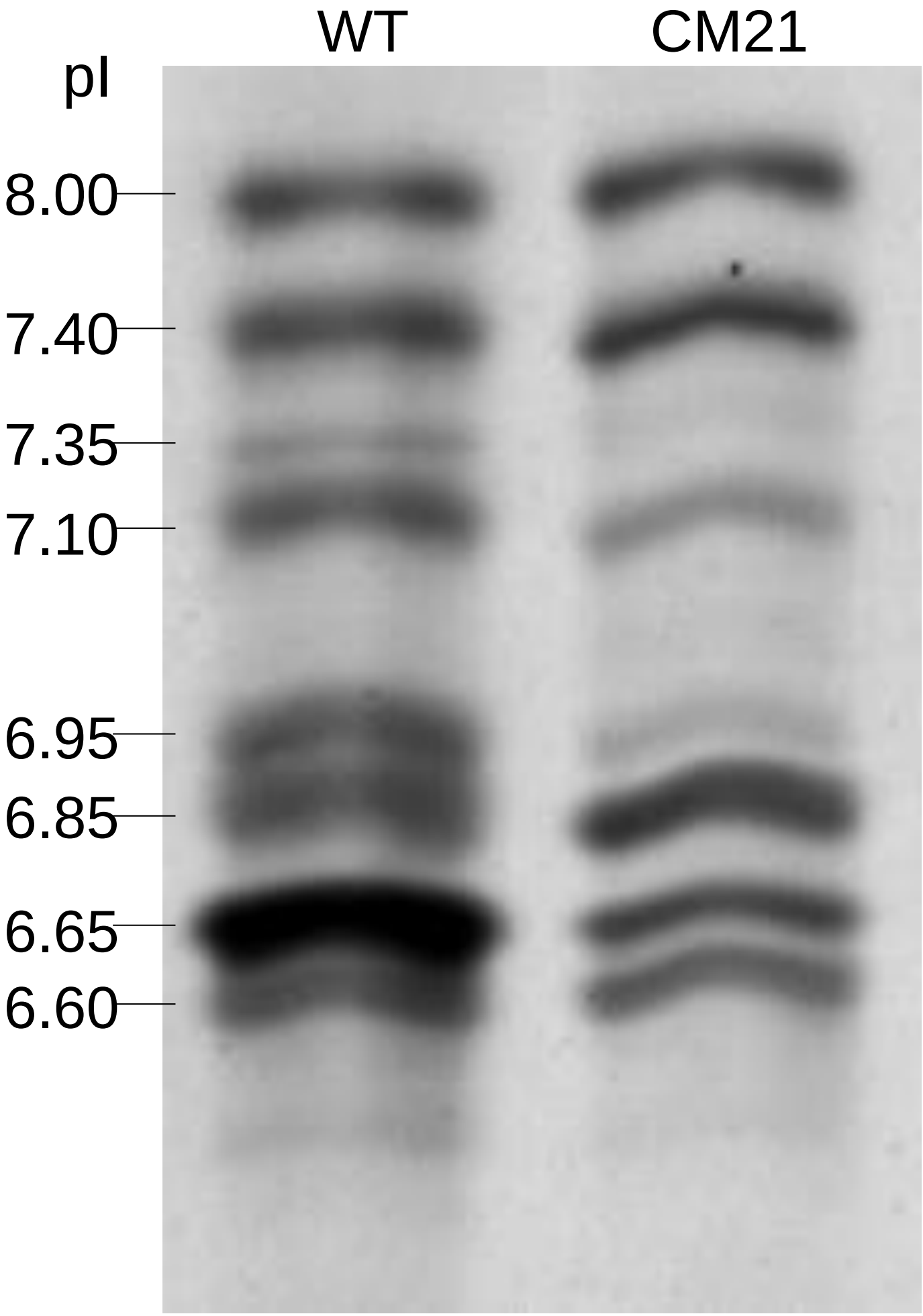


Fig. 2Ushijima et al

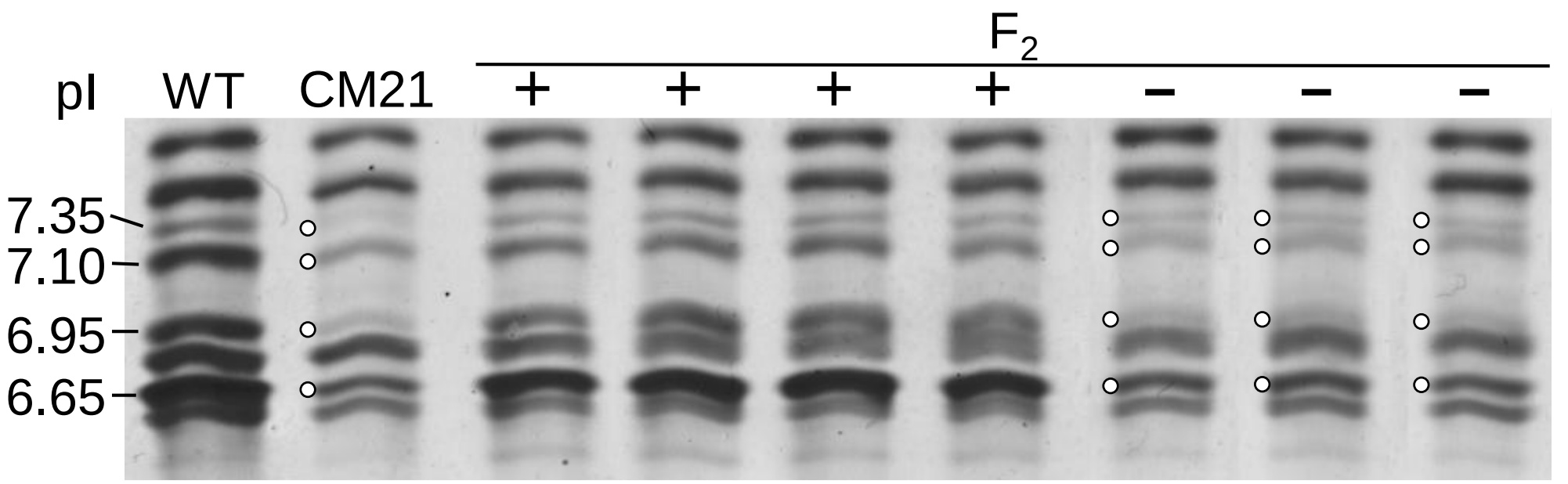


Fig. 3 Ushijima et al

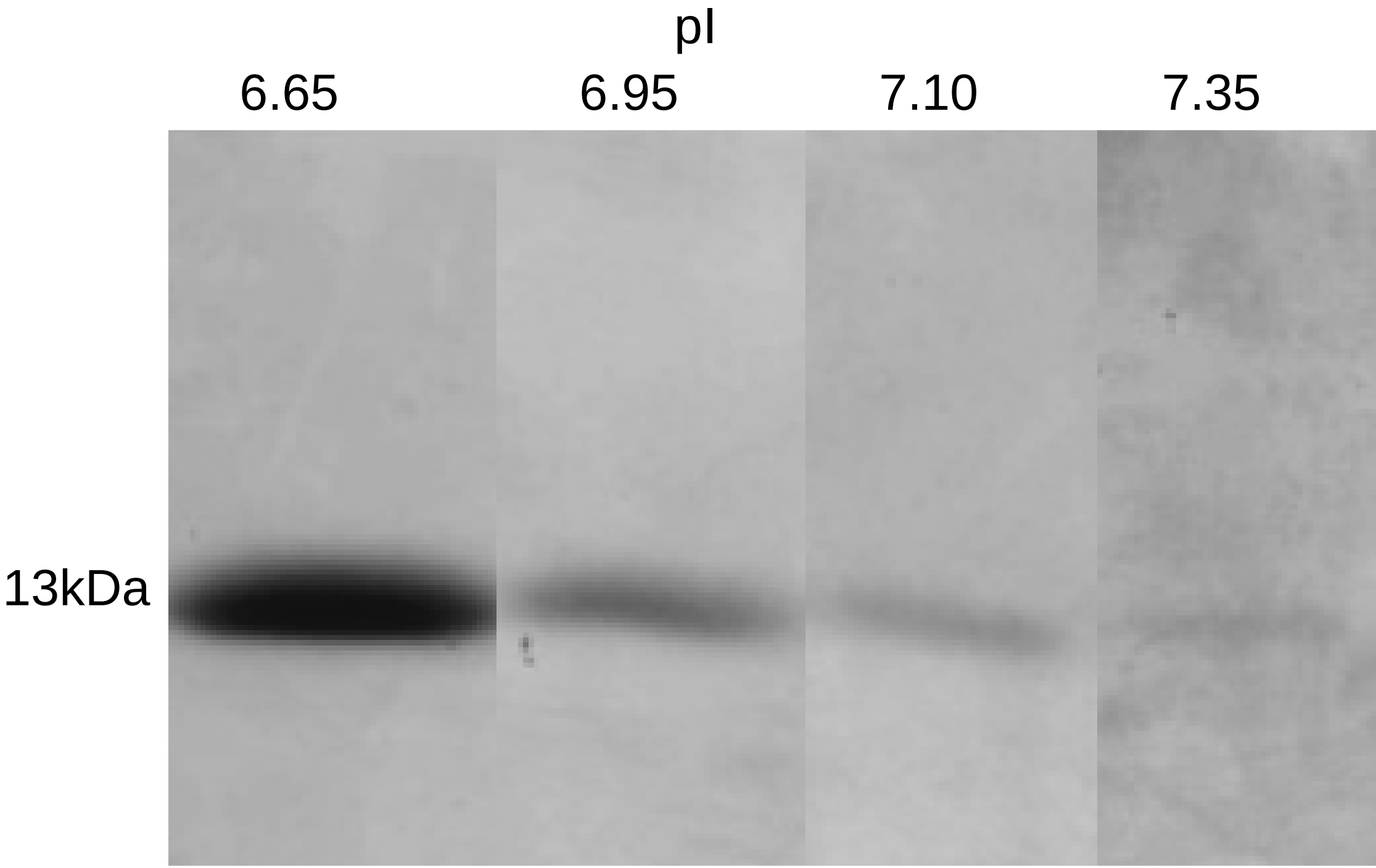


Fig. 4 Ushijima et al

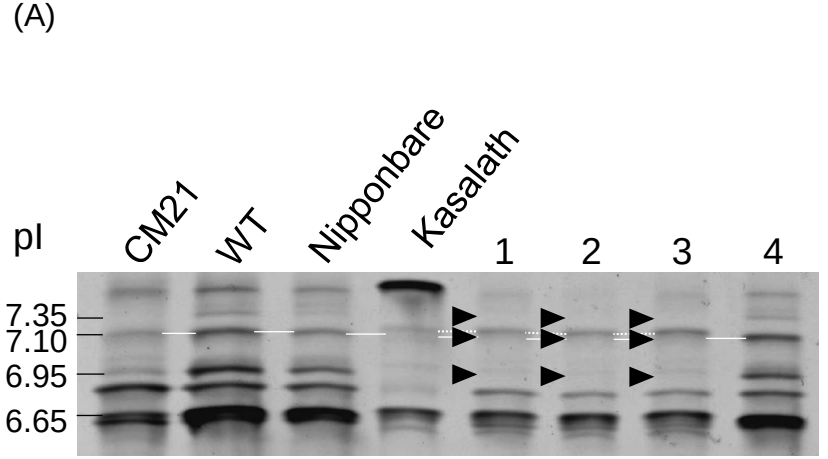


Fig. 5 Ushijima et al

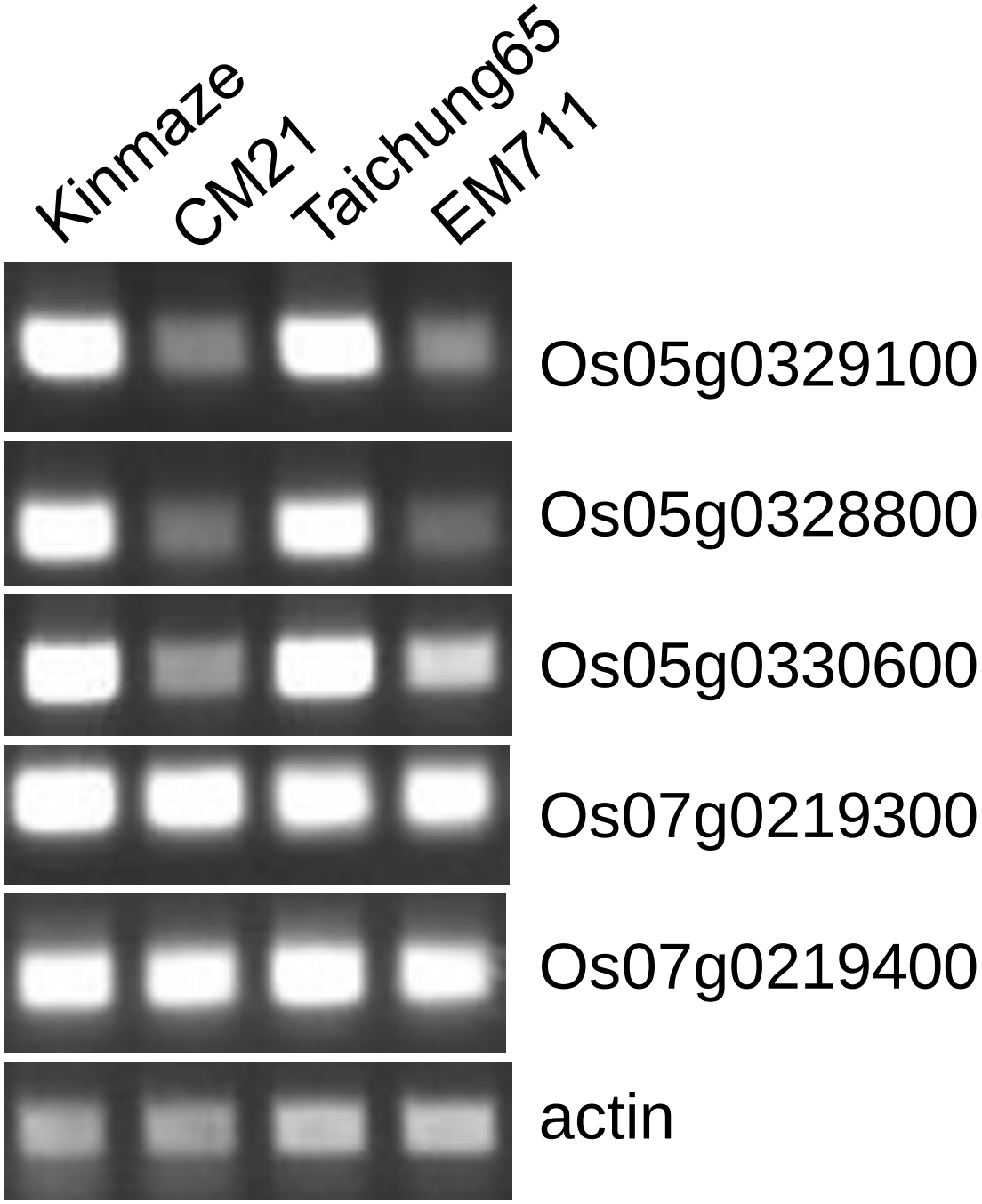
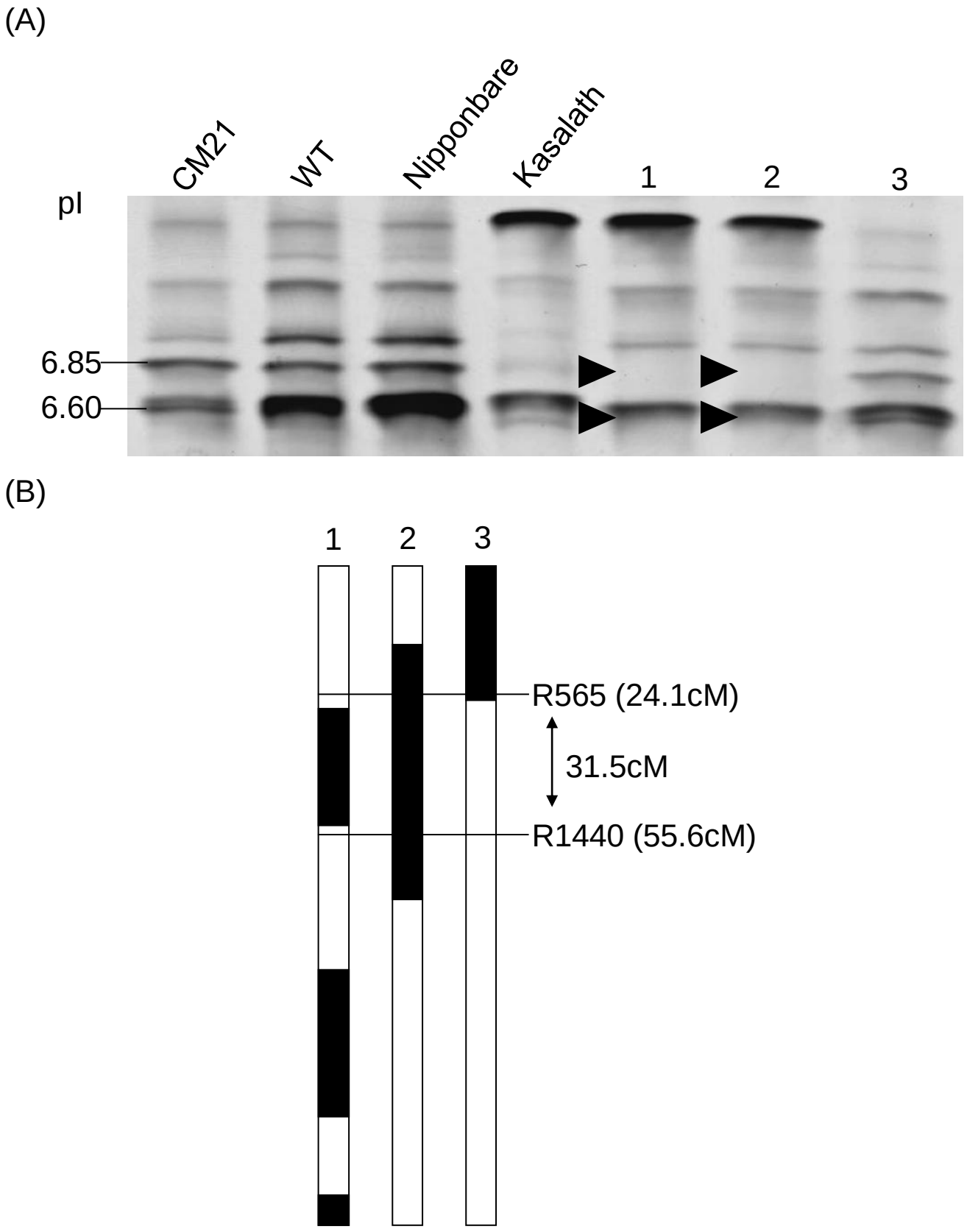
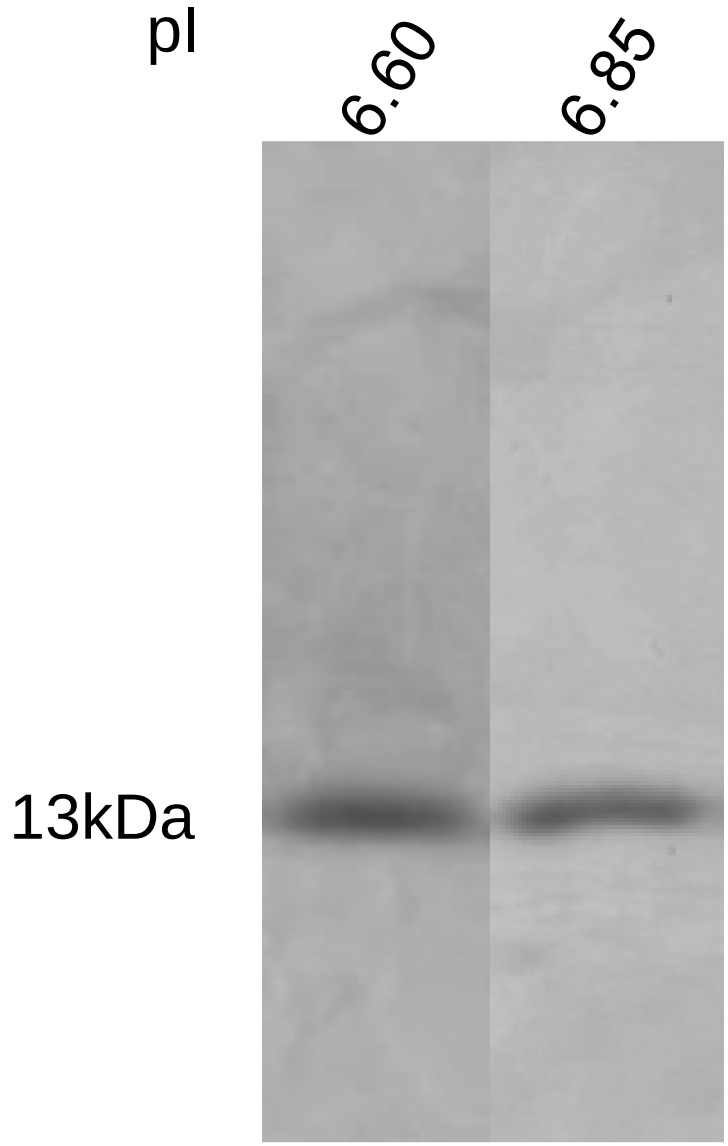

Kinmaze
CM21
Taichung65
EM711
Os05g0329100
Os05g0328800
Os05g0330600
Os07g0219300
Os07g0219400
actin

Fig6



Supplemental Fig. 1. Chromosome localization of genes encoding CysP prolamins with non-reduced levels in *esp1* mutants. Wild-type: WT; 1: CSSL99102; 2: CSSL99107; 3: CSSL9927. (A) IEF analysis of CysP prolamins in seeds of CSSLs. The arrowheads indicate the CysP prolamin bands deleted in CSSL99102 and CSSL99107. (B) Graphical genotype of chromosome 7 of the CSSLs. The open and closed boxes indicate the chromosomal segment derived from Nipponbare and Kasalath, respectively. R565 and R1440 are DNA markers on chromosome 7. The numbers in parentheses indicate the distance from the short arm end of the chromosome.



Supplemental Fig. 2. Two-dimensional electrophoresis (IEF/SDS) analysis of CysP prolamin polypeptides in wild-type showing the same level in the *esp1* mutant.

Table 1. Identification of genes by MALDI-TOF MS analysis of the prolamin polypeptides in the wild-type sample corresponding to those CysP prolamin polypeptides with reduced levels in the *esp1* mutant.

	Mass Value of tripsin digested peptid		Predicted Gene Name
pI of IEF band	Apparent	Theoretical	
6.65	1195.38	1212.33	Os05g0329100 (Os05g0329300)
	1713.35	1712.97	
	2096.80	2096.37	
	2779.31	2778.16	
	4978.26	4974.53	
6.95	1195.38	1212.33	Os05g0328800
	1713.43	1712.97	
	2096.80	2096.37	
	2787.80	2787.17	
	4975.21	4974.53	
7.10	1212.31	1212.33	Os05g0329100(Os05g032930(or Os05g0330600 and Os05g0329700
	1713.35	1712.97	
	2096.53	2096.37	
	2124.59	2124.43	
	2778.37	2778.16	
	4977.06	4974.53 (4988.55)	
7.35	1212.35	1212.33	Os05g0330600
	1713.35	1712.97	
	2096.44	2096.37	
	4981.67	4988.55	

The theoretical mass value indicates the mass value of the theoretical trypsin-digested peptide fragment deduced from the prolamin sequences of their DNA sequences in the NCBI database. The theoretical mass values were evaluated using the ExPASy Proteomics Server (http://expasy.org/tools/pi_tool.html). The gene ID corresponding to each pI band was assigned based on the combination of its theoretical mass value. The whole theoretical mass values of the trypsin-digested fragments deduced from the prolamin genes are shown in Supplemental Table 3.

Of note, Os05g0329100 and Os05g0329300 are the same sequence.

Supplemental Table 1. The sequence of the specific primers for the prolamin genes used for RT-PCR analysis.

Gene Name	sequence
Os05g0329100	5-'CTGTCCTGCTACAGCAACAG-3'
	5'-ATTGTTGCGCCACCAGCGC-3'
Os05g0328800	5'-CTGTCCTGCTACAGCAACAT-3'
	5'-ATTGTTGCGCCACCAGCGC-3'
Os05g0330600	5'-TCTGCGCAGTTTGATGTTTTG-3'
	5-'CTGGAGTTGTAGCTGCTGCA-3'
Os07g0219300	5'-TTGTCCAGGCCATAGCGCAC-3'
	5'-GCTATCGGGTGCACTATAGGACCA-3'
Os07g0219400	5'-TTGTCCAGGCCATAGCGCAG-3'
	5'-CAAGGGTGGTAATACTACCC-3'
actin	5'-TCCATCTTGGCATCTCTCAG-3'
	5'-GTACCCGCATCAGGCATCTG-3'

Supplemental Table 2. Rice prolamin genes obtained from the NCBI database and the Rice Annotation Project Database (RAP-DB).

Gene Name	Function	chr.	Predicted mature aa	Cysteine residues	Accession No. of supporting cDNA clone	Reference
Os03g0766100	10 kDa prolamin precursor	3	110	11	AK108254	
Os05g0328800	Prolamin 7	5	131	0	X53857	Shyur *et al.* (1990)
Os05g0329100	Prolamin	5	131	0	AK242260	
Os05g0329300	Prolamin 7	5	131	0	AF042201	
Os05g0329400	Similar to prolamin	5	134	2	(EST clone)	
Os05g0329700	Prolamin precursor	5	131	0	AK242910	
Os05g0330600	Prolamin PPROL 7 precursor	5	131	0	(EST clone)	
Os05g0331532	Similar to prolamin	5	134	0	AK242154	
Os05g0331550	Similar to prolamin	5	134	0	(EST clone)	
Os05g0332000	Prolamin 7	5	134	0	AK242285	
Os06g0507100	Similar to prolamin	6	130	11	AK242322	
Os06g0507200	Prolamin	6	130	13	AK107785	
Os07g0206400	13 kDa prolamin precursor	7	137	4	X14392	Masumura *et al.* (1989)
Os07g0206500	13 kDa prolamin precursor	7	137	4	X60979	Shyur *et al.* (1992)
Os07g0219300	Prolamin precursor (13 kDa prolamin)	7	132	0	D11385	Yamagata *et al.* (1992)*
Os07g0219400	Prolamin	7	132	0	X53856	Shyur *et al.* (1990)
Os07g0220050	Similar to prolamin	7	132	0	(EST clone)	
Os11g0535525	Similar to 10 kDa prolamin	11	111	9	AK242313	
Os12g0269101	Similar to prolamin	12	139	8	(EST clone)	
Os12g0269200	13 kDa prolamin precursor	12	137	8	AB016505	Mitsukawa *et al.* (1999)**

*Yamagata H, Nomura T, Arai S, Tanaka K, Iwasaki T (1992) Nucleotide sequence of a cDNA that encodes a rice prolamin. Biosci. Biotechnol. Biochem. 56: 537

**Mitsukawa N, Konishi R, Kidzu K, Ohtsuki K, Masumura T, Tanaka K (1999) Amino acid sequencing and cDNA cloning of rice seed storage proteins, the 13 kDa prolamins, extracted from Type I protein bodies. Plant Biotechnology 16: 103-113

Supplemental Table 3. Theoretical mass values of the trypsin-digested peptide fragments deduced from all prolamin sequences obtained from the NCBI database and the Rice Annotation Project Database (RAP-DB). The theoretical mass values that matched the apparent mass values of the trypsin-digested fragments detected by MALDI-TOF MS analysis of the bands with pIs 6.65, 6.95, 7.10, and 7.35, which were reduced in the *esp1* mutant (as shown in Table 1), are highlighted in red and the identified genes are highlighted in yellow. The theoretical mass values that matched the apparent mass values of the trypsin-digested fragments detected by MALDI-TOF MS analysis of the bands with pIs 6.60 and 7.85, which showed normal levels in the *esp1* mutant (as shown in Table 1) are highlighted in blue and the identified genes are highlighted in green.

	CysP Prolamin										CysR Prolamin							
	chr. 5							chr. 7			chr. 3	chr. 6		chr. 7		chr. 11	chr. 12	
All fragment*/Gene Nam	Os05g032880	Os05g0329100 (Os05g03293	Os05g032940	Os05g032970	Os05g033060	Os05g0331532(Os05g03315	Os05g0332000	Os07g021930	Os07g021940	Os07g0220050	Os03g0766100	Os06g0507100	Os06g0507200	Os07g0206400	Os07g0206500	Os11g0535525	Os12g0269100	Os12g0269200
636.74												636.74						
767.88	767.88	767.88		767.88	767.88	767.88	767.88											
794.91																	794.91	
800.94													800.94					
909.07												909.07						
917.01													917.01					
923.11												923.11						
1024.20													1024.20					
1052.20														1052.20				
1078.23															1078.23			
1086.17															1086.17			
1112.21														1112.21				
1112.25								1112.25	1112.25	1112.25								
1164.40													1164.40					
1185.31						1185.31	1185.31											
1189.34								1189.34	1189.34	1189.34								
1212.33	1212.33	1212.33		1212.33	1212.33	1212.33	1212.33											1226.35
1226.35																		
1258.41			1258.41															
1327.54												1327.54						
1423.55																	1423.55	
1512.72	1512.72	1512.72		1512.72	1512.72													
1653.83												1653.83						
1661.87									1661.87	1661.87								
1676.88													1676.88					
1691.90						1691.90	1691.90											
1709.99																	1709.99	1709.99
1712.97	1712.97	1712.97	1712.97	1712.97	1712.97													
1878.16							1878.16											
1892.19						1892.19												
2034.34							2034.34											
2062.36						2062.36												
2096.37	2096.37	2096.37	2096.37		2096.37													
2124.43				2124.43														
2130.39								2130.39										
2140.43										2140.43								
2154.45									2154.45									
2184.50																		2184.50
2254.59																	2254.59	
2274.71											2274.71							
2278.62																2278.62		
2421.78									2421.78	2421.78								
2439.83																		2439.83
2632.09			2632.09															
2661.25											2661.25							
2745.15														2745.15	2745.15			
2775.25																	2775.25	2775.25
2778.16		2778.16	2778.16	2778.16	2778.16			2778.16										
2787.17	2787.17																	
2794.20									2794.20	2794.20								
2801.20						2801.20	2801.20											
3298.68												3298.68						
3324.78													3324.78					
3853.31									3853.31	3853.31								
3861.33								3861.33										
3892.34						3892.34												
3915.48															3915.48			
3941.51														3941.51				
3950.38							3950.38											
4077.60								4077.60										
4966.84												4966.84						
4974.53	4974.53	4974.53	4974.53	4974.53														
4988.55					4988.55													
5346.11																		5346.11
5830.76													5830.76					
6762.75															6762.75			
6892.80														6892.80				
7038.16																	7038.16	
7365.88											7365.88							
10187.19																10187.19		

*: The mass values of trypsin-digested peptide fragments were deduced from all prolamin sequences.

Supplemental Table 4. Theoretical pI and molecular mass of the protein deduced from prolamin sequence except for the signal peptides.

Gene Name	pI	Molecular Mass (dalton)
Os05g0328800	8.23	14955.88
Os05g0329100 (Os05g0329300)	8.22	14946.87
Os05g0329700	8.22	14974.92
Os05g0330600	8.22	14960.90
Os07g0219300	7.02	15059.00
Os07g0219400	8.23	15079.12
Os07g0220050	8.23	15065.09

Supplemental Table 5. MALDI-TOF MS analysis identification of the genes of the prolamin polypeptides in the wild-type that corresponded to CysP prolamin polypeptides with similar levels in the *esp1* mutant.

pI of IEF band	Mass Value of tripsin digested peptide		Predicted Gene Name
	Apparent	Theoretical	
6.60	2130.47	2130.39	Os07g0219300
	2778.17	2777.16	
	3866.14	(3861.33)	
6.85	2143.81	2140.43 (2154.45)	Os07g0220050 or Os07g0219400
	2422.14	2421.78	
	2794.75	2794.20	
	3859.45	(3853.31)	