Functional analysis of genes involved in the regulation of development of reproductive organs in rice (Oryza sativa)
Chen, Y.

Citation
Chen, Y. (2011, December 20). *Functional analysis of genes involved in the regulation of development of reproductive organs in rice (Oryza sativa)*. Retrieved from https://hdl.handle.net/1887/18262

Identification of two CCCH type zinc finger proteins as repressors for expression of the seed glutelin gene *GluB-1* in rice (*Oryza sativa*)

Yi Chen*[1], Aijun Sun*[1,2], Mei Wang[1,3], Zhen Zhu[2], Pieter B.F. Ouwerkerk[1]

[1] Department of Molecular and Developmental Genetics, Institute Biology Leiden, Leiden University, Sylvius Laboratory, Sylviusweg 72, PO Box 9505, 2300 RA Leiden, The Netherlands
[2] State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, No.1 West Beichen Road, Chaoyang District, Beijing, 100101 China
[3] SU BioMedicine-TNO, Utrechtseweg 48, 3704 HE Zeist, PO Box 360, 3700 AJ Zeist, The Netherlands

* These authors contribute equally to the work

Abstract

Glutelins are the most abundant storage proteins in rice (*Oryza sativa*) grains and make up to 80% of total grain protein. The glutelin genes in rice are divided into four subfamilies, GluA, B, C and D. As a model, the promoter region of the *GluB-1* gene has been intensively analyzed in the past decade; yet limited knowledge of transcription factors has been uncovered till now. We describe here two novel Cys3His1 (CCCH or C3H) zinc finger genes, *OsGZF1* and *OsGZF2*, which were identified in a yeast one-hybrid screening using the core promoter region of *GluB-1* as bait and cDNA expression libraries prepared of developing rice panicle and seed as prey. The OsGZF proteins were found to be able to bind specifically to the bait in yeast and the interaction was confirmed *in vitro* by electrophoretic mobility shift assays (EMSA). Northern blots and analysis of promoter beta-glucuronidase (GUS) transgenics showed that both genes were predominantly expressed in a thin layer surrounding the scutellum of the developing embryos. GUS activity was also observed in embryo, shoot and root of the germinating seeds, indicating their diverse functions other than regulating *GluB-1* expression. *GluB-1* promoter driven *GUS* reporter constructs co-transformed with *CaMV 35S* promoter driven *OsGZF1* or *OsGZF2* respectively into rice seedling protoplasts demonstrated that both proteins can down-regulate reporter gene expression. *OsGZF1* and *OsGZF2* were also able to reduce activation caused by a *GluB-1* activator, RISBZ1 in transient assays. Furthermore, nitrogen concentration of grains from *OsGZF1* RNAi plants was significantly higher than in the wild type. The results strongly suggest that OsGZF1 and OsGZF2 have functions as repressors of the *GluB-1* promoter.

Introduction

Rice (*Oryza sativa*) grain is one of the most important staple food sources consumed everyday by billions of people in the world. Because the rice grain contains merely 5-12% of seed storage protein (Villareal and Juliano, 1978) improvement and increase of protein content and thus of the nutritional value of rice, is of great significance for developing countries. Other than their nutritional value for mankind, seed storage proteins also have crucial functions on the texture of the grain. Likely protein-starch interactions can impede starch gelatinization and disruption of the structure of proteins can increase the stickyness of cooked rice (Hamaker et al., 1991; Saleh and Meullenet, 2007). In addition, the protein bodies in japonica rice were found concentrated near the cell wall whereas those in indica rice were scattered around amyloplasts. This ultrastructural aspect of the grain may affect the texture of cooked rice (Kang et al., 2006). Therefore, understanding of the mechanism behind storage protein formation and its regulation is not only critical for scientific research but also tremendously helpful for improving nutritional quality of rice. Up to 80% of the total seed storage protein in rice is glutelin (Yamagata et al., 1982) and thus it is this protein that has a decisive role in determining grain quality and cooking traits. Encoded by about 15 genes per haploid genome, glutelin genes are divided into four subfamilies, GluA, B, C and D (Takaiwa, 1987, 1991b; Okita et al., 1989; Mitsukawa et al., 1998; Kusaba et al., 2003; Katsube-Tanaka et al., 2004; Kawakatsu et al., 2008). All rice glutelin genes are expressed exclusively during seed development and their mRNA reach maximum levels during the filling stage (10-16 days after flowering) (Okita et al., 1989; Takaiwa and Oono, 1990, 1991; Mitsukawa et al., 1998; Duan and Sun, 2005;

Kawakatsu et al., 2008).

The expression pattern of glutelin is essentially the same as for prolamin and albumin, which are other seed storage proteins, suggesting that they are coordinately regulated and controlled by the same transcriptional regulatory machinery (Nakase et al., 1996; Duan and Sun, 2005). Comparison of their promoters showed the presence of conserved *cis*-regulatory elements, not only in rice but also in the orthologous genes of other cereals. For instance, cereal prolamin genes have a conserved region at ~300 bp upstream of the transcriptional start. This so-called endosperm box is essentially composed of two closely located motifs: a prolamine box (PROL) class endosperm motif (Vicente-Carbajosa et al., 1997) and a GCN4-like motif (Kreis et al., 1985). These sequences are so well conserved that they can also be recognized by transcription factors from other cereals. For instance, the GCN4 motif in the rice *GluB-1* promoter can be recognized by the bZip protein Opaque-2 (O2) from maize (Wu et al., 1998b; Hwang et al., 2004). O2 can also activate high levels of transcription from the wheat GCN4-like motif in plant protoplasts and yeast cells (Holdsworth et al., 1995). The prolamin-box binding factor (PBF) of maize is able to bind to the PROL motif and interacts with maize O2 protein *in vivo* (Vicente-Carbajosa et al., 1997). Therefore, it is likely that, investigation of one particular *cis*-acting element or transcription factor in rice, will also provide valuable information on regulation of seed storage proteins in general.

Other well-known *cis*-acting elements identified in promoters of seed storage protein genes include AACA, ACGT, G-box, GCAA, and Skn-I. *In vivo* transient assays and promoter-GUS transgenic rice confirmed that these sequences are essential for binding of transcription factors and are affecting expression of downstream reporter genes (Kim and Wu, 1990; Takaiwa et al., 1991a, 1996; Zheng et al., 1993; Yoshihara and Takaiwa, 1996; Yoshihara et al., 1996; Suzuki et al., 1998; Washida et al., 1999; Wu et al., 2000; Kawakatsu et al., 2008). The regulation of one particular member of the glutelins, *GluB-1*, has been studied into more detail and as a result a 245 bp core promoter was identified (Takaiwa et al., 1996). Within this region, the so-called AACA and GCN4 motifs are the critical elements driving seed-specific expression. An ACGT motif and a Skn-I-like element also play important roles (Washida et al., 1999). Later on, gain-of-function analysis with transgenic rice confirmed that the GCN4 motif in the core promoter acts as a decisive element in determining endosperm-specific expression. The flanking AACA, ACGT and PROL-boxes are involved in quantitative regulation of *GluB-1* (Wu et al., 2000).

Despite the many studies on promoter regions of different storage proteins, only few corresponding transcription factors have been identified so far. O2 (Hartings et al., 1989; Schmidt et al., 1990) is one of the first intensively investigated seed storage protein activators in cereals. Target binding sites of maize O2 include sequences with or without an ACGT core (Lohmer et al., 1991; Schmidt et al., 1992; Ueda et al., 1992; Cord Neto et al., 1995; Muth et al., 1996). In rice, a bZIP protein RITA-1 is capable of binding to an ACGT element thereby activating reporter gene expression in transient assays. RITA-1 is highly expressed in the aleurone of developing endosperm which indicates its potential role in seed storage protein synthesis in that tissue (Izawa et al., 1994). In addition, the AACA sequence in glutelin gene promoters is the target site for OsMYB5 (Suzuki et al., 1998). Five different bZIP proteins, RISBZ1 to RISBZ5 are able to interact with the GCN4 motif and RISBZ1 is capable of activating the expression of a reporter gene preceded by a minimal promoter fused to a pentamer of the GCN4 motif (Onodera et al., 2001; Kawakatsu et al., 2008). The Dof (DNA binding with one finger) prolamin box-binding factor (RPBF) is able to recognize

AAAG/CTTT motifs in the *GluB-1* promoter. RISBZ1 and RPBF both can *trans*-activate GUS activity driven by promoters of different storage protein genes in transient assays, including *GluA-1, GluA-2, GluA-3*, *GluB-1*, *GluD-1*, 10 kD Prolamin, 13 kD Prolamin, 16 kD Prolamin, and *α* Globulin. Synergistic interactions between RISBZ1 and RPBF were also studied in transient assays, as well as by transgenic rice in which RISBZ1 and RPBF RNA were silenced (Yamamoto et al., 2006; Kawakatsu et al., 2008, 2009).

Although several activators of genes encoding seed storage proteins have been identified so far, none of these are shown to have negative regulatory properties. Here we report on two novel CCCH type zinc finger proteins, OsGZF1 and OsGZF2 that act as repressors of *GluB-1* expression. Both proteins can interact specifically with the promoter region of *GluB-1* and are highly expressed in seeds of filling stage which coincides with *GluB-1* expression. Transient assays in rice protoplasts showed functions of *OsGZF1* and *OsGZF2* in down-regulation of the *GluB-1* promoter. Furthermore, it was found that both OsGZF1 and OsGZF2 could significantly decrease the *GluB-1* activator RISBZ1-mediated activation of *GluB-1* promoter expression in a transient expression system. Furthermore, a loss-of-function approach using *OsGZF1* RNAi transgenic plants showed increased seed nitrogen concentration. Based on these results we conclude that OsGZF1 and OsGZF2 can both act as repressors of *GluB-1*. Their regulatory modes on levels of seed storage proteins are discussed.

Results

Identification of two novel CCCH zinc finger transcription factors interacting with the *GluB-1* core-promoter

In order to further understand the regulatory mechanism controlling transcriptional expression of the seed storage protein *GluB-1*, yeast one-hybrid screenings were employed to search for novel transcription factors. Two cDNA expression libraries, derived from seeds and panicle from the filling stage of rice, were constructed and used as preys. cDNA was synthesized from poly(A)$^+$ RNA and separated by a Sepharose drip column (Figure 1A). The fragments larger than 300 bp were used in the library construction. The cDNA insertion efficiencies were analyzed by restriction digestion of the inserted fragments from the pACTII plasmids. Panicle library has about 60% (14/23) of insertions from several hundred base pairs to about 5 kb. However, seed library has merely half of cloning efficiency (6/23) as panicle library (Figure 1B).

Seven baits of different lengths were generated from ATG upstream region of *GluB-1* (Figure 2A) and cloned before a *HIS3* reporter gene. The constructs were then integrated into the genome of yeast strain Y187 (Clontech) through homologous recombination. After checking on a CM medium lacking histidine, only a 59 bp bait ProGluB-1-S2 containing a GCN4 and a PROL motif was suitable for further screenings. On the contrary, all the other constructs showed no leaky expression of the HIS3 reporter which represents endogenous yeast proteins repressed the baits. To suppress leaky expression, strain Y187/ProGluB-1-S2 was tested on histidine-lacking medium supplied with a series of concentrations of 3-aminotriazole (3-AT) and 10mM of 3-AT was found enough to suppress the leaky expression. Using this condition, rice panicle and seed cDNA libraries were screened. About 10^6 transformants of each library resulted in 131 positive colonies after re-streaking the primary positive colonies on the same selective medium. Positive colonies that could grow again after this initial selection step were further analyzed either by colony PCR or plasmid rescue
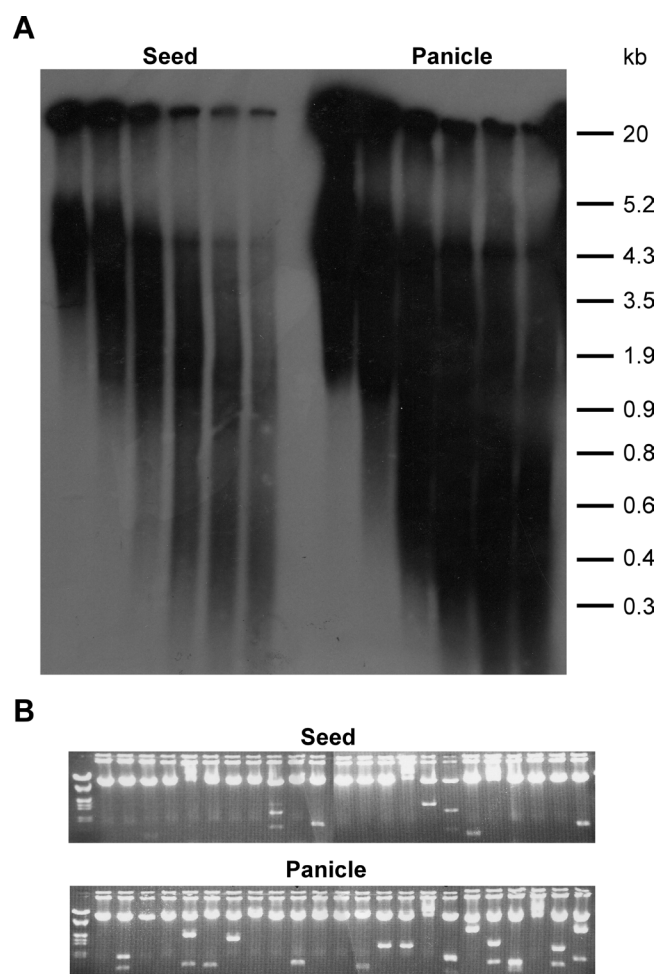
A

B

Figure 1. Construction of cDNA libraries of developing rice seeds and panicle. **A.** Results of a size fractioning of ^{32}P-dCTP-labelled cDNA. Total RNA of filling stage seed and panicle were isolated from Nipponbare rice. cDNAs were synthesized from poly(A)$^+$ RNA and separated by the Sepharose Drip Column. **B.** Efficiency of cDNA cloning and size distribution was estimated by restriction analysis of a random plasmid pool. 60% of the analyzed plasmids had an insert ranging from 300 to 5,000 bp.

followed by sequence analysis. Finally, 42 PCR products and 27 plasmids were sequenced and BLAST searched in Genbank (http://www.ncbi.nlm.nih.gov). Respectively seven and two colonies were identified encoding two zinc finger proteins among all genes which are in frame with the Gal4p activation domain (Gal4 AD) sequence in pACTII vector. Interestingly both of them belong to a large CCCH zinc finger family containing 71 members in rice (http://plntfdb.bio.uni-potsdam.de/v3.0/), we named them *OsGZF1* and *OsGZF2* after *Oryza sativa GluB-1*-binding Zinc Finger. These two clones in pACTII vector were then retransformed into strain Y187/ProGluB-1-S2 and the control strain YPO101 (Ouwerkerk and Meijer, 2001). A titration assay with a series of 3-AT concentrations was used to verify the interaction of OsGZF1 and 2 with the bait. As shown in Figure 2B, both of the transformed yeast stains were able to grow on selection medium with up to 50mM of 3-AT. But the growth of control strain with a *HIS3* gene preceded by a minimal promoter was totally suppressed by 25 mM of 3-AT. This indicates that these two CCCH proteins are able to interact with the bait sequence specifically.
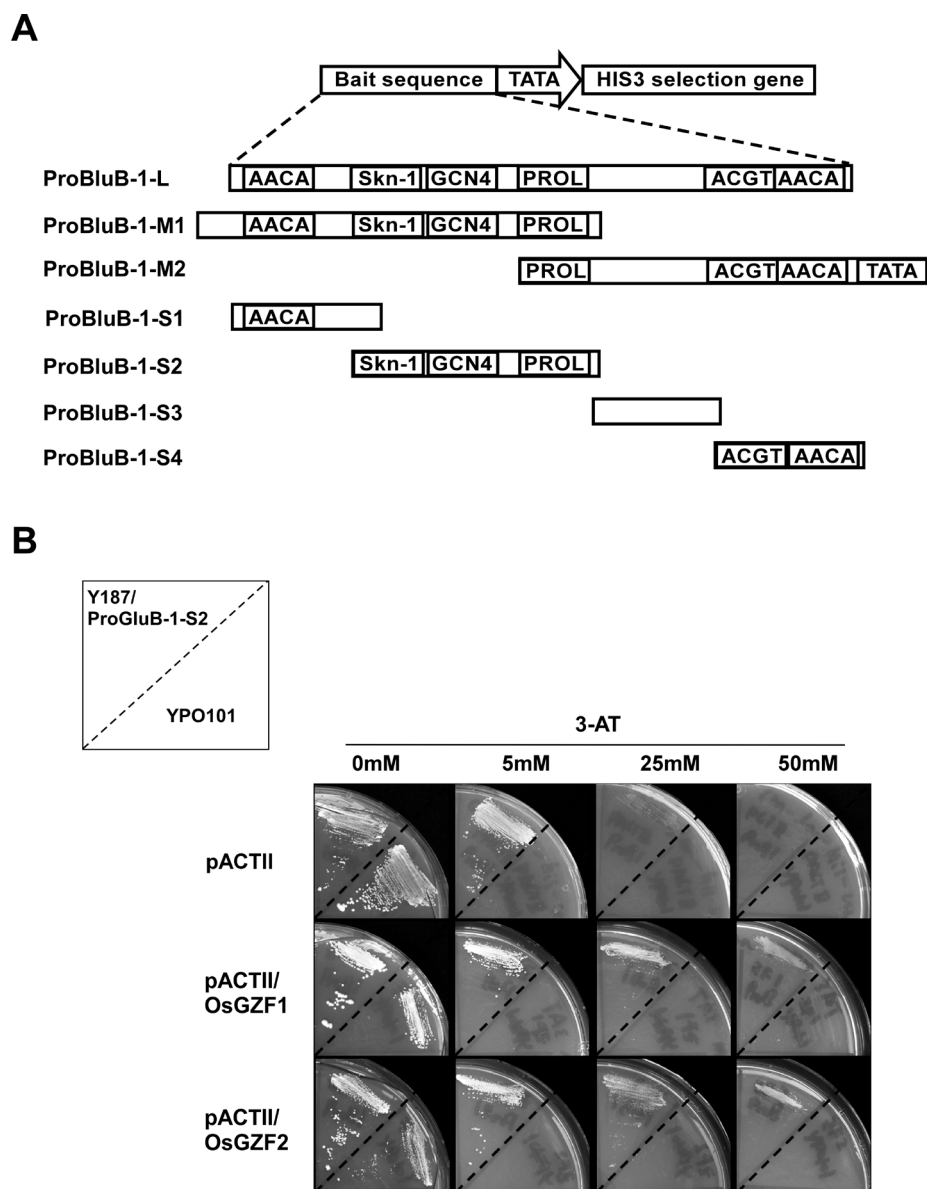
Figure 2. Yeast one-hybrid screening using different segments of *GluB-1* promoter as bait sequences. **A.** Schematic overview of the bait sequences used in the yeast screening. **B.** OsGZF1 and OsGZF2 exhibit strong interaction with bait ProGluB-1-S2. pACTII/OsGZF1 and pACTII/OsGZF2 were transformed into yeast strain Y187/ProGluB-1-S2 and YPO101 (containing the HIS3 reporter gene only with a minimal promoter but without a bait) and streaked on minimal selection medium (CM-Leu-His) with increasing concentrations of 3-AT. The empty pACTII plasmid with activation domain was used as control. Growth was monitored after seven days growth at 30°C.

Gene and protein structures of OsGZF1 and OsGZF2

Sequence comparisons of *OsGZF1* and *OsGZF2* in TIGR (http://rice.plantbiology. msu.edu/blast.shtml) showed that *OsGZF1* and *OsGZF2* correspond with loci LOC_Os07g47240 and LOC_Os11g28270 respectively. *OsGZF1* located on chromosome 7 from 28,232,263 to 28,233,649 bp which has two exons and only one intron near the C-terminus (Figure 3A). The full-length cDNA of *OsGZF1* is 1,341 bp long containing an open reading frame predicted to encode a 281 amino acid protein (Figure 3B). OsGZF1 has one CHCH motif and two CCCH motifs which are

A

OsGZF1

OsGZF2

B

```
   1 CTCACATCCTTCATCATGCATTAGGACATATATAGCCCTATATATGCACAACATTCATACCAATTAATGTTTAGTTTCTAACATGCATCTAGCCTAGCATTT
 101 TTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCAAAGCACATTGCAAACACAGGCTTAGCTTGCATATATACA
              M  A  M  H  H  M  H  D  V  F  A  Q  T  T  S  I  I  G  H  E  E  E  E  V  T  I  D  P  T
 201 TTTTTTTGTGGCCATGGCTATGCATCATATGCATGATGTTTTTGCACAGACGACGAGCATCATCGGCCATGAAGAAGAAGAAGTAACCATTGACCCGACC
           K  W  G  A  W  A  H  R  G  H  R  L  W  A  S  M  S  E  D  F  W  I  H  V  Y  K  V  Q  R  C  P  R  S
 301 AAGTGGGGCGCATGGGCTCATCGTGGCCACCGTCTCTGGGCATCCATGTCGGAGGATTTCTGGATACACGTCTACAAGGTGCAGCGGTGCCCGCGGTCGT
           S  S  H  D  W  T  S  C  P  Y  A  H  K  G  E  R  A  R  R  R  D  T  R  R  F  A  Y  A  A  V  S  C  P  D
 401 CCAGCCACGACTGGACGTCGTGCCCCTACGCGCACAAGGGCGAGCGCGCGCGGCGGCGCGACACGCGGCGCTTCGCGTACGCCGCCGTGTCGTGCCCGGA
           Y  R  P  R  E  A  A  P  G  A  V  P  S  C  A  H  G  L  R  C  R  Y  A  H  G  V  F  E  L  W  H  P
 501 CTACCGCCCGCGCGAGGCGGCGCCGGGCGCCGTGCCCAGCTGCGCCCACGGCCTCCGTTGCCGCTACGCCCATGGTGTCTTCGAGCTATGGCTGCACCCG
           S  R  F  R  T  R  M  C  S  A  G  T  R  C  P  R  R  I  C  F  F  A  H  S  A  A  E  L  R  D  D  P  N
 601 TCCCGCTTCCGCACGCGCATGTGCTCCGCCGGCACGCGGTGCCCTGCCCGGATCTGCTTCTTTGCCCACTCCGCCGCCGAGCTTAGGGACGATCCCAACT
           S  I  A  S  A  I  L  T  P  M  P  M  P  M  P  M  S  I  P  M  R  T  P  D  A  D  H  H  P  R  V  S  A  M
 701 CCATTGCGTCAGCCATCCTCACACCGATGCCAATGCCGATGCCGATGTCGATACCGATGCGGACTCCTGATGCTGATCATCATCCCAGGGTTTCTGCTAT
            R  D  Q  L  D  L  I  E  E  A  M  R  N  R  L  R  L  Y  S  N  A  N  A  N  A  N  V  G  S  S  S  A  T
 801 GCGTGATCAGCTTGATCTAATCGAAGAGGCCATGAGAAATAGGCTTCGCTTGTATTCTAATGCTAATGCTAATGCTAATGTTGGTTCTTCTTCGCTACC
            A  V  A  T  T  T  I  P  V  S  T  L  A  N  G  E  G  S  T  G  K  R  C  G  C  R  R  C  V  E  E  D
 901 GCTGTTGCTACCACCACCATACCGGTGTCAACATTGGCAAATGGTGAGGGTTCAACTGGAAAAAGGTGTGGTTGCCGCCGTTGTGTGGAGGAGGAAGACG
            A  L  L  N  G  Y  P  H  Y  D  L  I  M  D  L  V  D  E  *
1001 CTTTGCTCAATGGCTACCCACATTATGATCTCATCATGGACTTGGTGGATGAATAGATGGGTGTGTTACATACTCTTTTGCGATGCATTTGAAGGAGATG
1101 ACACAAGCTCGCAAATTGCCAAAATGGAGAAACTTGAGCGATTTGGGATGCTTTAAAAGTTGACACCCTATACATATGGTGCTGGATCATGTATCATATT
1201 TATGCCAATTTTTACTAGAAAAGTGTATGCGTGTGCATGCATATATAAAAATATTTATGATATTGTCGTCGGGCTAGCTCCGAGCCTATATATCAAGTGG
1301 TATCATTTTCTTTACAAATATTAAAAAAAAAAAAAAAAAAAA
```

C

```
   1 CTTCCTTCCTCCCTCTTCCGCCTCCGCCTCCTCCTCCTCCGCCCTTACCCCCGCGCGCGTGCACGCGACGGCGCGGCGATCTGATGCTGCCGCCGCTGCT
              M  A  A  A  P  S  A  G  G  V  G  E  G  S  S  S  S  A  A  A  A  A  A  A  A  A
 101 GCTGCAGGCGCCGCGGTGATGGCGGCGGCGCCGTCCGCCGGCGGCGCCGTTGGGGAGGGCTCTTCCTCCTCCGCCGCCGCCGCCGCTGCTGCTGCCGCGA
           T  I  G  P  H  V  V  D  E  E  A  M  W  Q  M  N  L  G  E  A  M  E  A  G  P  Y  P  E  R  I  G  E  P  D
 201 CGATCGGGCCCCACGTGGTCGACGAAGAGGCAATGTGGCAGATGAATCTAGGAGAAGCTATGGAAGCTGGGCCATACCCAGAGCGTATTGGAGAACCAGA
           C  S  Y  Y  M  R  T  G  L  C  R  F  G  M  T  C  K  F  N  H  P  A  D  R  K  M  A  V  A  A  A  R  M
 301 TTGTAGTTATTACATGAGGACTGGCTTGTGCAGGTTTGGGATGACCTGTAAATTTAATCACCCAGCAGATCGTAAGATGGCTGTTGCTGCTGCAAGGATG
           K  G  E  Y  P  Q  R  I  G  Q  P  E  C  Q  Y  Y  L  K  T  G  T  C  K  F  G  A  T  C  K  F  H  H  P
 401 AAGGGGGAATATCCTCAAAGAATTGGTCAACCTGAATGTCAATATTATTTGAAGACTGGAACATGCAAATTTGGAGCAACATGCAAGTTTCACCACCCCC
           R  E  K  A  A  I  A  T  R  V  Q  L  N  A  L  G  Y  P  L  R  P  N  E  K  E  C  A  Y  Y  L  R  T  G  Q
 501 GAGAAAAAGCTGCAATTGCAACCCGAGTACAGCTGAATGCTTTAGGCTACCCATTGCGGCCGAATGAGAAGGAATGCGCTTATTATTTAAGAACCGGACA
           C  K  F  G  S  T  C  K  F  H  H  P  Q  P  S  N  T  M  V  A  V  R  G  S  V  Y  S  P  G  Q  S  V  T
 601 GTGCAAATTTGGAGCACATGTAAGTTTCATCATCCACAGCCATCTAATACGATGGTTGCTGTACGTGGCTCTGTTTATTCACCTGGACAGTCAGTTACT
           S  P  S  Q  H  T  Y  P  G  A  V  T  N  W  P  L  S  R  S  A  S  F  I  A  S  P  R  W  P  G  H  S  S
 701 TCTCCTAGTCAGCATACTTACCCAGGGGCTGTAACAAACTGGCCCTTGTCAAGATCTGCTTCGTTTATTGCAAGTCCAAGGTGGCCAGGCCATTCTAGCT
           Y  A  Q  V  I  V  P  G  L  V  Q  V  P  G  W  N  P  Y  A  A  Q  I  G  S  S  S  S  D  D  Q  Q  R  T
 801 ATGCACAAGTGATTGTTCCTCCAGGGCTCGTTCAAGTTCCAGGGTGGAATCCTTATGCAGCACAAATTGGTTCTTCATCCTCAGACGATCAACAGAGGAC
           A  G  G  A  Q  Y  Y  T  G  S  R  H  S  E  T  P  N  M  G  D  Q  G  M  F  S  S  Y  Q  A  G  S  V  P
 901 AGCTGGAGGTGCACAATACTATACTGGCTCACGTCACAGTGAAACACCTAACATGGGTGACCAGGGAATGTTTTCATCATACCAAGCTGGTTCTGTTCCT
           L  G  L  Y  T  V  Q  R  E  S  I  F  P  E  R  P  D  Q  P  E  C  Q  F  Y  M  K  T  G  D  C  K  F  G
1001 CTTGGGCTATATACAGTACAGAGGGAGAGCATATTTCCAGAGCGACCTGACCAACCAGAATGCCAGTTCTATATGAAGACTGGGGACTGTAAGTTTGGTG
           A  V  C  K  F  H  H  P  K  E  R  I  I  P  T  P  N  C  A  L  S  S  L  G  L  P  L  R  P  G  E  P  I  C
1101 CTGTATGCAAGTTCCATCATCCCAAGGAGAGAATTATCCCTACTCCAAACTGTGCGTTGAGCTCATTAGGTCTTCCACTGCGTCCGGGAGAGCCCATATG
            T  F  Y  S  R  Y  G  I  C  K  F  G  P  N  C  K  F  D  H  P  M  G  T  V  M  Y  G  L  A  T  S  P  T
1201 CACATTCTATTCTCGCTATGGCATCTGCAAGTTTGGTCCGAATTGCAAATTTGATCATCCAATGGGAACTGTGATGTATGGTCTCGCAACATCACCAACT
            G  D  V  S  A  R  R  M  L  A  P  V  P  A  H  S  E  V  S  P  D  N  V  S  G  R  S  R  R  I  T  H  S
1301 GGTGATGTGTCTGCCCGACGTATGCTGGCACCTGTACCAGCACATTCAGAAGTATCGCCTGACAATGTCTCAGGGAGGTCTCGGAGGATCACCCATTCAG
            D  S  Q  Q  I  P  S  G  E  R  G  T  E  R  E  A  S
1401 ATTCCCAGCAAATACCCTCTGGTGAAAGAGGCACCGAGAGAGAGGCGTCCTAAGCCATCCTCGGTGCAAATGTTTCATGCTGCTTCCTGTTCTTGACTCT
1501 ICTGAGTAAATTACCCGCAACAAAACTTGAGGAGCCGAAGCTGTACAGGTCTCGCCCTTTTGCCTCACATTTGAGCTTTACATACCGGAGCTTGCTTCAAG
1601 ACTTCCAACATGTATCGCATTCAGGCAGGAAAAAGTTCTTGCTTGGTTAGGGATGCTCTTTTGCCCCGTGTAAACGGGACAACCGACCTTATTTAAACATC
1701 TATATATGGAACAAACTCACCTCTGGACATTCGATTTCGAGTGATGTGCCTTGATCCCTTTGTTGTAGTTATACAGGTGCCTTGCAATGGACAAAGCTTT
1801 TGCCAAACGGGGGATATGTGCAAATCGCTCTCCTCTCTGTTTCTTTCTTTTTTTTGCCTTCCCTCTAATTTAGCAGTTTCATAAGGGGCAACCCCCGCTT
1901 CTGGAATTGTGCATGTTGAGGATGCGGAAATAATAGGCAATTGTATAATCATCTACATATTTATACCCTGTCACTTGTTCCCATAAGAGAGCAGAAAATT
2001 CACCTGAAAAAAAAAAAAAAAAAAAAACTCGGGATCTGTGGAC
```

Figure 3. Gene and protein structures of OsGZF1 and OsGZF2. **A.** Schematic representation of exon and intron structure of *OsGZF1* and *OsGZF2*. **B** and **C**. cDNA and aligned protein sequence of OsGZF1 and OsGZF2. The zinc finger motifs are underlined. The Cys and His residues within the zinc finger motifs are marked with squares.

closely located together with 19 amino acid (AA) between the CHCH motif and the first CCCH motif and 16 AA between the two CCCH motifs (Figure 3B). OsGZF2 has seven exons and six introns spread over a region of over 9,000 bp long on chromosome 11 from 15,784,331 to 15,793,852 (Figure 3A). Comparing the *OsGZF2* cDNA sequence to the corresponding cDNA clone with Genbank Accession AK110345 of the KOME collection (http://cdna01.dna.affrc.go.jp/cDNA/) showed that there are two main differences altering the predicted protein sequence. First is that the coding region of AK1130345 starts at +315 bp instead of +136 bp as the predicted CDS of LOC_Os11g28270 because of 65 extra nucleotides located at +245 bp of AK110345 (Figure 4). This change results in alteration of the longest predicted open reading frame (ORF) which made the predicted protein of AK110345 41 AA shorter than LOC_Os11g28270 at the N-terminus. The second difference is that an extra A was found to be present at +1,108 bp which shifts the coding frame of the C-terminus of LOC_Os11g28270 (Figure 4). The predicted protein of LOC_Os11g28270 was 127 AA longer than AK110345. Therefore, based on the sequence comparison of *OsGZF2*, LOC_Os11g28270 and AK110345, we suggest that the cDNA length of *OsGZF2* is 2,039 bp encoding for a 407 AA protein (Figure 3C). OsGZF2 has five tandem located CCCH motifs, three of which are at the N-terminus with a 25 AA gap between the first two CCCH motifs and 26 AA between the second and the third CCCH motif, the other two CCCH motifs are adjacent to the C-terminus and also separated by a 26 AA gap (Figure 3C).

Since OsGZF1 and OsGZF2 have two and five CCCH motifs respectively, we compared the amino acid sequences of those motifs. Yet there is no similarity of the CCCH motifs between OsGZF1 and OsGZF2. Nonetheless, five CCCH fingers of OsGZF2 are highly similar as C-X_2-Y-X_3-G-X-C-X-F-G-X_2-C-K-F-X-H with nine conserved AAs, where X represents the unconserved AAs (Figure 5). The middle three motifs of the five even have 13 conserved AAs (C-X_2-Y-X_2-T-G-X-C-K-F-G-X_2-C-K-F-H-H). We also compared the sequence of OsGZF2 with its rice and Arabidopsis homologous proteins and they also show high similarity within the CCCH motifs (Figure 6).

Similar binding patterns of OsGZF1 and OsGZF2 to the *GluB-1* promoter *in vitro*
In order to verify the binding abilities of OsGZF1 and OsGZF2 to the core promoter of *GluB-1*, the four short baits from the yeast one-hybrid screens ProGluB-1-S1, 2, 3, 4 (S1a, S2a, S3a, S4a, Figure 7A) were used in Electrophoretic Mobility Shift Assays (EMSA) in combination with GST-tagged recombinant proteins of OsGZF1 and OsGZF2 that were overexpressed and purified from *E. coli*. As shown in Figure 7B, OsGZF1 and OsGZF2 have similar binding patterns. DNA-protein complexes were observed for all fragments tested with OsGZF1 and OsGZF2 proteins, thereby confirming the results of the yeast one-hybrid screens *in vitro*. Other authors (Takaiwa et al., 1996; Yoshihara et al., 1996; Wu et al., 1998b; Washida et al., 1999; Wu et al., 2000) showed the importance of the so-called AACA, GCN4, PROL and ACGT motifs in regulation of the *GluB-1* promoter. In order to validate if these boxes are also recognized by OsGZF1 and OsGZF2, single or double mutations were made in oligos ProGluB-1-S1, S2 and S4 as shown in Figure 7A. DNA-protein complexes were observed with all tested oligos, indicating the mutated motif was not essential for binding. Creation of mutations in the binding motif of ProGluB-1-S1 decreased binding. However, mutations in ProGluB-1-S2 and ProGluB-1-S4 increased the binding especially for ProGluB-1-S2 with double mutation in both GCN4 and PROL motives (Figure 7C and D).

```
             1                                                                                                  100
A            ..................................................................................................
B            .................CTTCCTTCCTCCCTCTTCCGCCTCCGCCTCCTCCTCCTCCGCCCTTACCCCCGCGCGCGTGCACGCGACGGCGCGGCGATCTG
C            TCGTTCCTCTCATTCCCCTTCCTTCCTCCCTCTTCCGCCTCCGCCTCCTCCTCCTCCGCCCTTACCCCCGCGCGCGTGCACGCGACGGCGCGGCGATCTG
             101                                                                 ▼                              200
A            ...................................ATGGCGGCGGCGCCGTCCGCCGGCGGCGTTGGGGAGGGGCTCTTCCTCCTCCGCCGCCGCCGCCGC
B            ATGCTGCCGCCGCTGCTGCTGCAGGCGCCGCGGTGATGGCGGCGGCGCCGTCCGCCGGCGGCGTTGGGGAGGGGCTCTTCCTCCTCCGCCGCCGCCGCCGC
C            ATGCTGCCGCCGCTGCTGCTGCAGGCGCCGCGGTGATGGCGGCGGCGCCGTCCGCCGGCGGCGTTGGGGAGGGGCTCTTCCTCCTCCGCCGCCGCCGCCGC
             201                                                                                                300
A            TGCTGCTGCTGCCGCGACGATCGGGCCCCACGTGGTCGACGAAG.......................................................
B            TGCTGCTGCTGCCGCGACGATCGGGCCCCACGTGGTCGACGAAG.......................................................
C            TGCTGCTGCTGCCGCGACGATCGGGCCCCACGTGGTCGACGAAGGAGCGCTCTGTTTTGGTCATTCTCACAATGCTGATATGTTTTATAATGTGGATACG
             301                             ▼                                                                  400
A            ........AGGCAATGTGGCAGATGAATCTAGGAGAAGCTATGGAAGCTGGGCCATACCCAGAGCGTATTGGAGAACCAGATTGTAGTTATTACATGAG
B            ........AGGCAATGTGGCAGATGAATCTAGGAGAAGCTATGGAAGCTGGGCCATACCCAGAGCGTATTGGAGAACCAGATTGTAGTTATTACATGAG
C            TTGTCGGTGAGGCAATGTGGCAGATGAATCTAGGAGAAGCTATGGAAGCTGGGCCATACCCAGAGCGTATTGGAGAACCAGATTGTAGTTATTACATGAG
             401                                                                                                500
A            GACTGGCTTGTGCAGGTTTGGGATGACCTGTAAATTTAATCACCCAGCAGATCGTAAGATGGCTGTTGCTGCTGCAAGGATGAAGGGGGAATATCCTCAA
B            GACTGGCTTGTGCAGGTTTGGGATGACCTGTAAATTTAATCACCCAGCAGATCGTAAGATGGCTGTTGCTGCTGCAAGGATGAAGGGGGAATATCCTCAA
C            GACTGGCTTGTGCAGGTTTGGGATGACCTGTAAATTTAATCACCCAGCAGATCGTAAGATGGCTGTTGCTGCTGCAAGGATGAAGGGGGAATATCCTCAA
             501                                                                                                600
A            AGAATTGGTCAACCTGAATGTCAATATTATTTGAAGACTGGAACATGCAAATTTGGAGCAACATGCAAGTTTCACCACCCCCGAGAAAAAGCTGCAATTG
B            AGAATTGGTCAACCTGAATGTCAATATTATTTGAAGACTGGAACATGCAAATTTGGAGCAACATGCAAGTTTCACCACCCCCGAGAAAAAGCTGCAATTG
C            AGAATTGGTCAACCTGAATGTCAATATTATTTGAAGACTGGAACATGCAAATTTGGAGCAACATGCAAGTTTCACCACCCCCGAGAAAAAGCTGCAATTG
             601                                                                                                700
A            CAACCCGAGTACAGCTGAATGCTTTAGGCTACCCATTGCGGCCGAATGAGAAGGAATGCGCTTATTATTTAAGAACCGGACAGTGCAAATTTGGGAGCAC
B            CAACCCGAGTACAGCTGAATGCTTTAGGCTACCCATTGCGGCCGAATGAGAAGGAATGCGCTTATTATTTAAGAACCGGACAGTGCAAATTTGGGAGCAC
C            CAACCCGAGTACAGCTGAATGCTTTAGGCTACCCATTGCGGCCGAATGAGAAGGAATGCGCTTATTATTTAAGAACCGGACAGTGCAAATTTGGGAGCAC
             701                                                                                                800
A            ATGTAAGTTTCATCATCCACAGCCATCTAATACGATGGTTGCTGTACGTGGCTCTGTTTATTCACCTGGACAGTCAGTTACTTCTCCTAGTCAGCATACT
B            ATGTAAGTTTCATCATCCACAGCCATCTAATACGATGGTTGCTGTACGTGGCTCTGTTTATTCACCTGGACAGTCAGTTACTTCTCCTAGTCAGCATACT
C            ATGTAAGTTTCATCATCCACAGCCATCTAATACGATGGTTGCTGTACGTGGCTCTGTTTATTCACCTGGACAGTCAGTTACTTCTCCTAGTCAGCATACT
             801                                                                                                900
A            TACCCAGGGGCTGTAACAAACTGGCCCTTGTCAAGATCTGCTTCGTTTATTGCAAGTCCAAGGTGGCCAGGCCATTCTAGCTATGCACAAGTGATTGTTC
B            TACCCAGGGGCTGTAACAAACTGGCCCTTGTCAAGATCTGCTTCGTTTATTGCAAGTCCAAGGTGGCCAGGCCATTCTAGCTATGCACAAGTGATTGTTC
C            TACCCAGGGGCTGTAACAAACTGGCCCTTGTCAAGATCTGCTTCGTTTATTGCAAGTCCAAGGTGGCCAGGCCATTCTAGCTATGCACAAGTGATTGTTC
             901                                                                                                1000
A            CTCCAGGGCTCGTTCAAGTTCCAGGGTGGAATCCTTATGCAGCACAAATTGGTTCTTCATCCTCAGACGATCAACAGAGGACAGCTGGAGGTGCACAATA
B            CTCCAGGGCTCGTTCAAGTTCCAGGGTGGAATCCTTATGCAGCACAAATTGGTTCTTCATCCTCAGACGATCAACAGAGGACAGCTGGAGGTGCACAATA
C            CTCCAGGGCTCGTTCAAGTTCCAGGGTGGAATCCTTATGCAGCACAAATTGGTTCTTCATCCTCAGACGATCAACAGAGGACAGCTGGAGGTGCACAATA
             1001                                                                                               1100
A            CTATACTGGCTCACGTCACAGTGAAACACCTAACATGGGTGACCAGGGAATGTTTTCATCATACCAAGCTGGTTCTGTTCCTCTTGGGCTATATACAGTA
B            CTATACTGGCTCACGTCACAGTGAAACACCTAACATGGGTGACCAGGGAATGTTTTCATCATACCAAGCTGGTTCTGTTCCTCTTGGGCTATATACAGTA
C            CTATACTGGCTCACGTCACAGTGAAACACCTAACATGGGTGCCAGGGAATGTTTTCATCATACCAAGCTGGTTCTGTTCCTCTTGGGCTATATACAGTA
             1101           ▼                                                                                   1200
A            CAGAGGGAGAGCATATTTCCAGAGCGACCTGACCAACCAGAATGCCAGTTCTATATGAAGACTGGGGACTGTAAGTTTGGTGCTGTATGCAAGTTCCATC
B            CAGAGGGAGAGCATATTTCCAGAGCGACCTGACCAACCAGAATGCCAGTTCTATATGAAGACTGGGGACTGTAAGTTTGGTGCTGTATGCAAGTTCCATC
C            CAGAGGG.GAGCATATTTCCAGAGCGACCTGACCAACCAGAATGCCAGTTCTATATGAAGACTGGGGACTGTAAGTTTGGTGCTGTATGCAAGTTCCATC
             1201                                                                                               1300
A            ATCCCAAGGAGAGAATTATCCCTACTCCAAACTGTGCCGTTGAGCTCATTAGGTCTTCCACTGCGTCCGGGAGAGCCCATATGCACATTCTATTCTCGCTA
B            ATCCCAAGGAGAGAATTATCCCTACTCCAAACTGTGCCGTTGAGCTCATTAGGTCTTCCACTGCGTCCGGGAGAGCCCATATGCACATTCTATTCTCGCTA
C            ATCCCAAGGAGAGAATTATCCCTACTCCAAACTGTGCCGTTGAGCTCATTAGGTCTTCCACTGCGTCCGGGAGAGCCCATATGCACATTCTATTCTCGCTA
             1301                                                                                               1400
A            TGGCATCTGCAAGTTTGGTCCGAATTGCAAATTTGATCATCCAATGGGAACTGTGATGTATGGTCTCGCAACATCACCAACTGGTGATGTGTCTGCCCGA
B            TGGCATCTGCAAGTTTGGTCCGAATTGCAAATTTGATCATCCAATGGGAACTGTGATGTATGGTCTCGCAACATCACCAACTGGTGATGTGTCTGCCCGA
C            TGGCATCTGCAAGTTTGGTCCGAATTGCAAATTTGATCATCCAATGGGAACTGTGATGTATGGTCTCGCAACATCACCAACTGGTGATGTGTCTGCCCGA
             1401                                                                                               1500
A            CGTATGCTGGCACCTGTACCAGCACATTCAGAAGTATCGCCTGACAATGTCTCAGGGAGGTCTCGGAGGATCACCCATTCAGATTCCCAGCAAATACCCT
B            CGTATGCTGGCACCTGTACCAGCACATTCAGAAGTATCGCCTGACAATGTCTCAGGGAGGTCTCGGAGGATCACCCATTCAGATTCCCAGCAAATACCCT
C            CGTATGCTGGCACCTGTACCAGCACATTCAGAAGTATCGCCTGACAATGTCTCAGGGAGGTCTCGGAGGATCACCCATTCAGATTCCCAGCAAATACCCT
             1501                                                                                               1600
A            CTGGTGAAAGAGGCACCGAGAGAGAGGCGTCCTAA..............................................................
B            CTGGTGAAAGAGGCACCGAGAGAGAGGCGTCCTAAGCCATCCTCGGTGCAAATGTTTCATGCTGCTTCCTGTTCTTGACTCTCTGAGTAAATTACCCGCA
C            CTGGTGAAAGAGGCACCGAGAGAGAGGCGTCCTAAGCCATCCTCGGTGCAAATGTTTCATGCTGCTTCCTGTTCTTGACTCTCTGAGTAAATTACCCGCA
```

```
     1601                                                                                 1700
A    ....................................................................................................
B    ACAAAACTTGAGGAGCCGAAGCTGTACAGGTCTCGCCCTTTTGCCTCACATTTGAGCTTTACATACCGGAGCTTGCTTCAAGACTTCCAACATGTATCGC
C    ACAAAACTTGAGGAGCCGAAGCTGTACAGGTCTCGCCCTTTTGCCTCACATTTGAGCTTTACATACCGGAGCTTGCTTCAAGACTTCCAACATGTATCGC
     1701                                                                                 1800
A    ....................................................................................................
B    ATTCAGGCAGGAAAAAGTTCTTGCTTGTTAGGGATGCTCTTTTGCCCCGTGTAAACGGGACAACCGACCTTATTTAAACATCTATATATGGAACAAACTC
C    ATTCAGGCAGGAAAAAGTTCTTGCTTGTTAGGGATGCTCTTTTGCCCCGTGTAAACGGGACAACCGACCTTATTTAAACATCTATATATGGAACAAACTC
     1801                                                                                 1900
A    ....................................................................................................
B    ACCTCTGGACATTCGATTTCGAGTGATGTGCCTTGATCCCTTTGTTGTAGTTATACAGGTGCCTTGCAATGGACAAAGCTTTTGCCAAACGGGGGATATG
C    ACCTCTGGACATTCGATTTCGAGTGATGTGCCTTGATCCCTTTGTTGTAGTTATACAGGTGCCTTGCAATGGACAAAGCTTTTGCCAAACGGGGGATATG
     1901                                                                                 2000
A    ....................................................................................................
B    TGCAAATCGCTCTCCTCTCTGTTTCTTTCTTTTTTTTGCCTTCCCTCTAATTTAGCAGTTTCATAAGGGGCAACCCCCGCTTCTGGAATTGTGCATGTTG
C    TGCAAATCGCTCTCCTCTCTGTTTCTTTCTTTTTTTTGCCTTCCCTCTAATTTAGCAGTTTCATAAGGGGCAACCCCCGCTTCTGGAATTGTGCATGTTG
     2001                                                                                 2100
A    ....................................................................................................
B    AGGATGCGGAAATAATAGGCAATTGTATAATCATCTACATATTTATACCCTGTCACTTGTTCCCATAAGAGAGCAGAAAATTCACCTGAAAAAAAAAAAA
C    AGGATGCGGAAATAATAGGCAATTGTATAATCATCTACATATTTATACCCTGTCACTTGTTCCCATAAGAGAGCAGAAAATTCACCTTGATGGTGCCGGC
     2101                                                                                 2200
A    ....................................................................................................
B    AAAAACTCGGGATCTGTGGAC.......................................................................
C    TTTAGTTTGTCATTTCATCTTGCAAAATTTTGTTCTGCTACTGAACCTTTTATCAATTGTGCATCATTTTTTTGCCATTATATGTGCACGGGAAGGATGG
     2201                                                                                 2300
A    ....................................................................................................
B    ....................................................................................................
C    TTACTAAAGAACAAGTAACAAGAGGTAGGGTTAGACTGGTTGAGACAGGTCCGGTTCCTAATTCTGTTACAGCAATCTATCATTTGCACAGCTATCAAGG
     2301                                                                                 2400
A    ....................................................................................................
B    ....................................................................................................
C    TTGCCAAACACCTCTATCCGGTGTTACGCCTCCTAAATCCTGATGAACGGAGCAGCTGCGTCAACACCTATGTCCTGATAAACGGAGCAGCTGCATCAGG
     2401                                                                                 2500
A    ....................................................................................................
B    ....................................................................................................
C    ATTCTGGAGTGTCAGGTAGGTGAAGGTATAATGTGGAGCAAAAGGAGTTCCAGATGACTCCTAATCCTATAATCATCCGATACGAGTAGGTTCGCCACAT
     2501                                                                                 2600
A    ....................................................................................................
B    ....................................................................................................
C    GCGGATGAAATGTAGTGGATGGTTTCTAAACTACAAATGTAGTGGATGGTTTCTTATAAATAGGATTATCAGATCATCAGCCCTTCAAAACTATGGGCCT
     2601                                                                                 2700
A    ....................................................................................................
B    ....................................................................................................
C    GGTTGGCACAGCTCTAGCTCCCGCTCCACCCCTCCTGGAGCTGGAGCTCAGCCAAACAGTTTTAGCTCCACCAAAAATGGGAGTGGAGTTGGGTAGAGCT
     2701                                                                                 2800
A    ....................................................................................................
B    ....................................................................................................
C    CTCTCACAGAATGTACTGGAGTTGTAGAGCTGGGTTTAGGCAGCTCCACAACTCCACTTCAGACTCAACTCCTAAAGTTATATTTAGGAGTTAGAGTTGT
     2801                                                               2874
A    ...........................................................................
B    ...........................................................................
C    ACTAAACAGGCCTTACAAATAGAACACTACTTATATGTCCACATTATAAGTTTGCCTAAGAGGGTATCATTTTC
```

Figure 4. Comparison of coding regions of *LOC_Os11g28270* (A), *OsGZF2* (B) and *AK110345* (C). Putative transcription start sites and the extra A in *OsGZF2* and *LOC_Os11g28270* are indicated with arrow heads.

OsGZF1 and OsGZF2 are localized in nucleus

To validate the putative function of OsGZF1 and OsGZF2 as transcription factors, subcellular localization of the proteins was analyzed in rice protoplasts using a transient expression system.

```
C3H-1    IGEPDCSYYMRTGLCRFGMTCKFNHPADRK
C3H-2    IGQPECQYYLKTGTCKFGATCKFHHPREKA
C3H-3    PNEKECAYYLRTGQCKFGSTCKFHHPQPSN
C3H-4    PDQPECQFYMKTGDCKFGAVCKFHHPKERI
C3H-5    PGEPICTFYSRYGICKFGPNCKFDHPMGTV
              ▲          ▲      ▲     ▲
```

Figure 5. Alignment and comparison of the five CCCH motifs of OsGZF2.

Figure 6. Sequence comparison of proteins from Arabidopsis and rice homologous with OsGZF2. 0
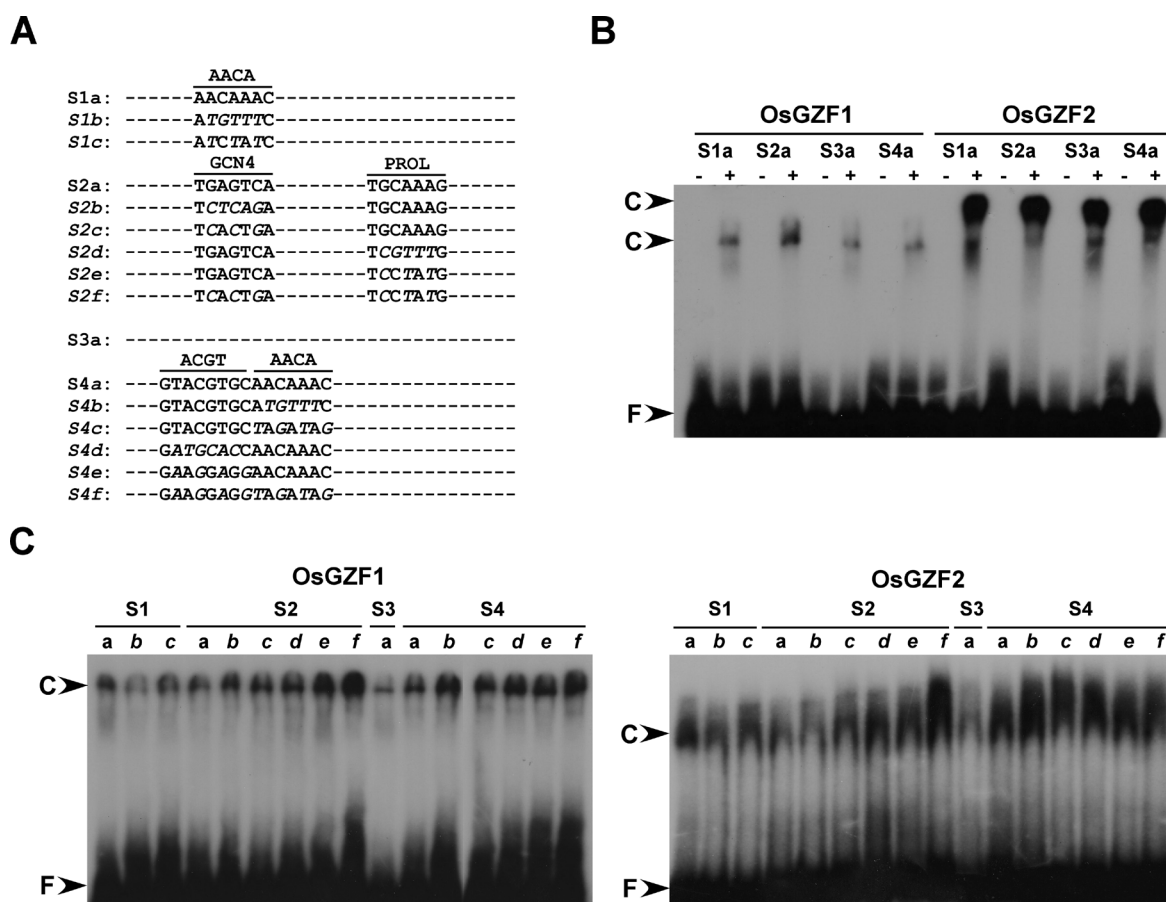
Figure 7. Interaction of OsGZF1 and OsGZF2 with *GluB-1* promoter *in vitro*. **A.** Oligo sequences used in the EMSA. S1a, S2a, S3a and S4a represent the four bait sequences ProGluB-1-S1, 2, 3, 4 used in the yeast one-hybrid assays. The other fragments were generated from the four short baits but with single or double mutations in the so called *cis*-elements displayed in Figure 1A. Single mutations were designed such that they converted the motifs of parts into their complementary sequences (*S1b, S2b, S2d, S4b, S4d*) or every other nucleotide (*S1c, S2c, S2e, S4c, S4e*). Double mutations were created by converting both motifs into their complementary sequences every other nucleotide (*S2f* and *S4f*). **B.** EMSAs were performed with recombinant GST-tagged OsGZF1 and OsGZF2 proteins and radio-labeled wild-type oligos. – and + represent without and with the proteins added to the reaction. **C.** Interaction of OsGZF1 and OsGZF2 to oligos without and with mutations. The arrow heads mark the positions of protein-DNA complexes (C) and free probes (F).

Initially we used two constructs by fusing the ORF of *OsGZF1* and *OsGZF2* in frame to the C-terminus of the Green Fluorescent Protein (GFP) reporter. As shown in Figure 8A to C, the OsGZF1-GFP signal was observed predominantly in the nucleus but no signal was observed for OsGZF2 (data not shown). However, when *OsGZF2* was fused to the N-terminus of GFP, a nuclear signal was observed too (Figure 8D to F). GFP signal of 35S::GFP control vector transformed protoplasts was visualized both in nucleus and cytoplasm (Figure 8G to I).

Overlapping expression patterns of *OsGZF1* and *OsGZF2*

Northern blot analysis was employed to assess the expression pattern of *OsGZF1* and *OsGZF2* on RNA level. RNA analyzed was isolated from different tissues of different stages. Both genes were
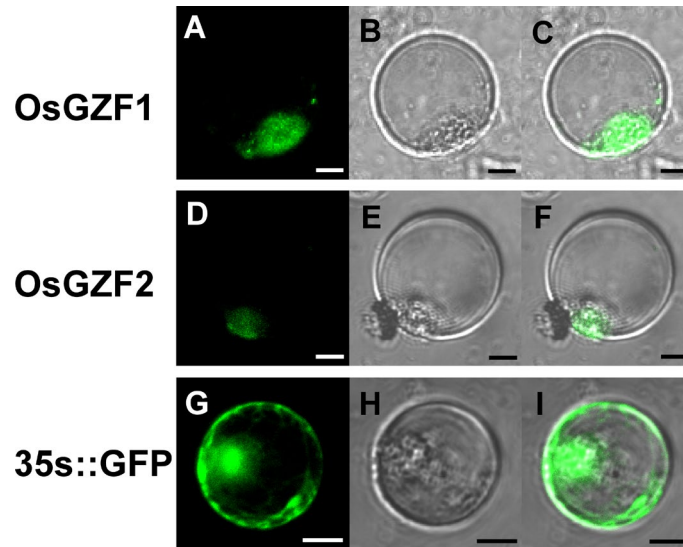
Figure 8. Nuclear localization experiments. Coding regions *OsGZF1* and *OsGZF2* were tagged to a *GFP* reporter gene and transformed into rice seedling protoplasts. OsGZF1/GFP and OsGZF2/GFP are nuclear localized**.** After overnight culture, cells were observed under fluorescence (**A**, **D** and **G**) and bright field (**B**, **F** and **H**). Merged picture of **A** and **B**, **D** and **E**, **G** and **H** are shown in **C**, **F** and **I** respectively. Bar = 10 μm
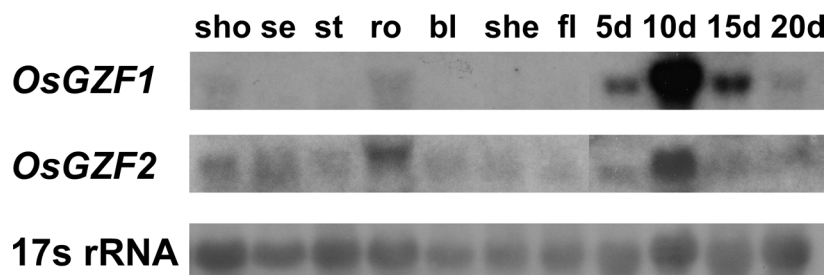
Figure 9. Northern blot analysis of *OsGZF1, OsGZF2* and *GluB-1* in different tissues. Total RNAs were isolated from seminal shoot (sho), seedling (se), stem (st), root (ro), leaf blade (bl), sheath (she), flower (fl), seed of 5 (5d), 10 (10d), 15 (15d) and 20 (20d) DAF. *OsGZF1* and *OsGZF2* and *GluB-1* probes were hybridized on the same membrane one by one.

found to be highly expressed in developing seeds. As shown in Figure 9, mRNA of *OsGZF1* and *OsGZF2* started to accumulate at DAF 5 (days after flowering) and peaked at 10 DAF. In grains of 15 DAF, expression of *OsGZF2* was already absent and the expression of *OsGZF1* decreased between 15 DAF and 20 DAF. Besides in developing grains, *OsGZF1* and *OsGZF2* were also found to be expressed lowly in seminal roots and shoots. We also observed low level of *OsGZF2* expression in more vegetative tissues than *OsGZF1*. According to previous reports, *GluB-1* shows no detectable expression at 3 DAF but start to accumulate at 6 DAF with a peak activity at 10 (Duan and Sun, 2005) to 15 (Onodera et al., 2001) DAF before declining towards maturation (20 DAF). Thus, the expression pattern of *OsGZF1* and *OsGZF2* mostly overlapped with *GluB-1* but the peaking time is slightly earlier than *GluB-1*.

In order to study the temporal and spatial expression pattern of the genes in more detail, 2 kb upstream promoter regions of *OsGZF1* and *OsGZF2* were amplified from wild type Nipponbare and cloned in front of the *GUS* reporter gene resulting in constructs ProOsGZF1::GUS and
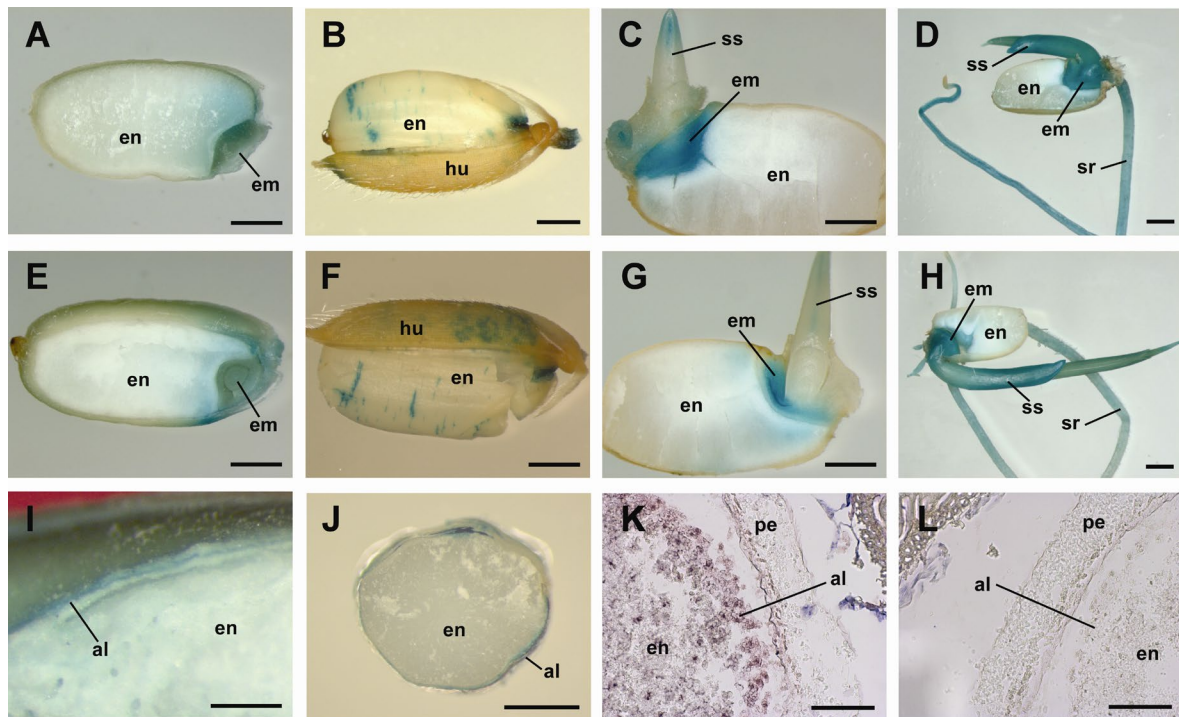
Figure 10. Analysis of the temporal and spatial expression pattern of *OsGZF1* and *OsGZF2* in different developing seed stages using ProOsGZF1::GUS and ProOsGZF2::GUS transgenics and RNA *in situ* hybridization. Histochemical staining of ProOsGZF1::GUS (**A** to **D**) and ProOsGZF2::GUS (**E** to **J**) plants. Seeds were cut longitudinally (**A**, **C**-**E**, **G**-**I**) or cross sectioned (**J**) and incubated in X-Gluc solution at 37°C overnight. **A** and **E**, developing seeds of 10 DAF. **B** and **F**, mature seeds. **C** and **G**, seeds at 1 day after germination (DAG). **D** and **H**, seeds at 3 DAG. **I** close-up of developing seed in **E**. **J** mature seed. **K** and **L** *In situ* hybridization of *OsGZF2* transcripts at the longitudinal sections of a seed from Nipponbare rice at 7 DAF. **K** hybridization with an *OsGZF2* antisense probe. **L** hybridization with an *OsGZF2* sense probe. al: aleurone layer. em: embryo. en: endosperm. hu: husk. pe: pericarp. sr: seminal root. ss: seminal shoot. Bar (A to F, H, J) = 1,000 μm; bar (I) = 200 μm; bar (K, L) = 50 μm.

ProOsGZF2::GUS. The constructs were introduced into rice varieties Nipponbare and Zhonghua 11 respectively. Five independent transgenic lines of each construct were histochemically analyzed for GUS activity which was shown to be predominant in seeds thereby confirming the data obtained with northern blot. GUS activity was detected mainly in a small endosperm layer around the scutellum region of the embryo during grain filling stages (five to ten DAF) for both ProOsGZF1::GUS and ProOsGZF2::GUS constructs (Figure 10A and E). A GUS signal was also observed in the aleurone of ProOsGZF2::GUS seeds of five to ten DAF (Figure 10I and J) and on the surface of mature seeds (Figure 10B and F), germinating embryo's, seminal shoots and roots (Figure 10C, D, G and H) for both genes, indicating their diverse function in these tissues. To further support results obtained with the promoter GUS results, we performed RNA *in situ* hybridization for *OsGZF2* on seven DAF old Nipponbare seeds. A strong signal was observed in the aleurone layer of the *OsGZF2* antisense probe hybridized seeds (Figure 10K), but not on the sense probe hybridized ones (Figure 10L). This result is consistent with the ProOsGZF2::GUS histochemical staining. For *OsGZF1* we were not able to set-up reliable *in situ* hybridizations.

OsGZF1 and OsGZF2 down-regulate reporter gene expression and reduce activation by RISBZ1 of *GluB-1* promoter activity

In order to elucidate the potential regulatory function of *OsGZF1* and *OsGZF2* in gene transcription, further studies were performed in yeast and a rice protoplast system. For the yeast experiments, coding regions of *OsGZF1* and *OsGZF2* were cloned in frame with the GAL4 binding domain (BD) in vector pAS2-1 (Clontech) and transformed into yeast strain PJ69-4A containing a *HIS3* and an *ADE2* reporter gene each preceded by a Gal4p binding sequence. The transformants were grown on selection medium (CM-His and 2mM 3-AT) for a week and none showed apparent differences compared to negative controls (empty pAS2-1, data not shown). All yeast transformants could grow weakly on the plates lacking histidine including the control strain which was caused by the leaky *HIS3* expression. However, application of 2mM 3-AT could completely suppress the growth and similar results were obtained when the yeast transformants were grown on plates lacking adenine to check for ADE2 activation which is a more stringent condition. In agreement with the results obtained from the histidine selection plates, no growth was found indicating that in yeast OsGZF1 and OsGZF2 show no activation properties.

Next, transient assays were employed to further determine the possible roles of OsGZF1 and OsGZF2 in regulation of the *GluB-1* promoter. For this, cDNAs of *OsGZF1* and *OsGZF2* were cloned into expression vector pRT101 under control of *CaMV 35S* promoter and in this way two effectors pRT101/OsGZF1 and pRT101/OsGZF2 were constructed (Figure 11A). The *GluB-1* promoter fragments of ProGluB-1-L and ProGluB-1-S2 (Figure 2A) used in the yeast one-hybrid screens were cloned in front of the *GUS* reporter gene to generate two reporter constructs ProGluB-1-L::GUS and ProGluB-S2::GUS (Figure 11A). Combinations of 6 µg of one of the effectors and 4 µg of one of the reporters were transformed together into protoplasts as indicated in Figure 11B. GUS activity was measured at different time point. To equilibrate any transformation efficiency bias, 3 µg of the 35S-Ω-Luc vector (firefly luciferase gene driven by the *CaMV 35S* promoter and Ω leader) was transformed at the same time with the effector and reporter as internal control. The luciferase standardized GUS activity of different effector and reporter combinations were linearised against the time interval and compared. As displayed in Figure 11B, both the pRT101/OsGZF1 and pRT101/OsGZF2 effector demonstrated a negative regulatory function on the two reporter constructs compared to the control vector. OsGZF1 down-regulates the GUS activity driven by ProGluB-1-L and ProGluB-S2 by 20% compared to the control. OsGZF2 reduced GUS activities of both reporters by more than 40%.

The above described experiments show that both OsGZF1 and OsGZF2 have repressing activities on *GluB-1* promoter driven reporter gene expression. To investigate this phenomenon in more detail, co-transformations were carried out with a *RISBZ1* overexpression construct. RISBZ1 is a transcription factor of the bZip type and is a known activator of *GluB-1* (Onodera et al., 2001; Yamamoto et al., 2006; Kawakatsu et al., 2009). In agreement with reported findings, also in our experimental set-up, RISBZ1 showed strong activation of *GluB-1* promoter GUS constructs. Transformation of 6 µg pRT101/RISBZ1 effector with ProGluB-1-L::GUS or ProGluB-1-S2::GUS could increase the GUS activity of both reporters 110 times compared to the control (Figure 11C). Then we transformed 3 µg of pRT101/RISBZ1, together with 3 µg of pRT101/OsGZF1 or pRT101/OsGZF2 effectors and also 4 µg of one of the reporters into protoplast respectively. Half amount of pRT101/RISBZ1 could still increase GUS activity 52 and 82 times when transformed
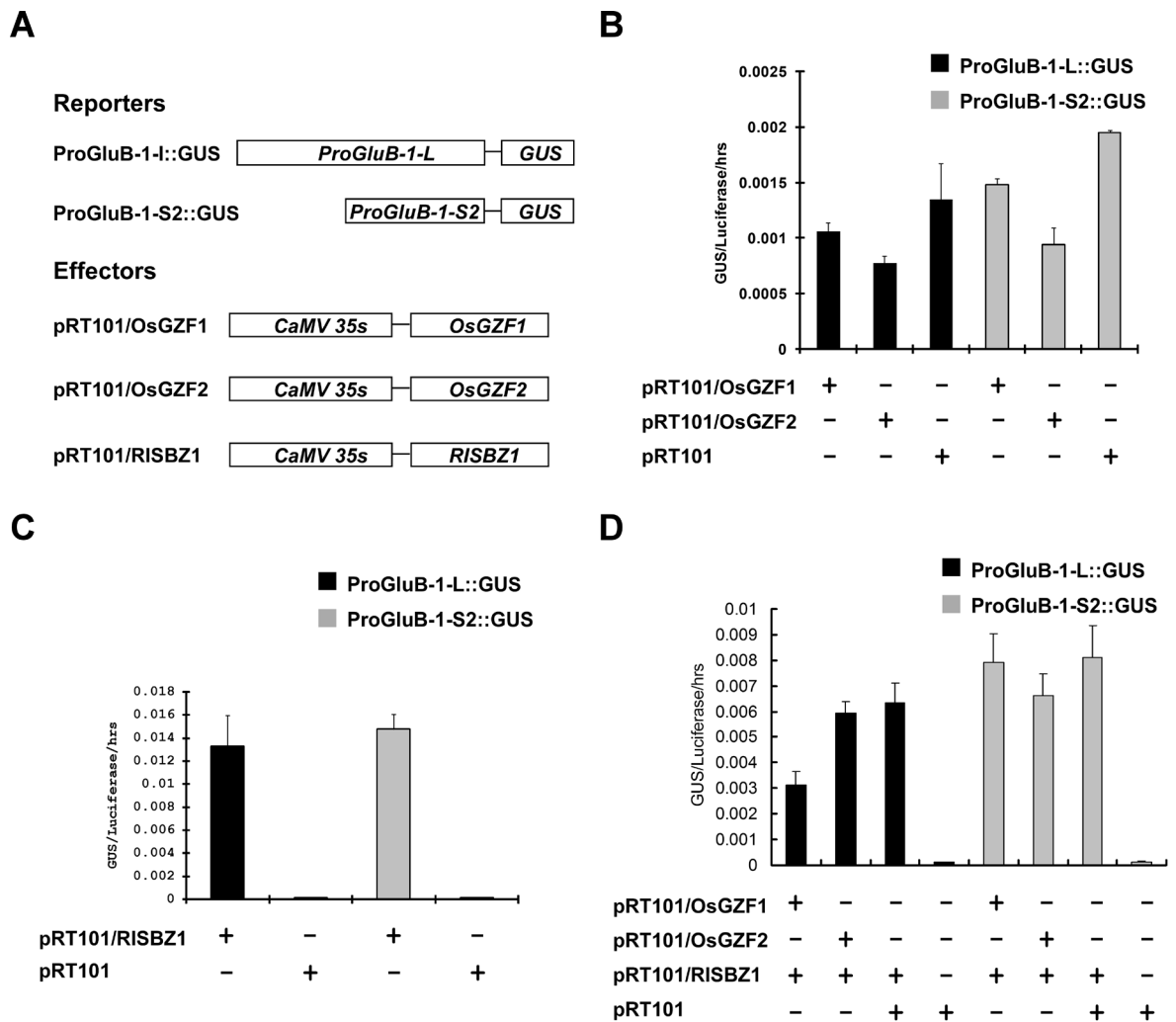
Figure 11. OsGZF1 and OsGZF2 down-regulate *GluB-1* promoter in transient assays. Rice seedling protoplasts were cotransformed with different combination of reporter and effector constructs as indicated. **A.** Schematic overview of reporter and effector constructs. **B.** Both OsGZF1 and OsGZF2 repress GUS activity driven by the *GluB-1* promoter. **C.** Effect of RISBZ1 on expression of ProGluB-1-L::GUS or ProGluB-S2::GUS reporters. **D.** Effects of OsGZF1 and OsGZF2 on the activation of ProGluB-1-L::GUS or ProGluB-S2::GUS by RISBZ1.

with ProGluB-1-L::GUS and ProGluB-1-S2::GUS. However, the OsGZF1 and OsGZF2 displayed a negative influence on the activation caused by RISBZ1 on these two reporters. OsGZF1 strongly down-regulated the activation of construct ProGluB-1-L::GUS by RISBZ1 and thus only 25 times of activation remained, whereas the effect of OsGZF2 on the activation of ProGluB-1-L::GUS by RISBZ1 was not significant. For the ProGluB-1-S2::GUS reporter, the results were different. OsGZF2 reduced activation by RISBZ1 by 19% while OsGZF1 showed no effects (Figure 11D).

A yeast two-hybrid system was applied in order to find out if OsGZF1 and OsGZF2 can interact with RISBZ1 *in vivo* and thus decreases its activation function on the *GluB-1* promoter. RISBZ1 was fused with GAL4 binding domain (pAS2-1) and activation domain (pACTII), respectively to form two constructs pAS2-1/RISBZ1 and pACTII/RISBZ1. Then the combination of pAS2-1/RISBZ1 with pACTII/OsGZF1 or 2 and pACTII/RISBZ1 with pAS2-1/OsGZF1 or 2 were

transformed into yeast strain PJ69-4A respectively. The transformants were grown on selection medium (CM-His and 2 mM 3-AT) for a week. However, no growth was observed on the plates supplied with 2 mM 3-AT (data not shown), and therefore no evidences for an interaction between the proteins was obtained.

Potential roles of *OsGZF1* in regulation of synthesis of glutelin

To further investigate the regulatory function of the two novel TFs in plants, we searched online databases for T-DNA and transposon mutants. For *OsGZF1* no mutant was found. However, for *OsGZF2* two T-DNA insertion lines 3D-00246/247 (Postech collection) and AEP E09 in (OTL collection) in japonica background were found. After aligning the flanking sequence tags of the two lines (D09173/176 and ER888440/441 respectively) with the genomic DNA sequence of LOC_Os11g28270, the position of inserted T-DNAs were determined at between the fifth and sixth exon for 3D-00276 and between the sixth and seventh exon for AEP E09, respectively. Homozygous plants were identified by Southern blotting and reduced *OsGZF2* expression was confirmed by northern blot with RNA isolated from 10 DAF grains. Seed total protein from five random seeds from different parts of the mature panicle was extracted and analyzed by SDS-PAGE. But no obvious protein change in the *OsGZF2* mutants was found (data not shown).

Since no mutant for *OsGZF1* was available, an RNAi strategy was used to create transgenic loss-of-function rice. For this an 800 bp fragment specific for *OsGZF1* was selected and expressed using the Hannibal system (Wesley et al., 2001) in Nipponbare rice. The transgenic plants DNA were analyzed by Southern blot and single insertion lines were selected for further studies. Total RNA was isolated from the 10 DAF transgenic seeds and four reduced *OsGZF1* plants (54, 59, 61 and 79) were identified (Figure 12B). We compared these four plants with two Nipponbare plants (NB-1, NB-2), and two transgenics but with unreduced *OsGZF1* expression (55, 60) were also analyzed as controls. Compared to Nipponbare plants, we did not observe obvious abnormal phenotype in seeds as well as vegetative organs (Figure 12A). The weight of 20 grains from the five *OsGZF1* low expression lines was compared with Nipponbare and no significant difference was found (Figure 12C). Seed proteins were isolated from five randomly selected T1 seeds of each line and analyzed by SDS-PAGE. The volume of isolation buffer was standardized by grain weight and equal amounts of extracts were loaded on gel. As displayed in Figure 12D, there are variations between the glutelin bands of two Nipponbare seeds NB-1 and NB-2. But all four low *OsGZF1* expressors exhibited an obvious increase in the intensity of the glutelin precursor band. Plants 54, 55 and 61 showed increased basic subunit bands and 54 showed also an increased acidic subunit bands compared to both of the Nipponbare control plants. However, plant 79 did not show obvious changes in either of the bands. In order to verify the results of SDS-PAGE, we measured the nitrogen concentration of OsGZF1-RNAi transgenic seeds by dry combustion on a Flash EA 2000 elemental analyzer. The nitrogen concentrations of 54 and 59 turned out to be significantly higher than that of the Nipponbare plants and the transgenic control plants 55 and 60. Plant 61 also showed a significantly higher nitrogen concentration compared to NB-1 and the transgenic controls, but the difference between plant 61 and NB-2 was not significant. Plant 79 did not have significant changes compared to the Nipponbare plants or the transgenic controls (Figure 12E). These results are consistent with the SDS-PAGE assays, indicating the important roles of OsGZF1 in the regulation of storage protein production.
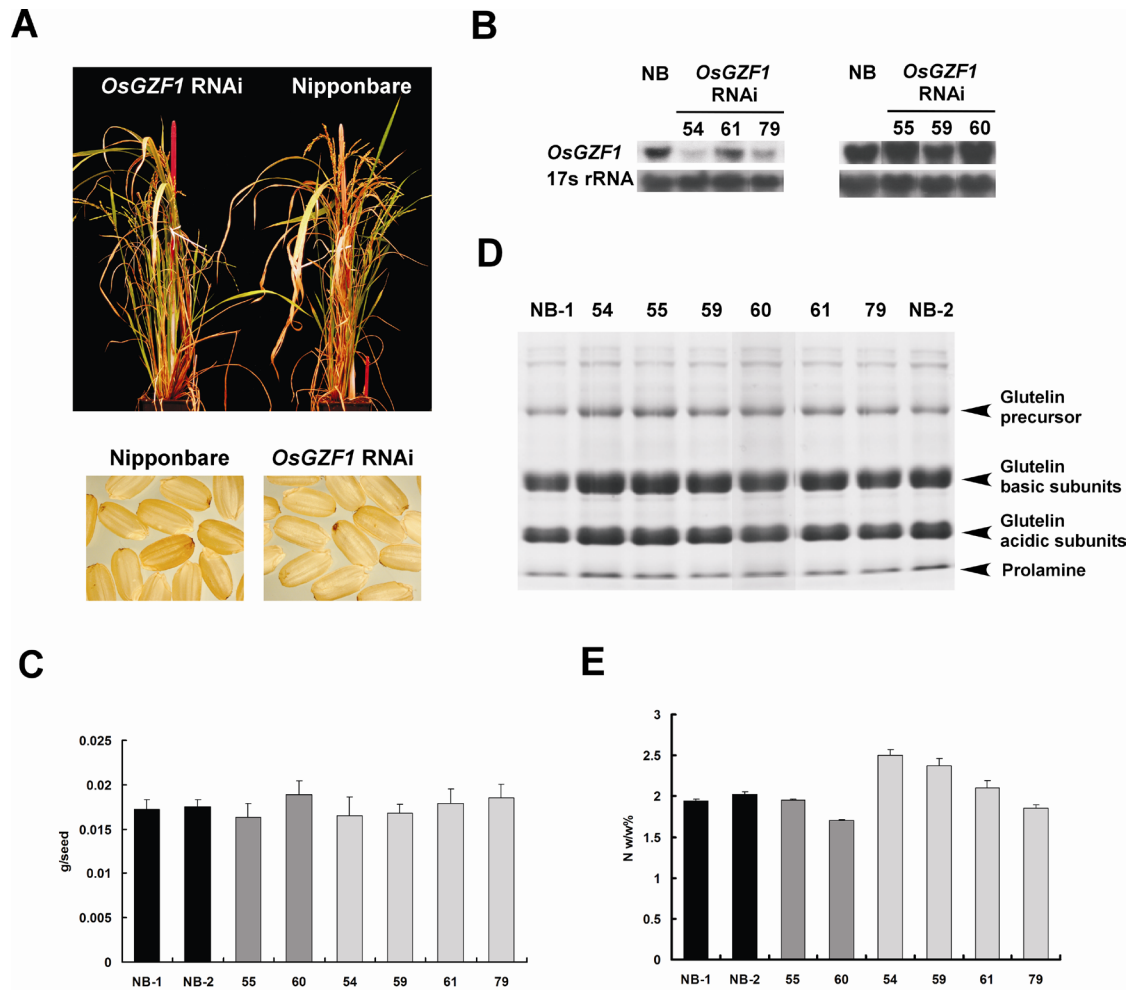
Figure 12. *OsGZF1* RNAi plants showed increased seed protein content. **A.** Overall comparison of *OsGZF1* RNAi and untransformed Nipponbare plants. **B.** Identification of low *OsGZF1* expression transgenics by northern blot. Total RNA of 10 DAF seeds was isolated and hybridized with a *OsGZF1* specific probe. **C.** Comparison of the weight of *OsGZF1* RNAi and untransformed seeds. Twenty random seeds of each plant were dehusked and weighted after drying at 50℃ for one week. **D.** Protein composition of *OsGZF1* RNAi and wild type seeds. Five random seeds of each plant were weighted and grinded thoroughly in isolation buffer. The volume of the buffer was added based on the dry weight. Equal volumes of the extracts were loaded on SDS-PAGE. **E.** Comparison of the nitrogen content of the *OsGZF1* RNAi and untransformed seeds. Three random seeds from each plant were grind into fine powder. Per sample, 4 mg of powder was used and the nitrogen concentration was measured by dry combustion on a Flash EA 2000 elemental analyser.

For a gain-of-function analysis, transgenic *OsGZF1* and *OsGZF2* overexpressing plants were made with *CaMV 35S* driven cDNAs and overexpression of the transgenes was confirmed by northern blot analysis. Only 8% (9/106) of the *OsGZF1* flowers set seeds and the rest of flowers showed premature stop of grain development within one week after fertilization. In contrast, *OsGZF2* overexpressors produced seeds normally. Seed protein content was checked in these plants by SDS-PAGE but no differences were found compared to wild type grains (data not shown).

Discussion

Regulatory functions of OsGZF1 and OsGZF2 on *GluB-1* expression
In this chapter we report on OsGZF1 and OsGZF2 to be the first identified repressors of a seed storage protein gene. All transcription factors known so far to be involved in regulation of cereal storage protein genes, for instance O2, RISBZ1 and PBF1, are activators. Several *in vitro* and *in vivo* experiments supported evidence for functions of OsGZF1 and OsGZF2 in regulation of the *GluB-1* promoter. Firstly, both the yeast one-hybrid assays and the EMSAs demonstrated that OsGZF1 and OsGZF2 can interact with the *GluB-1* promoter. Using a yeast one-hybrid system, seven *OsGZF1* and two *OsGZF2* cDNA clones respectively, were found and sequence analysis showed they had zinc-finger motifs which are consistent with a role in DNA binding and gene regulation. Secondly, using GFP-tagged proteins, both OsGZF1 and OsGZF2 showed nuclear localization which further supports a function in transcriptional gene regulation. Other support is from the expression patterns of the genes. Northern blot results showed that *OsGZF1* and *OsGZF2* were predominantly expressed in the developing seeds during filling stage, which is overlapping with the expression pattern of their potential target gene *GluB-1*. *OsGZF1* and *OsGZF2* are mainly expressed from 5 DAF onwards, peaking at 10 DAF but still detectable at 15 DAF. Histochemical assays of ProOsGZF1::GUS and ProOsGZF2::GUS transgenics revealed their main expressing location is in developing seeds and in a small domain in the endosperm closely around the embryo. In this region a series of ProGluB-1::GUS constructs showed expression too (Wu et al., 1998a, 1998b, 2000). Taken together, the data on the expression patterns showed that at least in some regions of grain development expression of *OsGZF1* and *OsGZF2* overlaps with *GluB-1* which would be consistent with functions in regulation.

Furthermore, the gain-of-function approach using a transient expression system clearly showed that both proteins could negatively control the *GluB-1* promoter. Considering that *GluB-1* is a seed-specific gene and the experiments in seedling protoplasts may not fully reflect its full transcriptional properties as in seed tissue, we added the well known transcriptional activator RISBZ1 of *GluB-1* into the assay in order to reconstitute a *GluB-1* regulatory system. RISBZ1 too is predominantly active in developing seeds and is functional through the GCN4 box in the core promoter of *GluB-1* (Onodera et al., 2001). At first, RISBZ1 was confirmed to be able to activate the *GluB-1* promoter in the rice protoplast system by co-transforming it with two constructs of the *GUS* reporter gene driven by the core *GluB-1* promoter and the bait sequence used in yeast one-hybrid respectively. Addition of OsGZF1 or OsGZF2 in the protoplast assays, however, reduced activation of the two *GluB-1* promoter-GUS constructs by RISBZ1. OsGZF1 decreased the activation by RISBZ1 function of the core promoter region of *GluB-1* by 50%, whereas OsGZF2 had no impact on this construct. On the other hand, OsGZF2 reduced the activation by RISBZ1 of the bait from yeast one-hybrid assays by 19%, while OsGZF1 showed no effect. OsGZF1 and OsGZF2 have three and five zinc fingers respectively and can interact with multiple sites as discussed earlier; they also do not interact with RISBZ1 *in vivo*. One possibility is that the binding of OsGZF1 and OsGZF2 to the *GluB-1* promoter blocks the GCN4 box in the *GluB-1* core promoter from interaction with RISBZ1. Such competitive DNA binding is one of the main ways in eukaryotes to limit the rate of transcription initiation from nearby promoters (Johnson, 1995). However, OsGZF1 and OsGZF2 demonstrated different effects on the two different *GluB-1* promoter constructs only differing in length for which we do not have a clear explanation yet.

Important *in planta* evidence supporting a role of OsGZF1 as repressor of *GluB-1* was from the analysis of seed storage protein content in *OsGZF1* loss-of-function plants made with RNAi technology. Using SDS-PAGE assays, an increased amount of precursor and two subunits of glutelin compared with untransformed Nipponbare seeds were observed. This result was further supported by measurements of the nitrogen concentration in seeds from OsGZF1 RNAi plants, in which an increased nitrogen concentration was found. Moreover, the changes in nitrogen concentration also indicated that OsGZF1 may be involved in the regulation of other glutelin genes in rice. Apart from *OsGZF1* loss-of function plants, we obtained and studied two T-DNA insertion mutants for *OsGZF2*. Seed protein of *OsGZF2* down-regulated plants were also analyzed by SDS-PAGE, but no change in protein composition was apparent in both mutants. Redundancy of OsGZF2 with OsGZF1 or other OsGZF proteins could be the reason for the absence of any effect on seed protein content in the two *OsGZF2* mutants.

The transient assay experiments with OsGZF1 and OsGZF2 indicated a negative regulatory function on the *GluB-1* promoter. OsGZF2 down-regulates the core promoter of *GluB-1* by 40% and OsGZF1 by 20%. However OsGZF1 could reduce the activation by RISBZ1 of the core promoter of *GluB-1* by 50% whereas OsGZF2 showed no influences. On the other hand, OsGZF1 RNAi seeds showed an increase in glutelin content whereas seeds from a T-DNA insertion mutant for OsGZF2 did not. These data indicate that the dominant repressing effects of OsGZF1 and OsGZF2 may act through affecting the function of other activators like RISBZ1. Further supporting evidence for a role in repressing expression of *GluB-1* is the expression patterns of *GluB-1*, *RISBZ1*, *OsGZF1* and *OsGZF2*. The expression of *GluB-1* reaches to highest level from 10 to 15 DAF (Onodera et al., 2001; Duan and Sun, 2005; Yamamoto et al., 2006), however *RISBZ1* start to accumulate from 5 DAF and remains at the same level till 15 DAF before declining (Onodera et al., 2001; Yamamoto et al., 2006) which is inconsistent with *GluB-1* expression. The results from the northern blot and promoter-GUS experiments showed that *OsGZF1* and *OsGZF2* are highly expressed from 5 to 10 DAF and peaks at 10 DAF, which is earlier than the peaking time of *GluB-1*. Promoter-GUS results also revealed that the expression region of *RISBZ1* (Onodera et al., 2001) partially overlapped with that of *OsGZF1* and *OsGZF2*. These results suggest that the function of RISBZ1 before 10 DAF may be affected by OsGZF1 and OsGZF2, and that the accumulation of *GluB-1* mRNA in this specific area may be negatively regulated by OsGZF1 and OsGZF2. Similar machinery has been demonstrated before in the regulation of the expression of storage protein genes in dicotyledons. The bean bZIP protein ROM1 functions as a DNA-binding repressor of the genes *DLEC2* and *PHSβ* (coding for phytohemagglutinin L-subunit and β-phaseolin) by reducing the activation by the *trans*-activator PvALF (Chern et al., 1996).

Interaction of OsGZF1 and OsGZF2 with the *GluB-1* promoter
CCCH type zinc finger proteins belong to a conserved group in pro- and eukaryotes. Most animal CCCH proteins investigated have been demonstrated to play important roles in post-transcriptional regulation of mRNAs by binding to the AU-rich element in the 3'-UTR (Guo et al., 2004; Bhandari and Saha, 2007; Kelly et al., 2007; Pagano et al., 2007; Prouteau et al., 2008; Viphakone et al., 2008; Hurt et al., 2009; Liang et al., 2009; Pomeranz et al., 2009; Schichl et al., 2009; Sinha et al., 2009; Stumpo et al., 2009; Tran et al., 2009). In plants, several CCCH proteins were also described as RNA-binding proteins (Li et al., 2001; Cheng et al., 2003; Addepalli and Hunt, 2007, 2008).

Compared to protein-RNA interactions, the knowledge of protein-DNA and protein-protein interaction ability of the CCCH zinc finger proteins is yet very poor. PEI1 is an Arabidopsis protein able to bind to specific DNA sequences (Li and Thomas, 1998). AtTZF1 can bind to both DNA and RNA *in vitro* (Pomeranz et al., 2009). GhZFP1 in cotton can interact with two other proteins and is induced by salt, drought and salicylic acid (Guo et al., 2009). In rice, OsLIC displayed binding activity to double-stranded DNA and single-stranded polyrA, polyrU and polyrG but not polyrC (Wang et al., 2008b). In this paper we describe two CCCH proteins in rice which demonstrated specific DNA binding ability both *in vitro* and *in vivo*.

Classic C2H2 type zinc finger proteins contain three or more repeating finger units and bind as monomers. Each finger binds a specific triplet base-pair sequence and tandemly linked fingers bind sequential triplets and interact with the same DNA strand. In this way, a strong and specific DNA-protein interaction is built up. However, it is not well understood if the CCCH zinc finger can bind specifically to DNA sequences like C2H2 fingers or that additional sequences are needed. But through a random binding site selection method it was demonstrated that recombinant CCCH type PEI1 protein in Arabidopsis can interact with specific DNA sequences (Li and Thomas, 1998). Moreover, many CCCH zinc finger proteins have tandemly arranged fingers too which is indicative of a high specificity of interaction.

According to our EMSA results, OsGZF1 and OsGZF2 can interact not only with the bait used in the yeast one-hybrid screens, but also can bind to the three other segments of the *GluB-1* core promoter region. Taken into account that OsGZF1 has three zinc finger motifs and OsGZF2 has five, and all the fingers are located with certain length of gaps in between, we suggest that these zinc fingers could interact with different part of the *GluB-1* promoter region at the same time. Nonetheless, mutations of any of the oligonucleotides, which have been demonstrated to be important for the downstream reporter gene expression (Yoshihara et al., 1996; Washida et al., 1999; Wu et al., 2000), had no influences on the interactions with OsGZF1 and OsGZF2 *in vitro*. We speculate that there are three potential reasons why mutations in AACA, GCN4 and PROL motifs have no impact on the interactions. The first is that oligonucleotides in the bait other than those motifs are the recognition sites for OsGZF1 and OsGZF2. This is reasonable because all these *cis-elements* are related to activation of downstream genes, but OsGZF1 and OsGZF2 showed a repression function *in vivo*. Secondly, it is also possible that the binding sites are located in the overlapping parts of the oligonucleotides used in yeast one-hybrid and EMSA which is probably the reason why OsGZF1 was able to interact with four different oligonucleotides while it has only three fingers. The third possibility is that because OsGZF1 and OsGZF2 have multiple fingers, there is always one able to interact with an oligonucleotide.

Potential roles of OsGZF1 and 2 in grain development
The biological functions of most members of the plant CCCH family are not clear so far. A reported Arabidopsis homologous gene of OsGZF1, PEI1, is highly conserved to OsGZF1 in the zinc-finger domain (Figure 13). PEI1 is an embryo-specific zinc-finger protein required for heart-stage embryo formation in Arabidopsis (Li and Thomas, 1998; Wang et al., 2008a), but we did not observe any expression within embryos in the histological assay of ProOsGZF1::GUS transgenic rice.

Cell differentiation in the embryo is taking place in the early stages after fertilization. At around 10 DAF, the size of the embryo increases rapidly and morphological differentiation is

```
            1                                                      50
OsGZF1      ....MAMHHM HDVFAQTTSI IGHEEEEVTI DPTKWGAWAH RGHRLWASMS
PEI1        MLKSASPMAF YDIGEQQYST FGYILSKPGN AGAYEIDPSI PNIDDAIYGS

            51                                                     100
OsGZF1      EDFWIHVYKV QRCPRSSSHD WTSCPYAHKG ERARRRDTRR FAYAAVSCPD
PEI1        DEFRMYAYKI KRCPRTRSHD WTECPYAHRG EKATRRDPRR YTYCAVACPA

            101                                                    150
OsGZF1      YRPREAAPGA VPSCAHGLRC RYAHGVFELW LHPSRFRTRM CSAGTRCPRR
PEI1        FRNG...... ..ACHRGDSC EFAHGVFEYW LHPARYRTRA CNAGNLCQRK

            151                                                    200
OsGZF1      ICFFAHSAAE LRDDPNSIAS AILTPMPMPM PMSIPMRTPD ADHHPRVSAM
PEI1        VCFFAHAPEQ LRQSEG.... .......... .......KHR CRYAYRPVRA

            201                                                    250
OsGZF1      RDQLDLIEEA MRNRLRLYSN ANANANVGSS SATAVATTTI PVSTLANGEG
PEI1        RGGGNGDGVT MRMDDEGYDT SRSPVRSGK. .......... ..DDLDSNEE

            251                                        289
OsGZF1      STGKRCGCRR CVEEEDALLN ....GYPHYD LIMDLVDE.
PEI1        KVLLKCWSRM SIVDDHYEPS DLDLDLSHFD WISELVD..
```

Figure 13. Comparison of protein sequences between OsGZF1 and its Arabidopsis homologous PEI1. The CCCH motifs are underlined.

finished. The transportation of substances to the embryo is performed through endosperm tissue both at the embryo development stage and during germination. Endosperm cells located closed to the embryo not only supply nutrients to the embryo but also are digested and absorbed by the embryo (Matsuo and Hoshikawa, 1993). The promoter GUS results showed that *OsGZF1* and *OsGZF2* are highly active in the endosperm region around the embryo during the early development stage. They also both showed negative regulation on *GluB-1* promoter *in vivo*. These data indicate that OsGZF1 and OsGZF2 may down regulate the promoter of *GluB-1* or affect the function of *GluB-1* activators such as RISBZ1 and thus decrease the expression level of *GluB-1* in this specific region.

The cell differentiation in the aleurone layer is finished around 10 DAF (Matsuo and Hoshikawa, 1993). *OsGZF2* is active in the aleurone layer after fertilization as shown by promoter-GUS assays and northern blot results showed its mRNA level peaking around ten day after flowering suggesting a potential role in gene regulation in this tissue and developmental phase. The influence of OsGZF2 on RISBZ1 function is much less than that of OsGZF1 in transient assays and no obvious alteration of seed protein content is found in OsGZF2 T-DNA mutants. These results suggest that the function of OsGZF2 might not be linked directly to the regulation of the protein composition but to the cell differentiation in the aleurone layer in contrast to OsGZF1. However, to further validate this hypothesis more assays need to be conducted. On the other hand, no *OsGZF2* expression was found in the aleurone layer of germinating seeds which means OsGZF2 is likely not involved in the mobilization of enzymes in the aleurone such as for instance α-amylase which is essential for breakdown of starch during germination. We tested if α-amylase activity in response to gibberellin acid (GA3) is impaired in *OsGZF2* loss-of function seeds during germination but no differences were found compared to the wild type (data not shown), which further supports that OsGZF2 is not needed for the activity of α-amylase during germination.

In summary, our results revealed that within the endosperm, regulation of *GluB-1* expression could occur differently in different regions. Although OsGZF1 and OsGZF2 both act *in vivo* as

repressors for *GluB-1* in developing seeds, promoter-GUS results show their expression patterns are not fully overlapping which may reflect to their biological functions. OsGZF1 is involved in determination of the glutelin final content in grains whereas OsGZF2 is more likely related to certain tissue differentiation. These results provide novel insight and knowledge in our understanding of the complex regulation mechanism of *GluB-1* which may have implications for other storage protein genes in rice given the resemblances between regulatory sequences in these genes.

Materials and methods

Plant materials

Japonica rice cultivars used in these studies were Nipponbare and Zhonghua 11. Rice plants were grown in 10×10 cm plastic pots in the green house at 28°C and 80% humidity. The light regime was 12hr day and 12hr night. Plant material for DNA and total RNA extraction were frozen in liquid nitrogen immediately after harvesting and stored at -80°C until use.

RNA isolation and cDNA library construction

Total RNA for cDNA library construction was isolated according to (Gao et al., 2001), except that the extraction buffer was composed of 100mM Tris-HCL, 100mM LiCl, 10mM EDTA, 1% SDS (pH 9.0), 2% PVP (Sigma, PVP40), 2% BSA (Sigma) and 10 mM DTT (Sigma). Metal and glassware were made RNase-free by baking at 180°C for 6 hours. Plastic ware was made RNAse freee by soaking in freshly prepared 50% bleach (12% sodium hypochlorite) for over 2 hours and rinsed with sterilized water. mRNA was isolated using the PolyATtract mRNA Isolation System III according to manufacturer's instructions (Promega). cDNA libraries of panicle and seed of filling stage were constructed from 5 µg of poly(A)$^+$ RNA with the HybriZAP-2.1 XR cDNA Synthesis and Library Construction Kits according to the manufacturer's instructions (Stratagene). cDNA was separated by the Sepharose Drip Column and the fragments larger than 300bp were combined, condensed and cloned between *EcoR*I and *Xho*I of *E.coli-yeast* shuttle vector λACTII (Memelink, 1997) instead of HybriZAP-2.1 from the kit. The average titration of five packages was ~2×10^5 and in total over 10^6 plaques was generated for each primary library. The primary libraries were then amplified to 10^{9}pfu/ml culture and the phage libraries were then circularized into pACTII-based plasmid libraries via Cre-recombinase-mediated mass excision in *E. coli* BNN132 (Ouwerkerk and Meijer, 2001).

Yeast one-hybrid screening

The 291bp (-346~-55) bait ProGluB-1-L was amplified using primers PGBL-for and PGBL-rev (Supplemental table 1) from Nipponbare genomic DNA with Phusion Polymerase (Finnzymes). The 216bp (-382, -166) and 169bp (-190,-21) overlapping baits ProGluB-1-M1 and ProGluB-1-M2 were amplified by two pairs of primers PGBM1-for and PGBM1-rev, PGBM2-for and PGBM2-rev (Supplemental Table 1). The products were cloned into pCR-Blunt II TOPO (Invitrogen) and sequenced (Baseclear, Leiden, The Netherlands). Four other *GluB-1* bait constructs ProGluB-1-S1 (-248,-194), ProGluB-1-S2 (-203, -144), ProGluB-1-S3 (-153, -104) and ProGluB-1-S4 (-109, -54) were based on four overlapping oligonucleotides with attached *Not*I and *Xba*I sites (Supplemental table 2). The PCR products were excised from the pTOPO vectors with *Not*I/*Xba*I and the annealed oligos were directly cloned into the integrative vector pINT1-HIS3NB (Genbank Accession

AY061966, (Meijer et al., 1998; Ouwerkerk and Meijer, 2011)) between *Not*I and *Xba*I. Next, linearized fragments containing the *HIS3* reporter and the *APT1* selection marker were transformed into yeast strain Y187 (*MATα, ura3-52, his-Δ200, ade2-101, trp1-901, leu2-3, 112met⁻, gal4 gal8 URA3::GAL1_{UAS}-GAL1_{TATA}-laZ, MEL1*; Harper et al., 1993; Clontech) and integrated at the non-essential *PDC6* locus via double cross-over. Positive colonies were selected on YAPD-G418 plates. Then positive strains were grown on CM plates lacking histidine to test leaky *HIS3* expression. Yeast one-hybrid screenings were performed according to previously described (Ouwerkerk and Meijer, 2001, 2011). One µg of pACTII-cDNA library was transformed into yeast and 10^6 transformants were screened. Selection was performed on CM plates lacking leucine and histidine and containing 10 mM 3-AT (a competitive inhibitor of the His3p reporter protein). Positive colonies were re-streaked on the same selective plates and the colonies still growing after this step were selected for plasmid isolation or colony PCR.

Colony PCR

Colony PCR was performed with primers COPCR-for and COPCR-rev (Supplemental Table 1) to amplify the cDNAs from the putative pACTII-cDNA clones. Phusion polymerase (Finnzymes) was used and the reaction system was set up according to manufacturer's instructions. Yeast cells were picked with autoclaved pipet tips directly from the plates and resuspended into the reaction transformation mixtures. The amplification was done with 98ºC 5 min, 98ºC 30 sec, 52ºC 30 sec, 72ºC 2 min, 35 cycles, 7ºC 10 min, 4ºC stop.

Electrophoretic Mobility Shift Assay

OsGZF1 cDNA was subcloned in frame with the GST-encoding sequence in GST expression vector pGEX-KG (Guan and Dixon, 1991) as *Nco*I/*Xho*I fragment from pACTII-OsGZF1. *OsGZF2* was cloned into pGEX-KG vetor between *Nco*I and *Xho*I sites using two fragments. The first fragment from N-terminal was amplified with primers ZF2GST-for and ZF2GST-rev (Supplemental Table 1). The PCR product was cloned into pCR-Blunt II TOPO (Invitrogen) and sequenced (Baseclear, Leiden, The Netherlands). Next the fragment was excised with *Nco*I/*Nde*I and ligated together with the second fragment from the C-terminus excised with *Nde*I and *Xho*I from pACTII-OsGZF2 vector into pGEX-KG between *Nco*I and *Xho*I. Overnight *E. coli* XL-1 cultures (10 ml) containing the expression constructs were diluted into 500ml LB medium and grown to OD600=0.5 at 37 ºC. The culture was then induced by adding IPTG to 1mM. After 3.5h (OD600=0.8) of growth at 29ºC, cells were pelleted, resuspended in 20 ml of $1 \times$ PBS buffer ($10 \times$ stock 1L: 80 g NaCl, 2 g KCl, 7.65 g Na_2HPO_4-$2H_2O$, 2 g KH_2PO_4), and lysed by mild sonication. Following centrifugation, the supernatants were filtered using a syringe through a 0.45 µm membrane filter. Samples were run over a plastic column with 0.5ml Glutathione-Sepharose (GE Healthcare) and washed with 10ml of 1×PBS buffer. Next, the bound GST fusion proteins were eluted with eight times volume of 0.25ml Glutathione Elution Buffer (10 mM glutathione, 50 mM Tris-HCl pH8.0) and frozen in liquid nitrogen after addition of 10% glycerol and store at -80ºC.

EMSAs were performed in a 10 µl system: 0.1-0.2 µg of purified fusion proteins, 0.1 ng P^{32}-end labeled probe (60 bp double strand oligonucleotides), variable amount of poly (dI-dC)-poly (dIdC) (Amersham-Pharmacia) depending on the protein/probe combination, 2 µl five times nuclear extraction buffer (Green et al., 1989), and water up to 10 µl. The reaction mixtures were incubated at

room temperature for 30 min and loaded on 5% acrylamide/bisacrylamide (37.5:1) gels in 0.5 × TBE buffer under tension. After electrophoresis, gels were vacuum dried on DE81 Whatman paper at 65ºC for one hour and autoradiographed. The DNA fragments were the same four sequences as used in the yeast one-hybrid assays, as well as 12 other mutant oligonucleotides derived from these with different mutations at the key motifs as described in the text and figure legend (sequences listed in Supplemental Table 3).

Southern and northern blotting
Ten µg of genomic DNA per sample was used for Southern blotting. Total RNA of vegetative tissues were isolated with TRIzol Reagent (Gibco-BRL) according to the manufacturer's instructions. Total RNA of seeds was isolated according to (Singh et al., 2003). Twenty µg of total RNA of different tissues of Nipponbare rice were separated by electrophoresis and transferred onto a cellulose membrane. *OsGZF1*, *OsGZF2* specific probes were then hybridized on the same membrane one after the other. Blotting and hybridization were as described previously (Memelink et al., 1987). For *OsGZF1*, an 801 bp probe was excised from vector pHannibal/OsGZF1 (constructed for the RNAi experiment) with *Kpn*I-*Xho*I. For OsGZF2, a 246 bp specific fragment cut with *Eco*RI and *Sal*I from pACTII-OsGZF2. The *GluB-1* probe was amplified from Nipponbare genomic DNA with primers GBNOR-for and GBNOR-rev (Supplemental Table 1).

Subcellular localization of OsGZF1 and OsGZF1
A GFP/OsGZF1 fusion construct for localization studies was made as following: A full length cDNA clone of *OsGZF1* was excised from pACTII/OsGZF1 with *Eco*RI/*Xho*I and ligated into *Bgl*II/*Xho*I digested pTH2-BN, (Kuijt et al., 2004) which is a derivative of pTH2 lacking the GFP stop codon (Chiu et al., 1996; Niwa et al., 1999), together with a *Bgl*II/*Eco*RI adapter. The adapter was made by annealing of two oligos 5'-GATCTGTCGACG-3' and 5'-GAATTCGTCGACA-3'. Construct GFP/OsGZF2 was made by amplifying the ORF of *OsGZF2* from pACTII/OsGZF2. The PCR product was sequenced (Baseclear, Leiden, The Netherlands), digested with *Sal*I/*Nco*I and ligated into *Xho*I/*Nco*I-digested pTH2. The vectors were then transformed into Nipponbare seedling protoplasts through a PEG-mediated method and incubated in dark at 25ºC overnight. GFP was visualized with an LSM 5 Exciter on an AXIO Observer confocal laser scanning microscope (ZEISS).

Binary vectors construction, plant transformation and detection of GUS expression
The binary vector to make the *OsGZF1* RNAi plants was constructed as follows. A ~800bp *OsGZF1* specific region was amplified on pACTII/OsGZF1 vector with primer pairs: sense fragment ZF1SN-for and ZF1SN-rev; anti-sense fragment ZF1ASN-for and ZF1ASN-rev (Supplemental Table 1). The fragments were cloned into pCR-Blunt II TOPO (Invitrogen) and verified by sequence analysis (Baseclear, Leiden, The Netherlands). Next, the sense and anti-sense segments were subcloned into vector pHannibal (Genbank Accession AJ311872; Wesley et al., 2001) separately as *Eco*RI/*Xho*I and *Cla*I/*BamH*I excised products resulting in construct pHannibal/OsGZF1. Then the fragment containing CaMV 35S promoter, sense, anti-sense part and OCS (octopine synthase) terminator was excised out with *Sac*I/*Spe*I and ligated into *Sac*I/*Xba*I cutted vector pC1300intA (Genbank Accession AF294976).

OsGZF1 and *OsGZF2* overexpression constructs were made by subcloning of *Eco*RI-*Xho*I fragments from pACTII/OsGZF1 and pACTII/OsGZF2, the original clones obtained in yeast one-hybrid screenings, into binary expression vector pC1300intB-35SnosEX (Genbank Accession AY560325; (Kuijt et al., 2004)).

Plasmids for expression pattern analysis were constructed as follows. The 5' regulatory regions of *OsGZF1* (-2,061 to 0) and *OsGZF2* (-2,603 to -1) were amplified from genomic DNA from cultivar Nipponbare using the following pairs of primers: PROZF1-for and PROZF1-rev; PROZF2-for and PROZF2-rev (Supplemental table 1). The PCR products were sequenced (Baseclear, Leiden, The Netherlands) and subcloned into pCAMBIA-1391Z (Genbank Accession AF234312; Hajdukiewicz et al., 1994) vector as *Hind*III/*Nco*I fragments.

Transgenic plants were made in either cultivar Nipponbare or Zhonghua 11. Transformation, histochemical detection of GUS activity, cytological techniques and microscopy were performed as previously described (Scarpella et al., 2000) except that *Agrobacterium tumefaciens* strain LBA 4404 was used instead of LBA 1119. Pictures were acquired using a Leica MZ12 stereo microscope equipped with a Sony 3CCD Digital Photo Camera DKC-5000.

RNA *in situ* hybridization

Digoxygenin (DIG) labeled anti-sense and sense RNA probes were prepared by *in vitro* transcription according to the instruction provided with the DIG RNA Labeling Kit (SP6/T7; Roche). cDNA used was amplified on the template of plasmid pACTII/OsGZF1 and pACTII/OsGZF2 with following primers: ZF1STASN-for and ZF1STASN-rev; ZF1STSN-for and ZF1STSN-rev; ZF2STASN-for and ZF2STASN-rev; ZF2STSN-for and ZF2STSN-rev (Supplemental Table 1). Young seeds of 7 DAF from Nipponbare rice were fixed in 3.7% FAA (5% acetic acid, 50% ethanol, and 3.7% formaldehyde in water), dehydrated through an ethanol series and embedded in paraffin. Paraffin blocks were cut into 7 µm thin sections and mounted on POLYSINE[TM] slides (Menzel GmbH & Co KG). Pretreatment of the sections was done as described by (Ferrandiz and Sessions, 2008). Subsequently, the sections were hybridized as described by (Caňas et al., 1994) with a modified hybridization temperature at 50°C. After hybridization, formamide washing and immunological detection were performed as described previously (Kouchi and Hata, 1993) with minor modifications. The slides were washed twice with $2 \times$ SSC, 50% deionized formamide at 50°C for 45 minutes after hybridization and washed twice with $1 \times$ SSC, 50% deionized formamide at 55°C for 60 minutes after RNAseA treatment. Then the sections were incubated with 1:1000 diluted alkaline phosphatase-conjugated antidigoxigenin antibody at 4°C overnight and the colour reaction was incubated at 30°C for 5 hours. Slides were mounted with a cover slip and pictures were made with a Zeiss Axioplan 2 upright microscope and a Sony 3CCD Digital Photo Camera DKC-5000.

Construction of reporter and effector plasmid and transient assay

Construction of effecter vectors was as follows, cDNA's of *OsGZF1* and *OsGZF2* were amplified from the pACTII/OsGZF1 and pACTII/OsGZF2 vector with primer sets: ZF1PR-for and ZF1PR-rev; ZF2PR-for and ZF2PR-rev (Supplemental Table 1). *RISBZ1* cDNA was amplified from cDNA of rice panicle 15 DAF (Agalou et al., 2008) with RISBZ1PR-for and RISBZ1PR-rev (Supplemental Table 1). The fragments were ligated into pCR-Blunt II TOPO (Invitrogen) and the sequences were determined (Baseclear, Leiden, The Netherlands). The cDNA of *OsGZF1, OsGZF2* and *RISBZ1*

were then subcloned into pRT101 (Töpfer et al., 1987) vector between *Eco*RI *and Kpn*I to form three effector constructs. Reporter vectors were constructed as follows. ProGluB-1-L promoter region was digested out from ProGluB-1-L/pINT1-HIS3NB vector by *Not*I/*Spe*I and cloned into *Eag*I/*Spe*I digested GUSSH-47 (Pasquali et al., 1994) vector. ProGluB-1-S2 was cloned into *Eag*I/*Xba*I digested GUSSH-47 vector by annealing the oligos used in the yeast one-hybrid assay.

Transient assays were performed as the method previously described (Chen et al., 2006) protoplasts isolated from two week old Nipponbare seedlings. GUS activity was measured by a Cytofluor 2350 fluorimeter (Millipore) according to (Jefferson, 1987). Plasmid FLUC (the firefly luciferase gene driven by the CaMV 35S promoter and Ω leader) was added as a control for transformation efficiency. Luciferase activity measurement was done according to the manufacturer's instructions (Promega) and the activity was read in a Perkin Elmer Victor 3. GUS activity was standardized against the luciferase activity.

Activation and two-hybrid assays in yeast

OsGZF1 and *OsGZF2* cDNAs were fused in frame with the GAL4 DNA-binding domain (BD) sequence in pAS2-1 (Clontech). The ORF of *OsGZF1* was cut from pACTII-OsGZF1 with *Nco*I/*Xho*I and cloned into *Nco*I/*Sal*I digested pAS2-1. For *OsGZF2*, a pair of primers ZF2PAS-for and ZF2PAS-rev (Supplemental Table 1) with an *Nco*I site at the beginning of the forward primer which is in frame with the ATG start codon was used to amplify the cDNA on pACTII-OsGZF2. PCR products were cloned into pCR-Blunt II TOPO (Invitrogen) and sequenced (Baseclear, Leiden, The Netherlands). The *Nco*I/*Nde*I digested fragment was then ligated together with the *Nde*I/*Xho*I fragment from pACTII/OsGZF2 into pAS2-1 vector between *Nco*I *and Sal*I. *RISBZ1* ORF was cut from the RISBZ1/pTOPO with E*co*RI/*Kpn*I and cloned into pUC18 vector (Genbank Accession L09136), then subcloned into pAS2-1 between *Eco*RI and *Bam*HI sites. The *Eco*RI/*Sal*I fragment of RISBZ1/pUC18 was also ligated in frame with the GAL4 activation domain in pACTIIA vector between *Eco*RI and *Xho*I. Finally, vectors or combination of vectors according to the text were transformed into yeast strain PJ69-4A (*MATα, trp1-901, leu2-3, ura3-52, his3-Δ200, gal4 gal80Δ, LYS2::GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, GAL2$_{UAS}$-GAL21$_{TATA}$-ADE2, met2::GAL7-LacZ* (Clontech)) as described (Ouwerkerk and Meijer, 2001, 2011). Positive colonies from CM-Trp plates were restreaked on CM-His, CM-His+2mM3-AT and CM-Ade plates and grow at 30°C for one week prior to scoring.

Seed protein isolation and SDS-PAGE

Five random seeds from different part of the mature panicle were weighed and grinded in mortar in isolation buffer (Kawakatsu et al., 2008) added according to the weight of the seeds. The mixtures were incubated overnight at 28°C under shaking (200 rpm) and centrifuged at 14,000 rpm for 10 min to get rid of debris. Finally, proteins were separated by 12% SDS-PAGE and the gel was stained with Commassie Brilliant Blue.

Measurement of nitrogen concentration in rice grains

Three random seeds from each plant were grind into fine powder. Per sample, 4 mg of powder was weighed in a 5×8 mm tin capsule and the nitrogen concentration was measured by dry combustion on a Flash EA 2000 elemental analyser (Thermo Electron Corporation, Rodana, Italy). The assays

were performed with three replicates and the nitrogen concentrations were analyzed statistically by Student's t-Test.

Acknowledgement

Reference

Addepalli B, Hunt AG (2007) A novel endonuclease activity associated with the Arabidopsis ortholog of the 30-kDa subunit of cleavage and polyadenylation specificity factor. Nucleic Acids Res **35:** 4453-4463

Addepalli B, Hunt AG (2008) Ribonuclease activity is a common property of Arabidopsis CCCH-containing zinc-finger proteins. FEBS Lett **582:** 2577-2582

Agalou A, Purwantomo S, Overnas E, Johannesson H, Zhu X, Estiati A, de Kam RJ, Engstrom P, Slamet-Loedin IH, Zhu Z, Wang M, Xiong L, Meijer AH, Ouwerkerk PBF (2008) A genome-wide survey of HD-Zip genes in rice and analysis of drought-responsive family members. Plant Mol Biol **66:** 87-103

Bhandari D, Saha P (2007) mRNA cycling sequence binding protein from Leishmania donovani (LdCSBP) is covalently modified by ubiquitination. FEMS Microbiol Lett **273:** 206-213

Caňas LA, Busscher M, Angenent GC, Beltran JP, van Tunen AJ (1994) Nuclear localization of the petunia MADS box protein FBP1 in petal and stamen cells. Plant J **6:** 597-604

Chen S, Tao L, Zeng L, Vega-Sanchez M, Umemura K,, Wang GL (2006) A highly efficient transient protoplast system for analyzing defence gene expression and protein-protein interactions in rice. Mol Plant Pathol. **7**: 417-427

Cheng Y, Kato N, Wang W, Li J, Chen X (2003) Two RNA binding proteins, HEN4 and HUA1, act in the processing of AGAMOUS pre-mRNA in Arabidopsis thaliana. Dev Cell **4:** 53-66

Chern MS, Eiben HG, Bustos MM (1996) The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. Plant J **10:** 135-148

Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J (1996) Engineered GFP as a vital reporter in plants. Curr Biol **6:** 325-330

Cord Neto G, Yunes JA, da Silva MJ, Vettore AL, Arruda P, Leite A (1995) The involvement of Opaque 2 on beta-prolamin gene regulation in maize and Coix suggests a more general role for this transcriptional activator. Plant Mol Biol **27:** 1015-1029

Duan M, Sun SS (2005) Profiling the expression of genes controlling rice grain quality. Plant Mol Biol **59:** 165-178

Ferrandiz C, Sessions A (2008) Nonradioactive In Situ Hybridization of RNA Probes to Sections of Plant Tissues. *In*, Vol 2008, pp pdb.prot4943

Gao J, Liu J, Li B, Li Z (2001) Isolation and purification of functional total RNA from blue-grained wheat endosperm tissues containing high levels of starches and flavonoids. Plant Mol Biol Reporter **19:** 185-186

Green P, Kay S, Lam E, Chua NH (1989) In vitro DNA footprinting. *In* Plant Mol Biol Manual. Kluwer Academic Publishers, Dordrecht, pp B11/11-22

Guo X, Carroll JW, Macdonald MR, Goff SP, Gao G (2004) The zinc finger antiviral protein directly binds to specific viral mRNAs through the CCCH zinc finger motifs. J Virol **78:** 12781-12787

Guo YH, Yu YP, Wang D, Wu CA, Yang GD, Huang JG, Zheng CC (2009) GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. New Phytol **183:** 62-75

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of Agrobacterium binary vectors

for plant transformation. Plant Mol Biol **25:** 989-994

Hamaker BR, Griffin VK, Moldenhauer KAK (1991) Potential Influence of a Starch Granule-associated Protein on Cooked Rice Stickiness. Journal of Food Science **56:** 1327-1329

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

Holdsworth MJ, Munoz-Blanco J, Hammond-Kosack M, Colot V, Schuch W, Bevan MW (1995) The maize transcription factor Opaque-2 activates a wheat glutenin promoter in plant and yeast cells. Plant Mol Biol **29:** 711-720

Hurt JA, Obar RA, Zhai B, Farny NG, Gygi SP, Silver PA (2009) A conserved CCCH-type zinc finger protein regulates mRNA nuclear adenylation and export. J Cell Biol **185:** 265-277

Hwang YS, Ciceri P, Parsons RL, Moose SP, Schmidt RJ, Huang N (2004) The maize O2 and PBF proteins act additively to promote transcription from storage protein gene promoters in rice endosperm cells. Plant Cell Physiol **45:** 1509-1518

Izawa T, Foster R, Nakajima M, Shimamoto K, Chua NH (1994) The Rice bZIP Transcriptional Activator RITA-1 Is Highly Expressed during Seed Development. Plant Cell **6:** 1277-1287

Jefferson R (1987) Assaying chimeric genes in plants: The GUS gene fusion system. Plant Mol Biol Reporter **5:** 387-405

Johnson AD (1995) The price of repression. Cell **81:** 655-658

Kang HJ, Hwang IK, Kim KS, Choi HC (2006) Comparison of the physicochemical properties and ultrastructure of japonica and indica rice grains. J Agric Food Chem **54:** 4833-4838

Katsube-Tanaka T, Duldulao JB, Kimura Y, Iida S, Yamaguchi T, Nakano J, Utsumi S (2004) The two subfamilies of rice glutelin differ in both primary and higher-order structures. Biochim Biophys Acta **1699:** 95-102

Kawakatsu T, Yamamoto MP, Hirose S, Yano M, Takaiwa F (2008) Characterization of a new rice glutelin gene GluD-1 expressed in the starchy endosperm. J Exp Bot **59:** 4233-4245

Kawakatsu T, Yamamoto MP, Touno SM, Yasuda H, Takaiwa F (2009) Compensation and interaction between RISBZ1 and RPBF during grain filling in rice. Plant J **59:** 908-920

Kelly SM, Pabit SA, Kitchen CM, Guo P, Marfatia KA, Murphy TJ, Corbett AH, Berland KM (2007) Recognition of polyadenosine RNA by zinc finger proteins. Proc Natl Acad Sci USA **104:** 12306-12311

Kim SY, Wu R (1990) Multiple protein factors bind to a rice glutelin promoter region. Nucleic Acids Res **18:** 6845-6852

Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol Gen Genet **238:** 106-119

Kreis M, Shewry PR, Forde BG, Forde J, Miflin BJ (1985) Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley and rye. *In* B Miflin, ed, Oxford Surveys of Plant Molecular and Cell Biology. Oxford University Press, Oxford, UK, pp 253-317

Kuijt SJH, Lamers GEM, Rueb S, Scarpella E, Ouwerkerk PBF, Spaink HP, Meijer AH (2004) Different subcellular localization and trafficking properties of KNOX class 1 homeodomain proteins from rice. Plant Mol Biol **55:** 781-796

Kusaba M, Miyahara K, Iida S, Fukuoka H, Takano T, Sassa H, Nishimura M, Nishio T (2003) Low glutelin content1: a dominant mutation that suppresses the glutelin multigene family via RNA silencing in rice. Plant Cell **15:** 1455-1467

Li J, Jia D, Chen X (2001) HUA1, a regulator of stamen and carpel identities in Arabidopsis, codes for a nuclear RNA binding protein. Plant Cell **13:** 2269-2281

Li Z, Thomas TL (1998) PEI1, an embryo-specific zinc finger protein gene required for heart-stage embryo formation in Arabidopsis. Plant Cell **10:** 383-398

Liang J, Lei T, Song Y, Yanes N, Qi Y, Fu M (2009) RNA-destabilizing factor tristetraprolin negatively regulates NF-kappaB signaling. J Biol Chem **284:** 29383-29390

Lohmer S, Maddaloni M, Motto M, Di Fonzo N, Hartings H, Salamini F, Thompson RD (1991) The maize regulatory locus Opaque-2 encodes a DNA-binding protein which activates the transcription of the b-32 gene. EMBO J **10:** 617-624

Matsuo T, Hoshikawa K (1993) Science of the Rice Plant. Volume One: Morphology, Vol 1, Tokyo

Meijer AH, Ouwerkerk PBF, Hoge JHC (1998) Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. Yeast **14:** 1407-1415

Memelink J (1997) Two Yeast/Escherichia coli lamba/ plasmid vectors designed for yeast one- and two-hybrid screens that allow directional cDNA cloning. Technical Tips Online TT01111 (http://research.bmn.com/). Trends Genet **13:** 376

Memelink J, Hoge JHC, Schilperoort RA (1987) Cytokinin stress changes the developmental regulation of several defence-related genes in tobacco. EMBO J **6:** 3579-3583

Mitsukawa N, Hayashi H, Yaiyiamoto K, Kidzu K, Konishi R, Masutvlura T, Tanaka K (1998) Molecular Cloning of a Novel Glutelin cDNA from Rice Seeds. Plant Biotechnology **15:** 205-211

Muth JR, Muller M, Lohmer S, Salamini F, Thompson RD (1996) The role of multiple binding sites in the activation of zein gene expression by Opaque-2. Mol Gen Genet **252:** 723-732

Nakase M, Yamada T, Kira T, Yamaguchi J, Aoki N, Nakamura R, Matsuda T, Adachi T (1996) The same nuclear proteins bind to the 5'-flanking regions of genes for the rice seed storage protein: 16 kDa albumin, 13 kDa prolamin and type II glutelin. Plant Mol Biol **32:** 621-630

Niwa Y, Hirano T, Yoshimoto K, Shimizu M, Kobayashi H (1999) Non-invasive quantitative detection and applications of non-toxic, S65T-type green fluorescent protein in living plants. Plant J **18:** 455-463

Okita TW, Hwang YS, Hnilo J, Kim WT, Aryan AP, Larson R, Krishnan HB (1989) Structure and expression of the rice glutelin multigene family. J Biol Chem **264:** 12573-12581

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

Ouwerkerk PBF, Meijer AH (2001) Yeast one-hybrid screening for DNA-Protein Interactions. Current Protocols in Molecular Biology 12.12.1-12.12.22

Ouwerkerk PBF, Meijer AH (2011) Yeast one-hybrid screens for detection of transcription factor DNA interactions. Methods in Molecular Biology **678:** 211-227

Pagano JM, Farley BM, McCoig LM, Ryder SP (2007) Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. J Biol Chem **282:** 8883-8894

Pasquali G, Ouwerkerk PBF, Memelink J (1994) Versatile transformation vectors to assay the promoter activity of DNA elements in plants. Gene **149:** 373-374

Pomeranz MC, Hah C, Lin PC, Kang SG, Finer JJ, Blackshear PJ, Jang JC (2009) The Arabidopsis tandem zinc finger protein AtTZF1 traffics between the nucleus and cytoplasmic foci and binds both DNA and RNA. Plant Physiol **152:** 151-165

Prouteau M, Daugeron MC, Seraphin B (2008) Regulation of ARE transcript 3' end processing by the yeast Cth2 mRNA decay factor. EMBO J **27:** 2966-2976

Saleh MI, Meullenet J-F (2007) Effect of Protein Disruption Using Proteolytic Treatment on Cooked Rice Texture propertied. Journal of Texture Studies **38:** 423-437

Scarpella E, Rueb S, Boot KJ, Hoge JHC, Meijer AH (2000) A role for the rice homeobox gene Oshox1 in provascular cell fate commitment. Development **127:** 3655-3669

Schichl YM, Resch U, Hofer-Warbinek R, de Martin R (2009) Tristetraprolin impairs NF-kappaB/p65 nuclear translocation. J Biol Chem **284:** 29571-29581

Schmidt RJ, Burr FA, Aukerman MJ, Burr B (1990) Maize regulatory gene opaque-2 encodes a protein with a "leucine-zipper" motif that binds to zein DNA. Proc Natl Acad Sci USA **87:** 46-50

Schmidt RJ, Ketudat M, Aukerman MJ, Hoschek G (1992) Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes. Plant Cell **4:** 689-700

Singh G, Kumar S, Singh P (2003) A quick method to isolate RNA from wheat and other carbohydrate-rich seeds. Plant Mol Biol Rep **21:** 93-93

Sinha S, Dutta S, Datta K, Ghosh AK, Mukhopadhyay D (2009) Von Hippel-Lindau Gene Product Modulates TIS11B Expression in Renal Cell Carcinoma: Impact on Vascular Endothelial Growth Factor Expression in Hypoxia. J Biol Chem **284:** 32610-32618

Stumpo DJ, Broxmeyer HE, Ward T, Cooper S, Hangoc G, Chung YJ, Shelley WC, Richfield EK, Ray MK, Yoder MC, Aplan PD, Blackshear PJ (2009) Targeted disruption of Zfp36l2, encoding a CCCH tandem zinc finger RNA-binding protein, results in defective hematopoiesis. Blood **114:** 2401-2410

Suzuki A, Wu CY, Washida H, Takaiwa F (1998) Rice MYB protein OSMYB5 specifically binds to the AACA motif conserved among promoters of genes for storage protein glutelin. Plant Cell Physiol **39:** 555-559

Takaiwa F (1987) Gene for a rice storage protein glutelin. Tanpakushitsu Kakusan Koso**:** 247-260

Takaiwa F, Oono K (1990) Interaction of an immature seed-specific trans-acting factor with the 5' upstream

region of a rice glutelin gene. Mol Gen Genet **224:** 289-293

Takaiwa F, Oono K (1991) Genomic DNA sequences of two new genes for new storage protein glutelin in rice. Jpn J Genet **66:** 161-171

Takaiwa F, Oono K, Kato A (1991) Analysis of the 5' flanking region responsible for the endosperm-specific expression of a rice glutelin chimeric gene in transgenic tobacco. Plant Mol Biol **16:** 49-58

Takaiwa F, Oono K, Wing D, Kato A (1991) Sequence of three members and expression of a new major subfamily of glutelin genes from rice. Plant Mol Biol **17:** 875-885

Takaiwa F, Yamanouchi U, Yoshihara T, Washida H, Tanabe F, Kato A, Yamada K (1996) Characterization of common cis-regulatory elements responsible for the endosperm-specific expression of members of the rice glutelin multigene family. Plant Mol Biol **30:** 1207-1221

Topfer R, Matzeit V, Gronenborn B, Schell J, Steinbiss HH (1987) A set of plant expression vectors for transcriptional and translational fusions. Nucleic Acids Res **15:** 5890

Tran TL, Castagne N, Dubosclard V, Noinville S, Koch E, Moudjou M, Henry C, Bernard J, Yeo RP, Eleouet JF (2009) The respiratory syncytial virus M2-1 protein forms tetramers and interacts with RNA and P in a competitive manner. J Virol **83:** 6363-6374

Ueda T, Waverczak W, Ward K, Sher N, Ketudat M, Schmidt RJ, Messing J (1992) Mutations of the 22- and 27-kD zein promoters affect transactivation by the Opaque-2 protein. Plant Cell **4:** 701-709

Vicente-Carbajosa J, Moose SP, Parsons RL, Schmidt RJ (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc Natl Acad Sci USA **94:** 7685-7690

Villareal RM, Juliano BO (1978) Properties of glutelin from mature and developing rice grain. Phytochemistry **17:** 177-182

Viphakone N, Voisinet-Hakil F, Minvielle-Sebastia L (2008) Molecular dissection of mRNA poly(A) tail length control in yeast. Nucleic Acids Res **36:** 2418-2433

Wang D, Guo Y, Wu C, Yang G, Li Y, Zheng C (2008) Genome-wide analysis of CCCH zinc finger family in Arabidopsis and rice. BMC Genomics **9:** 44

Wang L, Xu Y, Zhang C, Ma Q, Joo SH, Kim SK, Xu Z, Chong K (2008) OsLIC, a Novel CCCH-Type Zinc Finger Protein with Transcription Activation, Mediates Rice Architecture via Brassinosteroids Signaling. PLoS One **3:** e3521

Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, Takaiwa F (1999) Identification of cis-regulatory elements required for endosperm expression of the rice storage protein glutelin gene GluB-1. Plant Mol Biol **40:** 1-12

Wesley SV, Helliwell CA, Smith NA, Wang M, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. Plant J **27:** 581-590

Wu C-Y, Adach T, Hatano T, Washida H, Suzuki A, Takaiwa F (1998) Promoters of Rice Seed Storage Protein Genes Direct Endosperm-Specific Gene Expression in Transgenic Rice. Plant Cell Physiol **39:** 885-889

Wu C, Washida H, Onodera Y, Harada K, Takaiwa F (2000) Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. Plant J **23:** 415-421

Wu CY, Suzuki A, Washida H, Takaiwa F (1998) The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. Plant J **14:** 673-683

Yamagata H, Sugimoto T, Tanaka K, Kasai Z (1982) Biosynthesis of Storage Proteins in Developing Rice Seeds. Plant Physiol **70:** 1094-1100

Yamamoto MP, Onodera Y, Touno SM, Takaiwa F (2006) Synergism between RPBF Dof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. Plant Physiol **141:** 1694-1707

Yoshihara T, Takaiwa F (1996) cis-regulatory elements responsible for quantitative regulation of the rice seed storage protein glutelin GluA-3 gene. Plant Cell Physiol **37:** 107-111

Yoshihara T, Washida H, Takaiwa F (1996) A 45-bp proximal region containing AACA and GCN4 motif is sufficient to confer endosperm-specific expression of the rice storage protein glutelin gene, GluA-3. FEBS Lett **383:** 213-218

Zheng Z, Kawagoe Y, Xiao S, Li Z, Okita T, Hau TL, Lin A, Murai N (1993) 5' distal and proximal cis-acting regulator elements are required for developmental control of a rice seed storage protein glutelin gene. Plant J **4:** 357-366

Supplemental Table 1 Primers used in PCR

Code	Sequence
PGBL-for	5'-GCGGCCGCTTAAAGTTAGACTCATCTTCTCAAGCATAAGAG-3'
PGBL-rev	5'-ACTAGTCAATAAAAAAGCATCTTGGGATACATGG-3'
PGBM1-for	5'-GCGGCCGCTACGCAAAATGACAACATGC-3'
PGBM1-rev	5'-ACTAGTGAACTATACATTTCATGACTCAGCC-3'
PGBM2-for	5'-GCGGCCGCGGCTGAGTCATGAAATGTATAGTTC-3'
PGBM2-rev	5'-ACTAGTGAGTTCAAAGACAGACCAAGC-3'
COPCR-for	5'-CCCCACCAAACCCAAAAAAAG-3'
COPCR-rev	5'-GTTGAAGTGAACTTGCG-3 '
ZF2PAS-for	5'-CCATGGCAATGTGGCAGATGAATCTAGG-3'
ZF2PAS-rev	5'-GCGAGAATAGAATGTGCATATGGGC-3'
ZF2GST-for	5'-CCATGGCAATGTGGCAGATGAATCTAGG-3'
ZF2GST-rev	5'-GCGAGAATAGAATGTGCATATGGGC-3'
GBNOR-for	5'-GGCATCAACGCAGTAGCAG-3'
GBNOR-rev	5'-GGGGCTGACGAAGGTGTTAG-3'
ZF1SN-for	5'-CTCGAGTCTAGACTCACATCCTTCATCATGCAT-3'
ZF1SN-rev	5'-AAGCTTGAATTCTGATCAGCATCAGGAGTCCG-3'
ZF1ASN-for	5'-GGATCCCTCACATCCTTCATCATGCAT-3'
ZF1ASN-rev	5'-ATCGATTGATCAGCATCAGGAGTCCG-3'
PROZF1-for	5'-CCAAGCTTGCCACACAAAGTTTTATGCTCCC-3'
PROZF1-rev	5'-CCACTTGGTCGGGTCAATGGTTAC-3'
PROZF2-for	5'-GCAAGCTTAGAAAGGAGTTACCACCAAGG-3'
PROZF2-rev	5'-CATGCCATGGGCCTCTGCAATTTAATTATAAC-3'
ZF1STASN-for	5'-TAATACGACTCACTATAGGGGCGTGATCAGCTTGATCTAA-3'
ZF1STSN-rev	5'-TAATACGACTCACTATAGGGTCATCTCCTTCAAATGCATC-3'
ZF1STSN-for	5'-GCGTGATCAGCTTGATCTAA-3'
ZF1STASN-rev	5'-TCATCTCCTTCAAATGCATC-3'
ZF2STASN-for	5'-TAATACGACTCACTATAGGGCTTCCTTCCTCCCTCTTCCGCCTC-3'
ZF2STSN-rev	5'-TAATACGACTCACTATAGGGTCGACCACGTGGGGCCCGAT-3'
ZF2STSN-for	5'-CTTCCTTCCTCCCTCTTCCGCCTC-3'
ZF2STASN-rev	5'-TCGACCACGTGGGGCCCGAT-3'
ZF1PR-for	5'-TGCTACAACAAGTTCAGAT-3'
ZF1PR-rev	5'-GGGGTACCGTCAACTTTTAAAGCATCCC-3'
ZF2PR-for	5'-GTTCAGATACGCTAGCTTGG-3'
ZF2PR-rev	5'-GGGGTACCTATTTCCGCATCCTCAACAT-3'
RISBZ1PR-for	5'-CCATGGGTTGCGTAGCCGTA-3'
RISBZ1PR-rev	5'-GGGGTACCCCACCATGTCCCATTGTCCC-3'

Supplemental Table 2 Oligo sequences used as baits in yeast one-hybrid

Bait	Code	Sequence
ProGluB-1-S1	S1a top	5'-GGCCGCAGACATTTCTTTAACAAACTCCATTTGTATTACTCCAAAAGCACCAGAAGTTTGTT-3'
	S1a down	5'-CTAGAACAAACTTCTGGTGCTTTTGGAGTAATACAAATGGAGTTTGTTAAAGAAATGTCTGC-3'
ProGluB-1-S2	S2a top	5'-GGCCGCAGAAGTTTGTCATGGCTGAGTCATGAAATGTATAGTTCAATCTTGCAAAGTTGCCTTTCCT-3'
	S2a down	5'-CTAGAGGAAAGGCAACTTTGCAAGATTGAACTATACATTTCATGACTCAGCCATGACAAACTTCTGC-3'
ProGluB-1-S3	S3a top	5'-GGCCGCTTGCCTTTCCTTTTGTACTGTGTTTTAACACTACAAGCCATATATTGTCTT-3'
	S3a down	5'-CTAGAAGACAATATATGGCTTGTAGTGTTAAAACACAGTACAAAAGGAAAGGCAAGC-3'
ProGluB-1-S4	S4a top	5'-GGCCGCTGTCTGTACGTGCAACAAACTATATCACCATGTATCCCAAGATGCTTTTTTATTGT-3'
	S4a down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATAGTTTGTTGCACGTACAGACAGC-3'

Supplemental Table 3 Sequences of mutated oligos used in EMSA

Code	Sequence
S1*b* top	5'-GGCCGCAGACATTTCTTTATGTTTCTCCATTTCGTATTACACCAAAAGCACCAGAAGTTTGTT-3'
S1*b* down	5'-CTAGAACAAACTTCTGGTGCTTTTGGTGTAATACGAAATGGAGAAACATAAAGAAATGTCTGC-3'
S1*c* top	5'-GGCCGCAGACATTTCTTTTAGATAGTCCATTTCGTATTACACCAAAAGCACCAGAAGTTTGTT-3'
S1*c* down	5'-CTAGAACAAACTTCTGGTGCTTTTGGTGTAATACGAAATGGACTATCTAAAAGAAATGTCTGC-3'
S2*b* top	5'-GGCCGCAGAAGTTTGTCATGGCTCTCAGATGAAATGTATAGTTCAATCTTGCAAAGTTGCCTTTCCT-3'
S2*b* down	5'-CTAGAGGAAAGGCAACTTTGCAAGATTGAACTATACATTTCATCTGAGAGCCATGACAAACTTCTGC-3'
S2*c* top	5'-GGCCGCAGAAGTTTGTCATGGCTCACTGATGAAATGTATAGTTCAATCTTGCAAAGTTGCCTTTCCT-3'
S2*c* down	5'-CTAGAGGAAAGGCAACTTTGCAAGATTGAACTATACATTTCATCAGTGAGCCATGACAAACTTCTGC-3'
S2*d* top	5'-GGCCGCAGAAGTTTGTCATGGCTGAGTCATGAAATGTATAGTTCAATCTTCGTTTGTTGCCTTTCCT-3'
S2*d* down	5'-CTAGAGGAAAGGCAACAAACGAAGATTGAACTATACATTTCATGACTCAGCCATGACAAACTTCTGC-3'
S2*e* top	5'-GGCCGCAGAAGTTTGTCATGGCTGAGTCATGAAATGTATAGTTCAATCTTCCTATGTTGCCTTTCCT-3'
S2*e* down	5'-CTAGAGGAAAGGCAACATAGGAAGATTGAACTATACATTTCATGACTCAGCCATGACAAACTTCTGC-3'
S2*f* top	5'-GGCCGCAGAAGTTTGTCATGGCTCACTGATGAAATGTATAGTTCAATCTTCCTATGTTGCCTTTCCT-3'
S2*f* down	5'-CTAGAGGAAAGGCAACATAGGAAGATTGAACTATACATTTCATCAGTGAGCCATGACAAACTTCTGC-3'
S4*b* top	5'-GGCCGCTGTCTGTACGTGCATGTTTCTATATCACCATGTATCCCAAGATGCTTTTTTATTGT-3'
S4*b* down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATAGAAACATGCACGTACAGACAGC-3'

S4*c* top	5'-GGCCGCTGTCTGTACGTGCTAGATAGTATATCACCATGTATCCCAAGAT GCTTTTTTATTGT-3'
S4*c* down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATACTATCTAGC ACGTACAGACAGC-3'
S4*d* top	5'-GGCCGCTGTCTGATGCACCAACAAACTATATCACCATGTATCCCAAGAT GCTTTTTTATTGT-3'
S4*d* down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATAGTTTGTTGG TGCATCAGACAGC-3'
S4*e* top	5'-GGCCGCTGTCTGAAGGAGGAACAAACTATATCACCATGTATCCCAAGAT GCTTTTTTATTGT-3'
S4*e* down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATAGTTTGTTCC TCCTTCAGACAGC-3'
S4*f* top	5'-GGCCGCTGTCTGAAGGAGGTAGATAGTATATCACCATGTATCCCAAGAT GCTTTTTTATTGT-3'
S4*f* down	5'-CTAGACAATAAAAAAGCATCTTGGGATACATGGTGATATACTATCTACC TCCTTCAGACAGC-3'