Plant Biotechnology Journal (2004) **2**, pp. 113–125 Doi: 10.1046/j.1467-7652.2003.00055.x

Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice

Le Qing Qu and Fumio Takaiwa*

Department of Plant Biotechnology, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

Received 2 September 2003;
revised 14 October 2003;
accepted 15 October 2003.
**Correspondence* (fax 81 29 838 8373;
e-mail takaiwa@nias.afffrc.go.jp)

Summary

Using stable transgenic rice plants, the promoters of 15 genes expressed in rice seed were analysed for their spatial and temporal expression pattern and their potential to promote the expression of recombinant proteins in seeds. The 15 genes included 10 seed storage protein genes and five genes for enzymes involved in carbohydrate and nitrogen metabolism. The promoters for the glutelins and the 13 kDa and 16 kDa prolamins directed endosperm-specific expression, especially in the outer portion (peripheral region) of the endosperm, whilst the embryo globulin and 18 kDa oleosin promoters directed expression in the embryo and aleurone layer. Fusion of the GUS gene to the 26 kDa globulin promoter resulted in expression in the inner starchy endosperm tissue. It should be noted that the 10 kDa prolamin gene was the only one tested that required both the 5′ and 3′ flanking regions for intrinsic endosperm-specific expression. The promoters from the pyruvate orthophosphate dikinase (PPDK) and ADP-glucose pyrophosphorylase (AGPase) small subunit genes were active not only in the seed, but also in the phloem of vegetative tissues. Within the seed, the expression from these two promoters differed in that the PPDK gene was only expressed in the endosperm, whereas the AGPase small subunit gene was expressed throughout the seed. The GUS reporter gene fused to the alanine aminotransferase (AlaAT) promoter was expressed in the inner portion of the starchy endosperm, whilst the starch branching enzyme (SBE1) and the glutamate synthase (GOGAT) genes were mainly expressed in the scutellum (between the endosperm and embryo). When promoter activities were examined during seed maturation, the glutelin *GluB-4*, 26 kDa globulin and 10 kDa and 16 kDa prolamin promoters exhibited much higher activities than the others. The seed promoters analysed here exhibited a wide variety of activities and expression patterns, thus providing many choices suitable for various applications in plant biotechnology.

Keywords: embryo, endosperm, gene expression, promoter activity, specific expression.

Introduction

Plant biotechnology has been applied not only to the improvement of agronomic properties, such as pest/pathogen resistance, abiotic stress resistance, yield and quality, but also to the production of pharmaceutical and industrial materials. Several recombinant industrial enzymes and pharmaceuticals have been commercialized already, and the number is expected to increase in the near future (Hood and Jilka, 1999; Hood and Woodard, 2002). Plant production of recombinant proteins has many advantages, the most prominent of which are cost savings relative to animal transgenic systems, easy control of production scale in response to market size and freedom from animal-derived pathogens, including viruses and prions (Fischer and Emans, 2000; Giddings *et al*., 2000; Daniell *et al*., 2001).

For the production of recombinant protein in plants, seed systems have recently been proven to be more advantageous than systems using leaf or root (Delaney, 2002; Howard and Hood, 2002). The seed is a storage organ within which a small number of storage proteins are highly and stably accumulated in specific organelles designated as protein bodies. Taking advantage of these properties, the seed has been utilized as an ideal bio-reactor for the production of recombinant

proteins. Recombinant proteins accumulated in seed are highly stable, and do not require processing or purification, allowing direct oral delivery. When antibody or vaccine is expressed in seed, it has been reported to be highly stable, with no degradation for years even if stored at room temperature. Furthermore, vaccine delivered via seed is expected to induce antibody production through the mucosal immune system without purification and processing (Walmsley and Arntzen, 2000).

Rice is one of the most important crops and food resources in the world. Conventional breeding has long been employed to improve seed yield and quality. As a production system, rice seed has several advantages over other cereal crops, such as wheat, barley and rye, including easier storage and processing, greater biomass (yield per unit area) and lower producer costs (Stoger *et al*., 2002). In addition, transformation systems have already been established for rice (Hiei *et al*., 1998), and whole genome sequence information is now available (Goff *et al*., 2002).

For rice seed to be utilized as a platform for the production of recombinant proteins, it is important to use promoters tailored to the requirements of each particular protein and biotechnology application, because the promoter controls not only the time and place but also the level of expression. However, analyses of the *cis*-regulatory elements involved in endosperm-specific expression have been limited to a few glutelin genes, using heterologous transgenic tobacco and homologous transgenic rice (Takaiwa *et al*., 1991a, 1996; Zheng *et al*., 1993; Zhao *et al*., 1994; Croissant-Sych and Okita, 1996; Yoshihara *et al*., 1996; Wu *et al*., 1998a, 2000). For a few other rice storage protein promoters, only the spatial expression patterns have been observed (Wu *et al*., 1998b).

In this study, the tissue-specific expression of 10 seed storage protein genes and five genes for enzymes involved in carbohydrate and nitrogen metabolism was examined by introducing them into rice via *Agrobacterium*-mediated transformation. These genes included: three glutelin genes, *GluB-1*, *GluB-2* and *GluB-4* (Masumura *et al*., 1989a; Takaiwa *et al*., 1991b); three prolamin genes, 10 kDa, 13 kDa (*PG5a*) and 16 kDa (Masumura *et al*., 1989b; Nakase *et al*., 1996a; Mitsukawa *et al*., 1999); a 26 kDa globulin gene (*Glb-1*) (Nakase *et al*., 1996b); a rice gene encoding the small subunit of ADP-glucose pyrophosphorylase (AGPase) (Anderson *et al*., 1991); a rice pyruvate orthophosphate dikinase gene (PPDK) (Imaizumi *et al*., 1997); a rice alanine aminotransferase gene (AlaAT) (Kikuchi *et al*., 1999); a rice embryo globulin gene (REG-2) (Sun *et al*., 1996); a rice 18 kDa oleosin gene (Ole18) (Wu *et al*., 1998c); a rice glutamate synthase gene (GOGAT) (Goto *et al*., 1998); and a rice starch branching enzyme gene (SBE1) (Kawasaki *et al*., 1993). The soybean β-conglycinin gene (Doyle *et al*., 1986) and maize ubiquitin gene promoters (Christensen *et al*., 1992) were introduced as controls, representing a typical seed protein gene and constitutive promoter, respectively. The nucleotide sequences of these 15 promoters will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases with accession numbers AY427562 (REG-2), AY427563 (Ole18), AY427564 (AlaAT), AY427565 (GOGAT), AY427566 (AGPase small subunit), AY427567 (PPDK), AY427568 (SBE1), AY427569 (*GluB-1*), AY427570 (*GluB-2*), AY427571 (*GluB-4*), AY427572 (10 kDa prolamin), AY427573 (13 kDa prolamin), AY427574 (16 kDa prolamin), AY427575 (*Glb-1*) and AB008679 (soybean β-conglycinin). The major objective of this investigation was to characterize promoters that have the potential to drive high expression or specific spatial and temporal patterns of expression during rice seed maturation, in order to prepare a promoter set useful for the production of high value recombinant proteins in rice.

Results

Construction of promoter-GUS chimeric genes and isolation of transgenic plants

The aim of this study was not to investigate the putative regulatory elements, but to characterize the expression pattern and promoter activities of several genes expressed in seed. Fifteen promoters with sizes ranging from 0.8 to 2.4 kb were isolated by polymerase chain reaction (PCR) using genomic DNA or genomic clones as templates. The isolated promoters were linked to the GUS reporter gene (Figure 1) on the binary vector pGPTV-35S-HPT to make transcriptional fusions (Jefferson, 1987), which were subsequently introduced into rice (*Oryza sativa* L. cv. Kitaake) via *Agrobacterium tumefaciens*-mediated transformation. The presence of the desired promoter fusion was confirmed by PCR using genomic DNA isolated from leaves of independent transgenic rice lines, and positive lines (6 to 21) were subjected to analysis of the promoter properties.

Promoters of seed storage protein genes direct GUS expression in endosperm or embryo

To determine the sites of GUS reporter gene expression directed by the seed storage protein promoters, seeds of transgenic rice were examined by histochemical staining. Seeds collected at the stage of 17 days after flowering (DAF) from transgenic plants were longitudinally sectioned and stained with X-Gluc. The expression patterns are shown in

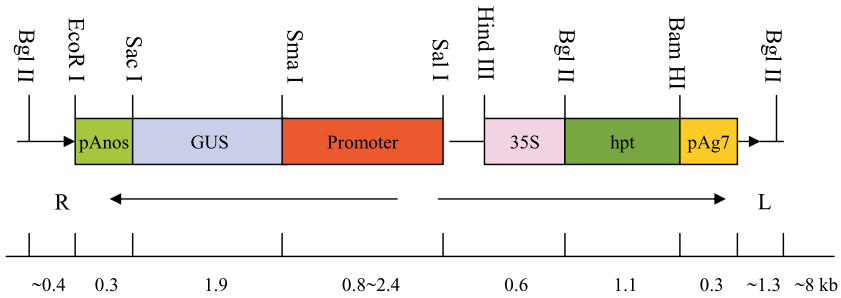

Figure 1 Schematic diagram of the chimeric gene construction for rice transformation. The 5′ flanking regions of various rice seed storage and non-storage protein genes were fused between two of the three restriction sites, *Hin*dIII, *Sal*I and *Sma*I, followed by the GUS reporter gene and the Nos terminator. The promoter refers to that of the 1.3 kb *GluB-1*, 2.3 kb *GluB-1*, *GluB-2*, *GluB-4*, 10 kDa prolamin, 13 kDa prolamin (*PG5a*), 16 kDa prolamin, 26 kDa *Glb-1*, REG-2, Ole18, soybean β-conglycinin, AlaAT, GOGAT, PPDK, AGPase small subunit and SBE1 genes.

Figure 2. The rice glutelin promoters (1.3 kb and 2.3 kb *GluB-1*, *GluB-2* and *GluB-4*; Figure 2a–d) and prolamin promoters (10 kDa, 13 kDa and 16 kDa; Figure 2e–g) directed the GUS gene to express in endosperm. GUS expression from glutelin and prolamin promoters was observed in the aleurone and subaleurone, but not in the embryo. Closer inspection of maturing seed of transgenic rice containing the glutelin and prolamin promoters revealed that GUS activity was strong in the outer portion and weak in the inner portion of the endosperm. The *GluB-1* promoters (both 1.3 kb and 2.3 kb) showed much higher activity in the region of the endosperm close to the embryo. The GUS expression directed by the 13 kDa prolamin promoter (*PG5a*) was strictly limited to the outer portion of the endosperm. The 26 kDa globulin *Glb-1* promoter directed the GUS expression in inner starchy endosperm tissue (Figure 2h). The GUS expression driven by the embryo storage protein promoters (REG-2, Ole18 and β-conglycinin; Figure 2i–k) was restricted to the embryo and aleurone with no expression in the endosperm. There was little difference among the GUS expression patterns mediated by these embryo storage protein promoters. It is interesting to note that the β-conglycinin promoter, from the dicot soybean, retained embryo-specific expression in the monocot rice, even though there have been numerous reports of differences in expression between dicotyledonous and monocotyledonous plants (Rathaus *et al*., 1993; Chowdhury *et al*., 1997). It is notable that GUS expression directed by the β-conglycinin promoter was very weak in rice, in marked contrast with the high expression observed in embryo and cotyledon when the same promoter was introduced into the dicot tobacco.

Generally, no GUS activity was detected in leaf, leaf sheath, stalk or roots of transgenic rice carrying fusions to seed storage protein promoters (data not shown). The one exception was the 10 kDa prolamin promoter, which did show some expression in vegetative organs (Figure 3). These results support the conclusion that the endosperm storage protein genes (except for 10 kDa prolamin) are expressed in an endosperm-specific manner and that expression of the embryo storage protein genes is restricted to the embryo and aleurone layer.

Although seed storage protein promoters showed either endosperm- or embryo-specific expression, non-storage protein promoters exhibited different expression patterns (Figure 2). The GUS gene directed by the AlaAT promoter was expressed in the central part of the starchy endosperm with stronger activity in the endosperm near the embryo (Figure 2l). The expression pattern of the PPDK-GUS transgene was similar to that of the endosperm storage protein (Figure 2o). The GUS gene directed by the AGPase small subunit promoter was expressed throughout the seed, including the pericarp, but especially in the inner starchy endosperm and the embryo (Figure 2n). In contrast, the GOGAT and SBE promoters directed GUS gene expression mainly in the scutellum (at the boundary between the embryo and endosperm) (Figure 2m,p).

Patterns of GUS expression in vegetative organs

Most promoters tested here exhibited endosperm- or embryo-specific expression of the GUS gene. However, GUS activity was also detected in vegetative tissues in the transgenic rice carrying the 10 kDa prolamin, PPDK and AGPase small subunit promoters (Figure 3), as well as the AlaAT promoter (Kikuchi *et al*., 1999). In these transgenic rice plants, GUS activity was detected in the phloem of large vascular bundles in leaf, leaf sheath and stalk, in addition to the endosperm or the whole seed (Figure 3a–c). GUS activity was also detected in the root endodermis in these transgenic rice plants. However, the expression pattern driven by the AGPase small subunit promoter differed slightly from that of the PPDK and 10 kDa prolamin promoters; specifically, the former promoter

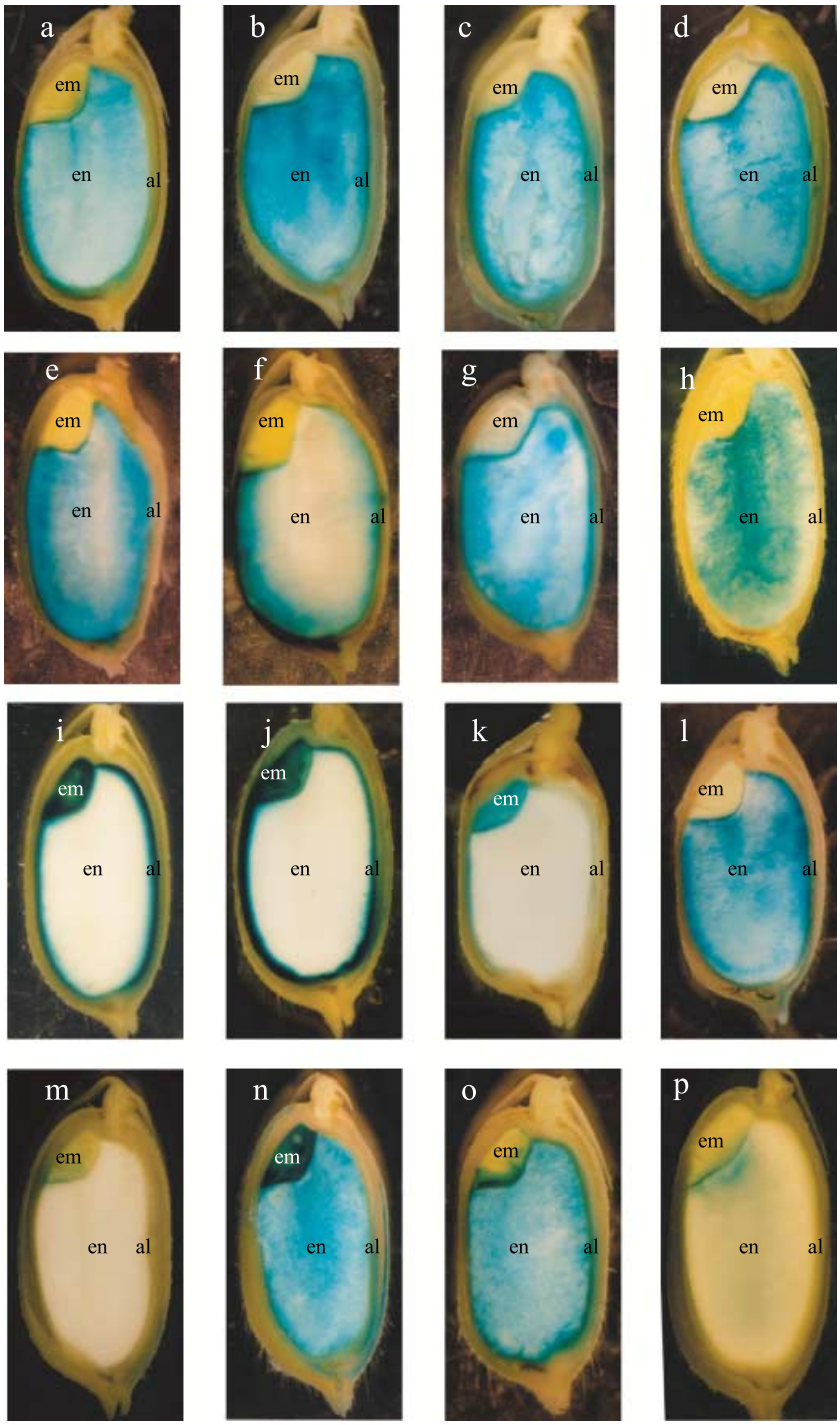

Figure 2 Histochemical analysis of GUS expression directed by various seed storage and non-storage protein gene promoters. GUS protein was detected by incubating hand-cut longitudinal sections of transgenic seeds in X-Gluc solution. (a) 1.3 kb *GluB-1* promoter; (b) 2.3 kb *GluB-1* promoter; (c) *GluB-2* promoter; (d) *GluB-4* promoter; (e) 10 kDa prolamin promoter; (f) 13 kDa prolamin (*PG5a*) promoter; (g) 16 kDa prolamin promoter; (h) 26 kDa *Glb-1* promoter; (i) REG-2 promoter; (j) Ole18 promoter; (k) β-conglycinin promoter; (l) AlaAT promoter; (m) GOGAT promoter; (n) AGPase small subunit promoter; (o) PPDK promoter; (p) SBE1 promoter. al, aleurone; em, embryo; en, endosperm.

drove strong GUS expression in the apical meristem, whereas the latter promoters resulted in root tissue that was uniformly stained. Furthermore, the GUS activities in the root differed, i.e. the AGPase small subunit promoter was stronger than the PPDK and 10 kDa prolamin promoters.

The intact 10 kDa prolamin gene is normally expressed in maturing endosperm and not detected in vegetative tissues. The ectopic expression of the GUS fusion observed here could be reverted back to the normal endosperm-specific expression by replacing the Nos terminator with the 0.3 kb region downstream of the stop codon of the native 10 kDa prolamin gene (data not shown). It should be noted that this replacement of the 3′ terminator region had little effect on promoter activities. These results indicate that the endosperm-specific expression of the 10 kDa prolamin gene requires both the 5′ and 3′ flanking regions.

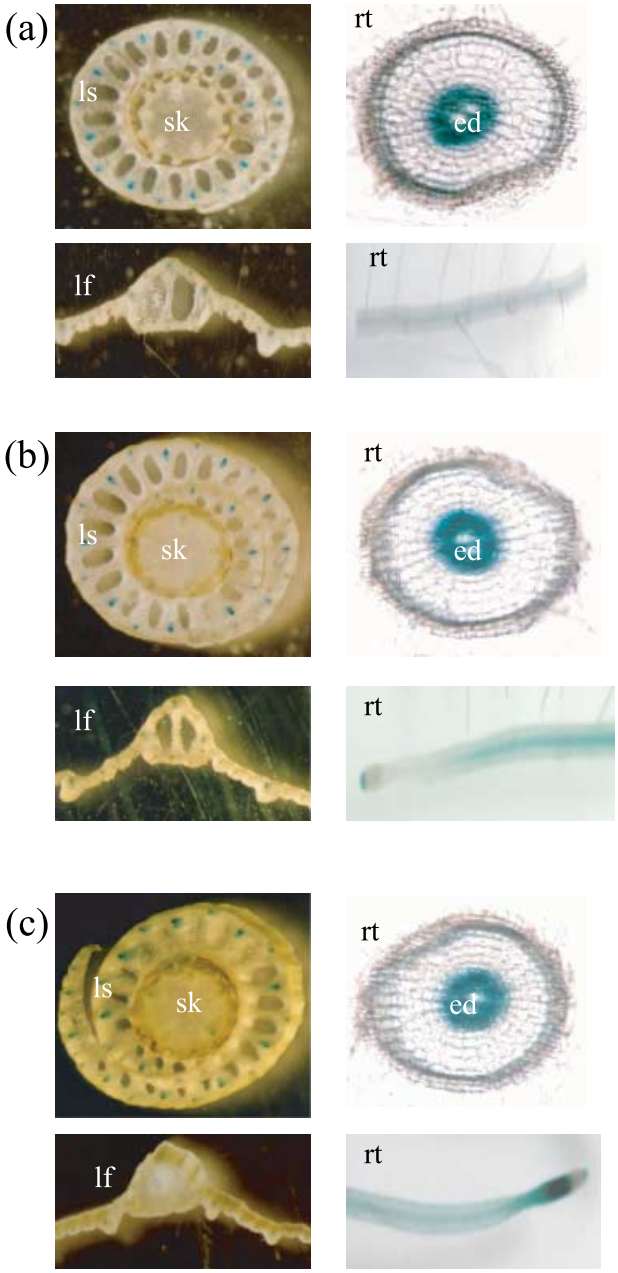

Figure 3 Histochemical GUS assay in vegetative tissues. ed, eudodersis; lf, leaf; ls, leaf sheath; rt, root; sk, stalk. (a) 10 kDa prolamin promoter; (b) PPDK promoter; (c) AGPase small subunit promoter.

Promoter activities during seed development

The distribution of transgene expression within developing seed was examined by histological staining of longitudinal sections of seed harvested at the developing stages of 7, 12 and 17 DAF. In Kitaake, an early heading variety, morphogenesis of the embryo and cellularization of the endosperm finish before 5 DAF; the endosperm cells then start to accumulate storage reserves (starch and storage proteins). The endosperm cells gradually expand by deposition of storage reserves and reach a maximum size at 12 DAF (mid-maturation stage); the fresh weight of seed reaches a maximum level at 17 DAF (late maturation stage). Seed completely maturates and desiccates around 23 DAF.

The expression pattern during seed maturation was examined in every transgenic line, and the results for one representative line for each seed promoter are shown in Figure 4. It is interesting to note that the site of the first detectable GUS expression was different for each construct. In the case of the glutelin and prolamin promoters, blue GUS staining was first observed in the peripheral region of the endosperm, i.e. in the aleurone and subaleurone. For the glutelin and 16 kDa prolamin promoters, but not for the 10 kDa and 13 kDa prolamin promoters, the staining then spread to the inner starchy endosperm as the seed matured (17 DAF) (Figure 4a–g). This expression pattern was in remarkable contrast with that observed in the 26 kDa *Glb-1* promoter, in which blue GUS staining was first detected in inner starch endosperm cells close to the embryo and the expression pattern was not changed during seed development (Figure 4h).

GUS gene expression directed by the promoters of the REG-2, Ole18 and β-conglycinin genes was observed by 7 DAF. The activity tended to start in the aleurone layer, followed by expression in the embryo. The expression of these promoters was restricted to the aleurone and embryo (Figure 4i–k).

The temporal patterns during seed maturation for transgenic lines representative of non-storage protein promoters are shown in Figure 4(l–p). GUS expression from the AlaAT promoter appeared first in the inner starchy endosperm tissue and eventually spread throughout the whole endosperm, but the embryos remained unstained (Figure 4l). The GUS activity driven by the SBE1 promoter was also limited to the inner starchy endosperm tissue, especially to the tissue near the embryo (Figure 4p). However, because of the very low GUS activity, blue staining could not be detected until 12 DAF. In contrast, when the AGPase small subunit gene promoter fusion was introduced, GUS staining initiated in the embryo and then spread to the central portion of the endosperm. Blue GUS staining was finally observed in the whole seed tissues towards maturation, with stronger staining in the embryo (Figure 4n). This expression profile during seed development was very similar to that observed for the ubiquitin promoter (Figure 4q). On the other hand, the PPDK promoter exhibited an expression pattern similar to that of the glutelin and prolamin gene promoters (Figure 4o). GUS activity directed by the GOGAT promoter was restricted to the scutellum with little change during seed development,

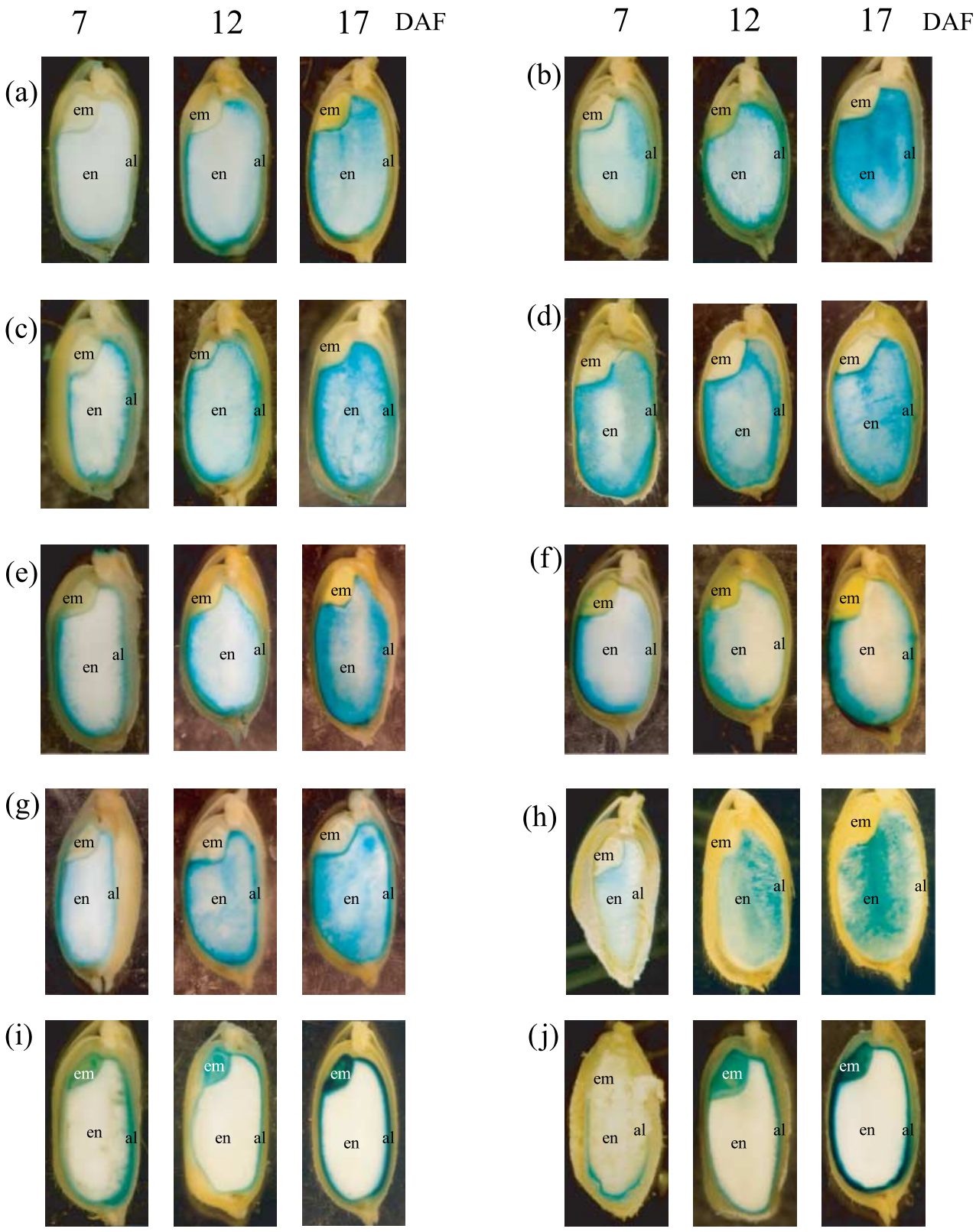

Figure 4 Developmental time course of GUS activity directed by seed promoters during seed maturation. The histochemical X-Gluc staining of longitudinal sections of transgenic rice seeds at 7, 12 and 17 days after flowering (DAF). (a) 1.3 kb *GluB-1* promoter; (b) 2.3 kb *GluB-1* promoter; (c) *GluB-2* promoter; (d) *GluB-4* promoter; (e) 10 kDa prolamin promoter; (f) 13 kDa prolamin promoter (*PG5a*); (g) 16 kDa prolamin promoter; (h) 26 kDa *Glb-1* promoter; (i) REG-2 promoter; (j) Ole18 promoter; (k) β-conglycinin promoter; (l) AlaAT promoter; (m) GOGAT promoter; (n) AGPase small subunit promoter; (o) PPDK promoter; (p) SBE1 promoter; (q) ubiquitin promoter.

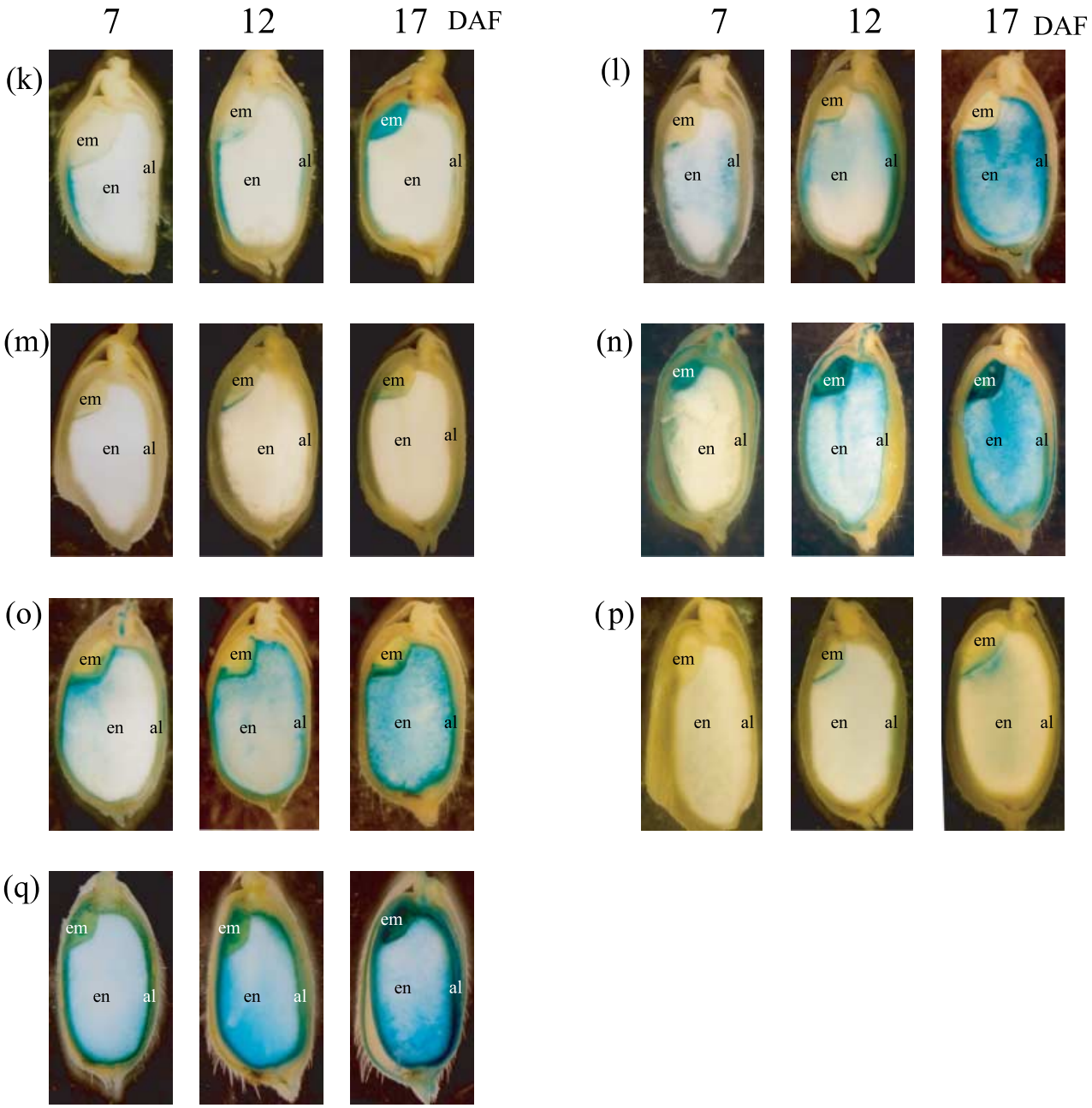

Figure 4 *Continued*.

except that GUS activity could not be detected at 7 DAF (Figure 4m).

Quantitative analysis of promoter strength

To evaluate the potential strength of the various promoters, GUS expression levels were determined in the maturing seed (17 DAF) from independent transgenic plants for each construction. As shown in Figure 5, the promoter activities varied widely. On the basis of the promoter strength, the seed promoters analysed here can be classified into four groups. The high GUS activity group has four members: the *GluB-4*, 10 kDa prolamin, 16 kDa prolamin and *Glb-1* promoters. The average GUS activities for these promoters were 44.8 ± 16.5, 38.8 ± 10.8, 27.1 ± 12.7 and 28.6 ± 11.8 pmol (4 MU/min/μg protein), respectively. The numbers of transgenic rice analysed for these four constructions were 19, 10, 8 and 20 lines, respectively. The medium GUS activity group has two members: the 2.3 kb *GluB-1* and AGPase small subunit gene promoters. Their GUS activities are weaker than those observed

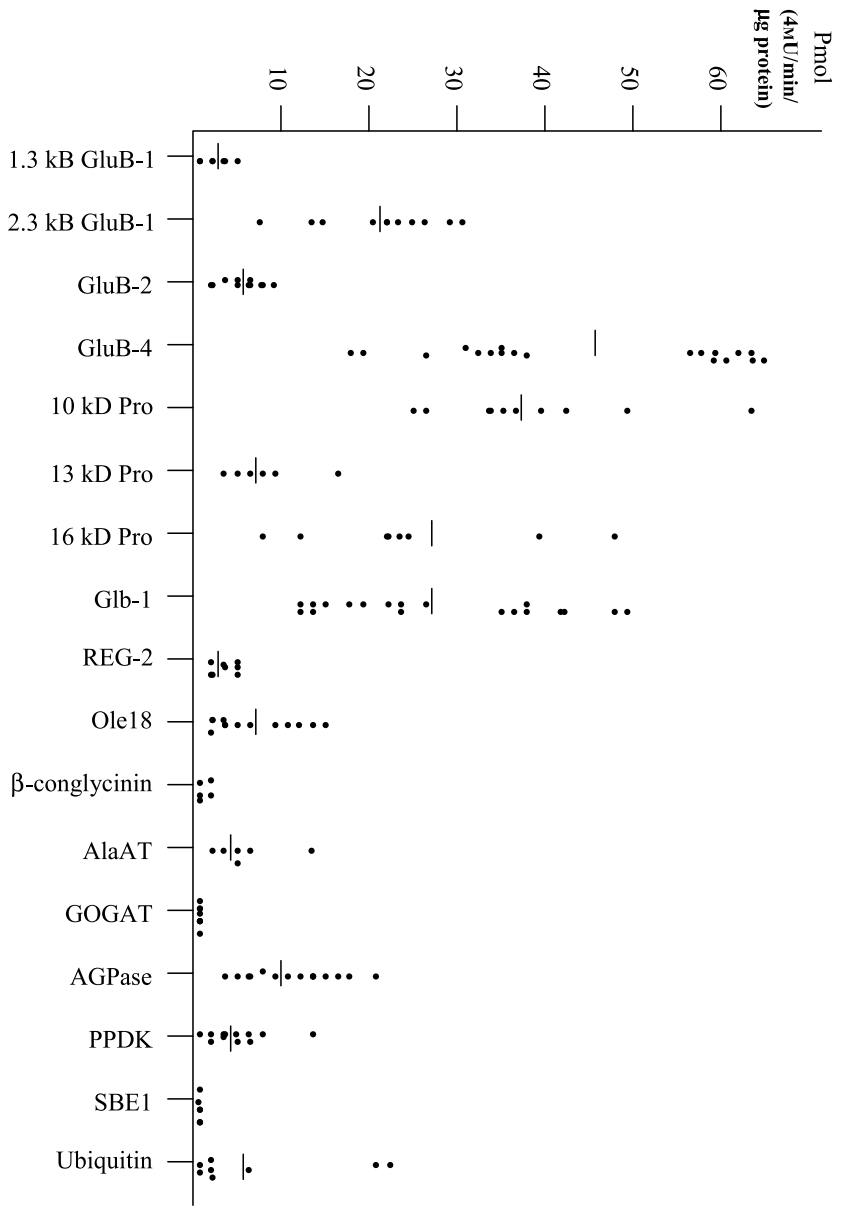

Figure 5 GUS activities expressed by various promoters in maturing seeds at 17 days after flowering (DAF). GUS activity is expressed as pmol (4 mU/min/μg protein).

in the high activity group, but are much stronger than those of the others. The average GUS activities for the 2.3 kb *GluB-1* and AGPase small subunit promoters were 21.3 ± 7.0 and 10 ± 4.7 pmol (4 mU/min/μg protein) for 10 and 18 independent lines, respectively. Seven promoters, namely the 1.3 kb *GluB-1*, *GluB-2*, 13 kDa prolamin, REG-2, Ole18, AlaAT and PPDK promoters, were tentatively classified into the relatively low GUS activity group. The average GUS activities for these promoters were 2.1 ± 1.2, 5.5 ± 2.2, 7.4 ± 5.5, 2.4 ± 1.2, 2 ± 4.6, 5.9 ± 4.0 and 4.0 ± 3.0 pmol (4 mU/min/μg protein), respectively. The GUS expression levels of these seven promoters were obtained by investigating 9, 12, 6, 12, 14, 6 and 21 independent lines, respectively. The remaining three promoters (the GOGAT, SBE1 and β-conglycinin gene promoters) were grouped into the low GUS activity group. These promoters directed very weak GUS expression with activities less than 1 pmol (4 mU/min/μg protein), when the average GUS activities were examined with 18, 16 and 11 independent lines for the corresponding constructs. GUS activity expressed from the ubiquitin promoter, which was used as a control, averaged 7.4 ± 8.5 pmol (4 mU/min/μg protein) in maturing seed. This level is similar to that obtained from the relatively low GUS activity group promoters, even though the ubiquitin promoter has been used in many applications as a universal promoter.

For comparison, the activities of the PPDK and AGPase small subunit promoters were also examined in vegetative

tissues. The average GUS activities in leaf, stalk and sheath were 8.7 ± 6.8, 3.7 ± 3.6 and 16.3 ± 13.9 pmol (4 MU/min/μg protein), respectively, for the PPDK promoter, and 12.5 ± 5.0, 40.2 ± 28.5 and 23.2 ± 16.6 pmol (4 MU/min/μg protein), respectively, for the AGPase small subunit promoter. These promoter activities were similar to or higher than those obtained from maturing seeds. On the other hand, although vegetative expression was observed for the 10 kDa prolamin promoter, the GUS activities [leaf, 3.1 ± 1.1; stalk, 6.0 ± 2.9; and sheath, 2.3 ± 1.0 pmol (4 MU/min/μg protein)] were much lower than those observed in maturing seeds. Although the PPDK, AGPase small subunit and 10 kDa prolamin genes were constitutively expressed, the expression levels in different tissues varied depending on the gene.

Discussion

Protein and starch are the major storage reserves in rice seed. It is important to note that all the genes involved in the synthesis of these storage components are mainly expressed in seed. It has been reported that the expression pattern and transcript level are primarily regulated by the individual gene promoter, although the final amount of gene product is determined at both the transcriptional and post-transcriptional levels. Recent advances in plant biotechnology have made it possible not only to improve traditional agricultural products through increased pest/herbicide/disease resistance, yield and nutritional value, but also to use plants as bioreactors for the production of a wide range of recombinant proteins for pharmaceutical and industrial uses. However, a shortage of suitable promoters is a major factor limiting the production of recombinant proteins in cereal seed, not only at the desired levels, but also in the desired spatial and temporal manner. Wu *et al*. (1998b), using transgenic rice lines produced via electroporation, evaluated the endosperm specificity and activity of four kinds of rice seed storage protein gene promoter. As many genes that are highly expressed in maturing rice seed have now been identified by microarray analysis (Zhu *et al*., 2003) and serial analysis of gene expression (SAGE) (Gibbings *et al*., 2003), it is worthwhile extending this investigation to other seed component promoters.

The major objective of this investigation was to use stable transgenic rice lines to characterize the expression pattern and to evaluate the potential promoter activity of several seed genes involved in metabolism and the biosynthesis of seed storage components. Based on their pattern of expression within the seed (Figures 2 and 4), the tested promoters can be classified into four groups: (1) The Outer Endosperm-Specific Group containing eight members: 1.3 kb *GluB-1*, 2.3 kb *GluB-1*, *GluB-2*, *GluB-4*, 10 kDa prolamin, 13 kDa prolamin (*PG5a*), 16 kDa prolamin and PPDK; (2) The Inner Endosperm-Specific Group containing four members: *Glb-1*, AlaAT, GOGAT and SBE1; (3) The Embryo-Specific Group containing three members: REG-2, Ole18 and β-conglycinin; and (4) The Whole Seed Group containing only one member: AGPase small subunit promoter (Figure 2).

The developmental stage at which GUS activity began to be detectable varied among the promoters used here and was positively correlated with promoter activity, i.e. the stronger the promoter, the earlier GUS was expressed. In transgenic rice containing the glutelin, prolamin and PPDK promoter fusions, GUS activity could be detected at 7 DAF, initially in the aleurone, subaleurone and outer portion of the endosperm (Figure 4), and then spreading to all tissues of the endosperm. In the case of the embryo storage protein promoters (REG-2 and Ole18), GUS activity was first detected in the aleurone at 7 DAF, followed by expression in the embryo at later stages of seed development. This temporal pattern of staining, first in the aleurone layer and then in the embryo, did not vary during seed development (Figure 4). A change in the GUS expression pattern, which started at the embryo and then spread to the inner starchy endosperm, was observed for the AGPase small subunit promoter (Figure 4). The temporal expression pattern of the AGPase small subunit promoter is unique among the promoters tested. The AlaAT promoter directed expression in the inner endosperm at the early maturation stage (7 DAF) and the expression level gradually increased towards 17 DAF, resulting in expression throughout the whole endosperm (Figure 4). GUS expression directed by the GOGAT and SBE1 promoters could be detected weakly at mid-development (12 DAF), and was restricted to the endosperm and embryo boundary and to the inner endosperm near the embryo, respectively (Figure 4), reflecting the weakness of the promoters, although the SBE1 enzyme has been reported to be expressed abundantly in the mid-maturation stage of developing seeds (Kawasaki *et al*., 1993).

The GUS activities varied significantly among the promoters. A wide range of activities were also observed among the individual transformants for each construct, reflecting differences in insertion position and gene dosage effects (Hobbs *et al*., 1993). The highest GUS activities were obtained by using the *GluB-4*, 10 kDa prolamin, 16 kDa prolamin and 26 kDa globulin promoters (Figure 5). The *GluB-1*, *GluB-2*, AGPase small subunit, PPDK and AlaAT promoters also directed high GUS activity in the endosperm, but their activities were much lower than those of the four strongest promoters. Nevertheless, it is notable that these promoter

activities, although in the relatively low activity group, are similar to that of the maize ubiquitin promoter, which is often used as a universal promoter. Taken together, the strong endosperm-specific promoters developed here are expected to perform well in the production of recombinant proteins whilst avoiding any detrimental effects on vegetative tissues.

The embryo storage protein (REG-2, Ole18 and β-conglycinin) promoters showed expression strictly limited to the embryo and aleurone layer (Figure 2). This observation is particularly interesting, because the aleurone is a layer of cells that differentiates from the starchy endosperm during seed development. Although it was observed by histochemical analysis that these promoters drove strong expression of the GUS gene in the embryo and aleurone layer, the activities were not very high on a whole seed basis. This result might be accounted for by the fact that the expression is restricted to a small portion in the rice seed. These embryo storage protein promoters are expected to be an ideal choice for the engineering of seed lipid metabolism, which mainly occurs in the embryo and aleurone.

In the transformants carrying fusions to the seed storage protein gene promoters (*GluB-1*, *GluB-2* and *GluB-4* glutelins, 13 kDa and 16 kDa prolamins, *Glb-1*, REG-2 and Ole18), strong GUS activity was detected in maturing seeds, whereas the leaves and stems showed only background or very weak GUS activity (data not shown). These results indicated that the 5′ flanking regions of these genes were sufficient to determine endosperm or embryo specificity. This is in contrast with the results for the 10 kDa prolamin promoter, which directed GUS gene expression in all tissues of the plant, except for the embryo (Figures 2 and 3). This unusual ectopic expression was prevented by replacing the Nos terminator with the native 10 kDa prolamin terminator. This is the first example of a seed storage protein gene requiring both the 3′ non-coding region as well as the 5′ flanking region for endosperm-specific expression.

This study also showed that, for the endosperm storage protein genes tested, the 5′ flanking regions contained all the elements required for endosperm-specific expression. The GCN4 and AACA motifs are conserved in all glutelin gene promoters isolated so far (Takaiwa *et al*., 1996). The importance of these two motifs in controlling gene expression in the endosperm has been demonstrated (Takaiwa *et al*., 1996; Yoshihara *et al*., 1996). The GCN4 motif is also observed in the PPDK, SBE1 and 13 kDa (*PG5a*) and 16 kDa prolamin promoters. This motif acts as a key element conferring aleurone- and subaleurone-specific expression (Wu *et al*., 1998a). The promoters of at least eight rice seed storage protein genes (*GluB-1*, *GluB-2*, *GluB-4*, *GluA-1*, *GluA-3*, 10 kDa prolamin, 13 kDa (*PG5a*) prolamin and 16 kDa prolamin) have been shown to be expressed predominantly in aleurone and subaleurone cells in transgenic plants. It is likely that the GCN4 motif conserved in these promoters may contribute greatly to this expression pattern. This hypothesis does not contradict the evidence that the embryo-specific REG-2, Ole18 and soybean β-conglycinin gene promoters do not contain the GCN4 motif. The AGPase small subunit and GOGAT promoters also do not contain the GCN4 motif, but they contain various other potential *cis*-elements, such as the endosperm motif, –300 element and typical DNA repeats in the 5′ flanking regions (Anderson *et al*., 1991; Goto *et al*., 1998). Further investigation will be required to fully understand the regulatory mechanisms underlying the endosperm specificity of these promoters.

The activities of the PPDK, AGPase small subunit and AlaAT promoters were relatively high, whereas those of the GOGAT and SBE1 promoters were very low. The PPDK and AGPase small subunit promoters directed the GUS gene to express in vegetative organs as well as in seed (Figure 3). These promoters might be useful for the production of transgenic rice where the final products are accumulated in vegetative tissues.

It is interesting to note that the activity of the *GluB-4* promoter was the highest amongst the promoters tested, despite its relatively short length of 1.4 kb. In light of the fact that about 250 bp of the upstream region overlaps with the coding region of the α-expansin gene, it is notable that the region of about 1 kb between the α-expansin and *GluB-4* coding regions includes all the elements necessary to provide the strongest promoter activity.

It has been demonstrated for the *GluB-1* gene that at least three *cis*-elements containing the GCN4, ACGT and AACA motifs within the –197 bp proximal promoter are the minimal elements required to direct endosperm-specific expression (Wu *et al*., 1998a, 2000). It is noteworthy that, although the 2.3 kb *GluB-1* and 1.3 kb *GluB-1* promoters used in this study were derived from the same *GluB-1* promoter and their expression patterns were similar, their promoter activities were quite different, i.e. the former was about 10-fold stronger than the latter (Figure 5). The only notable difference between the two promoters is the presence of one more ACGT motif in the additional 1 kb fragment of the 2.3 kb promoter. The activity of the *GluB-2* promoter was much lower than that of the 2.3 kb *GluB-1* promoter, even though their 5′ flanking sequences are relatively conserved up to about 600 bp from the transcriptional start site (Takaiwa *et al*., 1991a). This difference in activity may also be explained by the absence of the ACGT motif in the *GluB-2* promoter, although it contains four GCN4 motifs and three AACA motifs which are required

for endosperm-specific expression. Furthermore, it has been demonstrated by loss-of-function experiments that the removal of the ACGT motif in the proximal *GluB-1* promoter causes a remarkable decrease in promoter activity (Washida *et al*., 1999; Wu *et al*., 2000). Taking these results together, it seems that the ACGT motif might be involved in the quantitative regulation of gene expression.

To date, many improvements in rice seed quality have been achieved by introducing deficient nutrients through genetic engineering (Zheng *et al*., 1995; Goto *et al*., 1999; Katsube *et al*., 1999; Ye *et al*., 2000). Our results demonstrate various expression patterns and levels of activity among seed promoters, thus providing alternative choices for expressing foreign genes of interest at a desirable level in the targeted tissue.

Experimental procedures

Plasmid construction

Promoters for the genes of interest were isolated from genomic DNA by PCR using specific primers designed according to information from public databases. The genes and the sizes of the corresponding promoters were as follows: 10 kDa prolamin, 0.8 kb; 13 kDa prolamin (*PG5a*), 0.9 kb; 16 kDa prolamin, 0.9 kb; 26 kDa globulin, 1.0 kb; *GluB-4*, 1.4 kb; rice embryo globulin (REG-2), 1.3 kb; rice 18 kDa oleosin (Ole18), 1.3 kb; rice glutamate synthase gene (GOGAT), 0.8 kb; rice pyruvate orthophosphate dikinase (PPDK), 0.8 kb; rice ADP-glucose pyrophosphorylase small subunit (AGPase), 2.0 kb; rice starch branching enzyme (SBE1), 2.0 kb; and soybean β-conglycinin, 1.0 kb. The 1.3 kb *GluB-1*, 2.3 kb *GluB-1*, 2.4 kb *GluB-2* and 1.0 kb rice alanine aminotransferase (AlaAT) gene promoters were already available in our laboratory.

Various promoter fragments were inserted into a modified binary vector, pGPTV-35S-HPT, containing the hygromycin phosphotransferase (*HPT*) gene as a selectable marker. The modified vector was constructed from the pGPTV-HPT binary vector (Becker *et al*., 1992) by exchanging the Nos promoter for the 0.8 kb CaMV 35S promoter as the promoter of the *HPT* gene. The seed promoters to be tested were introduced upstream of the *UdiA* gene which encodes β-glucuronidase (GUS) in the modified binary vector.

Production of transgenic plants

Transgenic rice plants (*Oryza sativa* cv. Kitaake, an early heading variety) were produced by *Agrobacterium*-mediated transformation. Plasmids constructed as described above were transferred into *Agrobacterium tumefaciens* strain EHA105 by electroporation. Five-week-old calli derived from a mature rice seed were treated with the transformed *A. tumefaciens* for 3 days. The infected calli were successively cultured for 4 weeks each in N6 selection and MS regeneration medium containing hygromycin. The regenerated seedlings were transplanted to a greenhouse as described previously (Goto *et al*., 1999).

Analysis of GUS gene expression

For histochemical analysis, maturing seeds at 7, 12 and 17 DAF were sectioned longitudinally with a razor blade and incubated in 50 mM sodium phosphate buffer (pH 7.0) containing 0.5 mM X-Gluc (5-brom-4-chloro-3-indolyl glucuronide) and 20% methanol at 37 °C. The optimum incubation times for the staining reaction varied from 30 min to overnight depending on the abundance of the GUS activity.

Fluorometric assays of GUS activities were conducted according to Jefferson (1987). Maturing seeds at 17 DAF were homogenized in GUS extraction buffer (50 mM $NaPO_4$, pH 7.0, 10 mM 2-mercaptoethanol, 10 mM Na_2-EDTA, 0.1% sodium dodecylsulphate, 0.1% Triton X-100). After centrifugation, 10 μL of the supernatant was mixed with 90 μL of an assay buffer containing 1 mM 4-methyl umbelliferryl β-D-glucuronide (MUG). After incubation at 37 °C for 1 h, the reaction was stopped by adding 900 μL of 0.2 M Na_2CO_3. Fluorometer values were compared with those of a 4-methyl umbelliferone (4 mU) dilution series. Protein content was determined with a Bio-Rad Protein Assay kit using bovine serum albumin as the standard. Three seeds from each independent transgenic plant were assayed.

Acknowledgements

We thank Ms M. Utsuno, Ms Y. Suzuki, Ms H. Ito and Ms X. Wang for technical assistance. This work was supported in part by the Research and Development Program for New Bio-Industry Initiatives of the Bio-orientated Technology Research Advancement Institution (BRAIN) to F. Takaiwa.

References

Anderson, J.M., Larsen, R., Laudencia, D., Kim, W.T., Morrow, D., Okita, T.W. and Preiss, J. (1991) Molecular characterization of the gene encoding a rice endosperm-specific ADP glucose pyrophosphorylase subunit and its developmental pattern of transcription. *Gene*, **97**, 199–205.

Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) New plant binary vectors with selectable markers located proximal to the left T-DNA border. *Plant Mol Biol*. **20**, 1195–1197.

Chowdhury, M.K.U., Parveez, G.K.A. and Saleh, N.M. (1997) Evaluation of five promoters for use in transformation of oil palm (*Elaeis guineensis* Jacq.). *Plant Cell Report*, **16**, 277–281.

Christensen, A.H., Sharrock, R.A. and Quail, P.H. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol*. **18**, 675–689.

Croissant-Sych, Y. and Okita, T.W. (1996) Identification of positive and negative regulatory *cis*-elements of the rice glutelin Gt3 promoter. *Plant Sci*. **116**, 27–35.

Daniell, H., Stephen, J., Streatfied, S.J. and Wycoff, K. (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci*. **6**, 219–226.

Delaney, D.E. (2002) Choice of crop species and development of transgenic product lines. In: *Plants as Factories for Protein Production* (Hood, E.E. and Howard, J.A., eds), pp. 139–158. Netherlands: Kluwer Academic.

Doyle, J.J., Schuler, M.A., Godette, W.D., Zenger, V., Beachy, R.N. and Slightom, J.L. (1986) The glycosylated seed storage proteins of *Glycin max* and *Phaeolus vulgaris*. Structural homologies of genes and protein. *J. Biol. Chem*. **261**, 9228–9238.

Fischer, R. and Emans, N. (2000) Molecular farming of pharmaceutical proteins. *Transgenic Res*. **9**, 279–299.

Gibbings, J.G., Cook, B.P., Dufault, M.R., Madden, S.L., Khuri, S., Turnbull, C.J. and Dunwell, J.M. (2003) Global transcript analysis of rice leaf and seed using SAGE technology. *Plant Biotechnol J*. **1**, 271–285.

Giddings, G., Allison, G., Brooks, D. and Carter, A. (2000) Transgenic plants as factories for biopharmaceuticals. *Nature Biotechnol.* **18**, 1151–1156.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., Briggs, S. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Japonica*). *Science*, **296**, 92–100.

Goto, F., Yoshihara, T., Shigemoto, N., Toki, S. and Takaiwa, F. (1999) Iron fortification of rice seed by the soybean ferritin gene. *Nature Biotechnol.* **17**, 282–286.

Goto, S., Akagawa, T., Kojima, S., Hayakawa, T. and Yamaya, T. (1998) Organization and structure of NADH-dependent glutamate synthase gene from rice plants. *Biochim. Biophys. Acta*, **1387**, 298–308.

Hiei, Y., Komari, T. and Kubo, T. (1998) Agrobacterium-mediated transformation. In: *Molecular Biology of Rice* (Shimamoto, K., ed.), pp. 235–256. Tokyo: Springer.

Hobbs, S.L.A., Warkentin, T.D. and DeLong, C.M.O. (1993) Transgene copy number can be positively or negatively associated with transgene expression. *Plant Mol. Biol*. **21**, 17–26.

Hood, E.E. and Jilka, J.M. (1999) Plant-based production of xenogenic proteins. *Curr. Opin. Biotechnol*. **10**, 382–386.

Hood, E.E. and Woodard, S.L. (2002) Industrial proteins production from transgenic plants. In: *Plants as Factories for Protein Production* (Hood, E.E. and Howard, J.A., eds), pp. 119–135. Netherlands: Kluwer Academic.

Howard, J.A. and Hood, E.E. (2002) Introduction to molecular farming. In: *Plants as Factories for Protein Production* (Hood, E.E. and Howard, J.A., eds), pp. VII–X. Netherlands: Kluwer Academic.

Imaizumi, N., Ku, M.S.B., Ishihara, K., Samejima, M., Kaneko, S. and Matsuoka, M. (1997) Characterization of the gene for pyruvate orthophosphate dikinase from rice, a C3 plant, and a comparison of structure and expression between C3 and C4 genes for this protein. *Plant Mol. Biol*. **34**, 701–716.

Jefferson, R.A. (1987) Assaying chimeric genes in plant: the GUS gene fusion system. *Plant Mol. Biol. Report*, **5**, 387–405.

Katsube, T., Kurisaka, N., Ogawa, M., Maruyama, N., Ohtsuka, R., Utsumi, S. and Takaiwa, F. (1999) Accumulation of soybean glycinin and its assembly with the glutelins in rice. *Plant Physiol*. **120**, 1063–1073.

Kawasaki, T., Mizuno, K., Baba, Th and Shimada, H. (1993) Molecular analysis of the gene encoding a rice starch branching enzyme. *Mol. Gen. Genet*. **237**, 10–16.

Kikuchi, H., Hirose, S., Toki, S., Akama, K. and Takaiwa, F. (1999) Molecular characterization of a gene for alanine aminotransferase from rice (*Oryza sativa*). *Plant Mol. Biol*. **39**, 149–159.

Masumura, T., Kidzu, K., Sugiyama, Y., Mitsukawa, N., Hibino, T., Tanaka, K. and Fujii, S. (1989a) Nucleotide sequence of a cDNA encoding a major rice glutelin. *Plant Mol. Biol*. **12**, 723–725.

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

Mitsukawa, N., Konishi, R., Uchiki, M., Masumura, T. and Tanaka, K. (1999) Molecular cloning and characterization of a cysteine-rich 16.6-kDa prolamin in rice seeds. *Biosci. Biotechnol. Biochem*. **63**, 1851–1858.

Nakase, M., Hotta, H., Adachi, T., Aoki, N., Nakamura, R., Masumura, T., Tanaka, K. and Matsuda, T. (1996b) Cloning of the rice seed α-globulin-encoding gene: sequence similarity of the 5′-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. *Gene*, **170**, 223–226.

Nakase, M., Yamada, T., Kira, T., Yamaguchi, J., Aoki, N., Nakamura, R., Matsuda, T. and Adachi, T. (1996a) 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.

Rathaus, C., Bower, R. and Birch, R.G. (1993) Effects of promoter, intron and enhancer elements on transient gene expression in sugar-cane and carrot protoplasts. *Plant Mol. Biol*. **23**, 613–618.

Stoger, E., Sack, M., Perrin, Y., Vauero, C., Torres, E., Twyman, R.M., Christou, P. and Fischer, R. (2002) Practical considerations for pharmaceutical antibody production in different crop systems. *Mol. Breed.* **9**, 149–158.

Sun, J.L., Nakagawa, H., Karita, S., Ohyama, K. and Hattori, T. (1996) Rice embryo globulins: amino-terminal amino acid sequences, cDNA cloning and expression. *Plant Cell Physiol*. **37**, 612–620.

Takaiwa, F., Oono, K. and Kato, A. (1991a) 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. and Kato, A. (1991b) Sequences

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, H. and Washida, H. (1996) Characterization of common *cis*-regulatory elements responsible for the endosperm specific expression of members of a rice glutelin multigene family. *Plant Mol. Biol*. **30**, 1207–1221.

Walmsley, A.M. and Arntzen, C.J. (2000) Plants for delivery of edible vaccines. *Curr. Opin. Biotechnol.* **11**, 126–129.

Washida, H., Wu, C.Y., Suzuki, A., Yamanouchi, U., Akihama, T., Harada, K. and 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.

Wu, C.Y., Adachi, T., Hatano, T., Washida, H., Suzuki, A. and Takaiwa, F. (1998b) Promoters of rice storage protein genes direct endosperm-specific gene expression in transgenic rice. *Plant Cell Physiol*. **39**, 885–889.

Wu, C.Y., Suzuki, A., Washida, H. and Takaiwa, F. (1998a) 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**, 673–983.

Wu, L.S., Wang, L.D., Chen, P.W., Chen, L.J. and Tzen, J.T. (1998c) Genomic cloning of 18 kDa oleosin and detection of triacylglycerols and oleosin isoforms in maturing rice and postgerminative seedlings. *J. Biochem*. **123**, 386–391.

Wu, C.Y., Washida, H., Onodera, Y., Harada, K. and 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.

Ye, X., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P. and Potrykus, I. (2000) Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science*, **287**, 303–305.

Yoshihara, T., Washida, H. and 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.

Zhao, Y., Leisy, D.J. and Okita, T.W. (1994) Tissue-specific expression and temporal regulation of the rice glutelin Gt3 gene are conferred by at least two spatially separated *cis*-regulatory elements. *Plant Mol. Biol*. **25**, 429–436.

Zheng, Z., Kawagoe, Y., Xiao, S., Li, Z., Okita, T.W., Hau, T.L., Lin, A. and Murai, N. (1993) 5′ distal and proximal *cis*-acting regulatory elements are required for developmental control of a rice seed storage protein glutelin gene. *Plant J*. **4**, 357–366.

Zheng, Z., Sumi, K., Tanaka, K. and Murai, N. (1995) The bean seed storage protein β-phaseolin is synthesized, processed, and accumulated in the vacuolar type-II protein bodies of transgenic rice endosperm. *Plant Physiol*. **109**, 777–786.

Zhu, T., Budworth, P., Chen, W., Provrt, N., Chang, H.-S., Su, W., Estes, B., Zou, G. and Wang, X. (2003) Transcriptional control of nutrient portioning during rice grain filling. *Plant Biotechn J*. **1**, 59–70.