Natural variation of OsGluA2 is involved in grain protein content regulation in rice

Yihao Yang1,2,3,4, Min Guo1,2,3,4, Shengyuan Sun1,2, Yelu Zou2, Shuangyi Yin1, Yannan Liu2, Shuzhu Tang3, Minghong Gu1,2, Zefeng Yang1,2 & Changjie Yan1,2,3

Grain protein content (GPC) affects rice nutrition quality. Here, we identify two stable quantitative trait loci (QTLs), qGPC-1 and qGPC-10, controlling GPC in a mapping population derived from indica and japonica cultivars crossing. Map-based cloning reveals that OsGluA2, encoding a glutelin type-A2 precursor, is the candidate gene underlying qGPC-10. It functions as a positive regulator of GPC and has a pleiotropic effect on rice grain quality. One SNP located in OsGluA2 promoter region is associated with its transcript expression level and GPC diversity. Polymorphisms of this nucleotide can divide all haplotypes into low (OsGluA2LET) and high (OsGluA2HET) expression types. Population genetic and evolutionary analyses reveal that OsGluA2LET, mainly present in japonica accessions, originates from wild rice. However, OsGluA2HET, the dominant type in indica, is acquired through mutation of OsGluA2LET. Our results shed light on the understanding of natural variations of GPC between indica and japonica subspecies.
rice (*Oryza sativa* L.) is the most important human food crop in the world, providing over 21% of the calorific needs for the world’s population and up to 76% of the calorific intake for the South East Asian. With the improvement of people’s living standards, rice consumers are paying much attention to good grain quality. Rice grain quality is composed of four aspects, grain appearance, milling quality, nutrition quality, and cooking and eating quality. Of which, nutrition quality as well as cooking and eating quality are two most important indices that are widely concerned by the consumers. Three major components in endosperm, i.e. starch (70–80%), protein (7–10%), and lipids (~1%), predominantly determine the nutrition quality and cooking and eating quality. Therefore, unraveling the genetic basis underlying the synthesis of starch, protein, and lipids is a prerequisite for the improvement of rice grain quality. In recent decades, great achievements have been made in the dissection of starch synthesis pathway and the genetic network controlling starch variations. The genetic basis for grain protein content (GPC), another important factor affecting rice nutrition quality as well as eating and cooking quality remains largely unclear.

Rice grain protein consists of two categories, functional protein (~10%) and seed storage protein (SSP, ~90%). According to the solubility-linked physical properties, SSPs are classified as four fractions: albumins, globulins, prolamins, and glutelins. Among them, glutelin is the most abundant one, which comprises about 60–80% of the total SSPs. The nutritional value of rice glutelin is superior to other rice storage proteins due to its higher lysine content and greater digestibility by the humans. Therefore, any large change of glutelin content will certainly affect the grain nutrition quality. Glutelins can be further divided into four groups (GluA, GluB, GluC, and GluD) based on their amino acid sequence similarity, and synthesized as a 57 kDa precursor and then cleaved into a 37–39 kDa acidic subunit and a 22–23 kDa basic subunit in the cytoplasm.

Asian cultivated rice is generally divided into two main subspecies, *indica* and *japonica*. Investigations have revealed that both *indica* and *japonica* cultivars have substantial variation in GPC with a range from 4.9% to 19.3% and from 5.9% to 16.5%, respectively, and that the mean of GPC in *indica* is generally higher than that of *japonica*. The difference of GPC between *indica* and *japonica* might be due to their different genetic architectures. However, the mechanism underlying the GPC differences remains largely unclear.

Significant efforts have been taken toward dissecting the genetic mechanism of rice GPC. Consequently, many genes (regulators) affecting GPC have been isolated from various mutants in rice, along with hundreds of QTLs have been reported. However, GPC belongs to a typical quantitative trait with complex genetic structure and is susceptible to environmental factors, especially to nitrogen fertilization in the late growth stages. Thus, the QTLs for rice GPC detected using different genetic populations are frequently inconsistent, and few of them can be repeatedly detected. To date, only one QTL (*qPC1*), on the long arm of chromosome 1, has been cloned. *PC1* encodes a putative amino acid transporter OsAAP6 and functions as a positive regulator of rice GPC. Beyond that, little progress has been achieved in the breeding improvement of rice GPC due to lacking of gene targets.

In this study, we illustrate the putative genetic mechanism of rice GPC variation through QTL mapping, gene cloning, and functional detection, and demonstrate that glutelin content is the major contributor to the variation of rice GPC between *indica* and *japonica* subspecies. Two environmentally stable QTLs for GPC, *qGPC-1* and *qGPC-10*, are identified. *qGPC-10*, encoding a glutelin type-A2 precursor *OsGluA2*, is isolated through map-based cloning. We reveal that *OsGluA2* functions as a positive regulator of GPC through complementation, knockout (CRISPR-Cas9), and natural variation analyses. Our evolutionary and population genetic analyses show that *OsGluA2* significantly contributes to the genetic difference between *indica* and *japonica* subspecies.

**Results**

**Glutelin content contributes largely to GPC variation.** We firstly surveyed GPC in a core collection of 402 rice accessions, including 205 *indica* and 197 *japonica* cultivars, in two environmental conditions. The results showed that GPC differs tremendously among rice cultivars, ranging from 5.33% to 14.83%, 81.5% of cultivars concentrated in the range of 7.5–11.5% (Fig. 1a). Moreover, as expected, environmental factor has a significant effect on GPC as the mean of GPC in 2013 (10.56 ± 0.07%) is significantly higher than that in 2012 (8.66 ± 0.06%) (Supplementary Table 1).

GPC variation was further analyzed in *indica* and *japonica* subspecies, respectively. The results indicated that there is a significant difference between the two species in both two environmental conditions. The GPC of *indica* cultivars is 9.29 ± 0.09%, while 7.99 ± 0.07% in *japonica* in 2012. The same tendency was also found in 2013; the GPC are 11.19 ± 0.08% and 9.91 ± 0.08% for *indica* and *japonica*, respectively (Supplementary Table 1). Using the values of *x* ± s.d. as quantile, all the rice accessions were classified into three types, high, medium and low types. We noticed that the frequency distributions in the two subspecies are unbalanced in three types (*x*² = 6.474, *P* = 0.039). For *indica* subspecies, it possesses a higher proportion of high GPC cultivars; inversely, higher proportion of low GPC accessions occurs in *japonica* subspecies (Supplementary Table 2). Thus, it was concluded that *indica* cultivars exhibited higher GPC than *japonica*, consistent with previous findings.

To address why the GPC in *indica* subspecies is generally higher than that in *japonica*, 43 conventional *indica* and 60 *japonica* cultivars from 402 accessions with similar heading date and plant height were chosen to establish a subpopulation subject to further assay of four fractions of SSP by the Kjeldahl method. Statistical analysis revealed that glutelin and albumin + globulin contents show significant difference between the two subspecies, explaining 73.9% and 15.8% of the total variance of GPC, while no significant differences were detected for prolamin content in the test population (Fig. 1b, c). Thus, it is very clear that glutelin content variation acts as a major contributor to the GPC variation between *indica* and *japonica* subspecies, suggesting that glutelin content could be directly used as a target for rice nutrition quality improvement.

**Identification of two major QTLs for GPC variation.** To understand the genetic mechanisms of GPC in rice, a chromosomal segment substitution line (CSSL) population derived from a cross of an *indica* cultivar Habataki and a *japonica* cultivar Sasanishiki was employed for QTL mapping (Supplementary Table 3). In this population, a total of 18 QTLs for GPC were identified across three environmental conditions with a threshold of LOD ≥ 2.5. Among them, *qGPC-1* and *qGPC-10* were repeatedly identified under all three environmental conditions, while *qGPC-3*, *qGPC-8*, and *qGPC-12* were detected in two environmental conditions, and others can only be identified in one environmental condition, suggesting they might be sensitive to environmental factors (Fig. 2a and Supplementary Table 4). The alleles from Habataki on *qGPC-3*, *qGPC-8*, *qGPC-10*, and *qGPC-12* can increase GPC, while the allele of *qGPC-1* will decrease GPC.

**Functional characterization of *OsGluA2***

Using T-DNA and CRISPR-Cas9, we knocked out *OsGluA2* in *OsGluA2* overexpression lines and restored the activity of GluA protein in T3 generation. Our findings indicated that *OsGluA2* positively regulates GPC through complementation and knockout (CRISPR-Cas9), and the activity of *OsGluA2* could increase GPC. Our results showed that *OsGluA2* significantly contributes to the genetic difference between *indica* and *japonica* subspecies.
Delimiting qGPC-10 to a 35-kb long region. Because qGPC-10 is a stably inherited QTL with large effect for GPC, map-based cloning has been carried out for qGPC-10. According to the preliminary mapping results, the CSSL line SL431 harboring qGPC-10 was selected as parent to cross with Sasanishiki to develop an F₂ population for fine mapping (Supplementary Fig. 1a, b). The grain protein profiles of two parents were analyzed by using SDS-PAGE analysis. The densitometry analysis indicates that SL431 has higher accumulation of proglutelins, acidic and basic subunits, as well as prolams than those in Sasanishiki (Supplementary Fig. 1c). This observation was further confirmed by the results of protein fraction analysis (Supplementary Fig. 1d). Unexpectedly, a dramatic segregation of GPC in F₂ population containing 2085 plants was observed, ranging from 5% to 17% (Fig. 2f). Putatively encoding a precursor glutelin type-A2 precursor (referred to hereafter as OsGluA2), LOC_Os10g26050, LOC_Os10g26060, LOC_Os10g26070, and LOC_Os10g26110 were predicted. The projection of the gene OsGluA2 was responsible to qGPC-10.
We screened for the KO-OsGluA2 with two-base deletion in the second exon predicted to produce truncated poly-peptides with 187 amino acids (Supplementary Fig. 3b–d). As expected, grains produced from homozygous KO-OsGluA2 plants showed a lower GPC phenotype (Fig. 2h). Meanwhile, there were no significant differences between the NILs and transgenic plants in other agronomic traits and yield components (Supplementary Fig. 3a and Supplementary Table 7). All these evidences confirm that OsGluA2 is responsible to qGPC-10.

Higher expression level of OsGluA2 promotes GPC. In order to illustrate how OsGluA2 regulates GPC, qRT-PCR was used to...
investigate the expression patterns of OsGluA2 between NILs (NIL-qGPC-10H and NIL-qGPC-10S) (Fig. 3a). OsGluA2 transcripts were abundant in the endosperms at 15–25 days after flowering (DAF), similar to the previous results. Furthermore, significant differences of expression level were observed at 15 DAF, higher in NIL-qGPC-10H than in NIL-qGPC-10S during the endosperm development. The different expression level of OsGluA2 in NILs implies that the sequence variations in promoter region may affect the expression level of OsGluA2, and thus resulting in the diver GPCs.

The expression pattern of OsGluA2 was further analyzed by using transgenic plants carrying qGPC-10S::GUS (OsGluA2 promoter from Sasanishiki) and qGPC-10H::GUS (OsGluA2 promoter from SL431), respectively. The GUS staining in the endosperm gradually went darker along with grain filling, and the GUS activity of qGPC-10H::GUS was higher than that of qGPC-10S::GUS at 15, 20, and 25 DAF, respectively, consistent with the result obtained by qRT-PCR (Fig. 3b, c).

OsGluA2 has pleiotropic effects on rice grain quality. To explore the genetic effect of OsGluA2 on rice grain quality, the NILs and transgenic plants were investigated for GPC, SSP fractions, as well as starch content and physico-chemical properties (Table 1). The GPC, as well as glutelin, albumin, and prolamin contents in NIL-qGPC-10H are significantly higher than NIL-qGPC-10S, with an exception of globulin contents. Similar results were also observed in complementation (HpHc) and knockout (KO-OsGluA2) transgenic progenies. These results clearly indicate that OsGluA2 has a noticeable effect on GPC and most of the SSP fractions, with the largest effect on glutelin variations (\(P = 2.5 \times 10^{-4}\), independent-samples \(t\)-test). Hence, it is considered that the change in glutelin quantity is the major contributor to the total GPC variations in NILs and transgenic plants. Additionally, the starch contents and amylase contents (AC) were slightly changed subjecting to the genetic background. However, the gel consistencies (GCs) were altered significantly. When the GPC increases, the GC dramatically decreases in NILs (\(P = 3.7 \times 10^{-4}\), independent-samples \(t\)-test) and complementation test (\(P = 4.8 \times 10^{-3}\), independent-samples \(t\)-test). Similarly, when in KO (+) plants decreases, GC increases (\(P = 7.2 \times 10^{-3}\), independent-samples \(t\)-test). The fact that the GPC is negatively correlated with GC implies that GPC might affect rice eating and cooking quality as well.

To determine whether the alternation in the accumulation of SSPs and starch are due to the transcriptional changes of other genes related to grain storage materials, we examined the transcription levels of 32 genes, including 16 hallmark genes involved in SSP synthesis and 16 genes related to starch metabolism at 15 DAF (Fig. 3d). In general, the expression of genes encoding glutelins (GluA1, GluA2, GluB1, GluC, and RAG2) and prolamins (17-kD prolamins) are up-regulated at various extent in NIL-qGPC-10H when compared with NIL-qGPC-10S. Similarly, these genes are noticeably down-regulated in KO-OsGluA2 plants. At the same time, the transcriptional level of genes related to starch metabolism also changed to various extents. These results strongly suggest that OsGluA2 has significant associating effect on the expression of a large portion of other genes participating in grain storage materials metabolism in rice developing grains.

Moreover, the total amino acid contents of grains were also measured both in the NILs and transgenic plants (Supplementary Fig. 4). Compared with the corresponding levels in NIL-qGPC-10H and transgene-negative HpHc(-) plants, the amounts of aspartic acid, threonine, glutamic acid + glutamine, isoleucine, lysine, arginine and proline amino acids and total content of amino acids were significantly increased in NIL-qGPC-10H and transgene-positive HpHc(+), plants, respectively, whereas the reverse was true for the KO-OsGluA2. These results suggest that OsGluA2 enhances GPC by increasing the grain storage protein content and the total amount of amino acids, and thus improves nutritional quality.

GPC enriched grains have more and larger protein body II. We anatomically determined the effect of OsGluA2 on protein body (PB) formation by transmission electronic microscope (Fig. 3e). In the developing endosperms at 15 DAF, two types of PBs can be well discerned in ultrathrin sections, including prolamins-containing PB I and and glutelin/globulin-containing PB II. In the developing endosperms of NIL-qGPC-10H and HpHc(+), the size (section areas) of PB II were more expanded, whereas the opposite was true for KO(-), comparing to the control plants (Fig. 3f). These results indicate that OsGluA2 has enhanced the total amount of glutelin content, resulting in the size enlargement and increased number of PB IIs.

One SNP is associated with transcriptional level and GPC. To identify the natural variation of OsGluA2, we firstly sequenced the genomic region of OsGluA2 from Sasanishiki and SL431. We found no nucleotide differences between the two cultivars in the coding regions of OsGluA2. Instead, six polymorphisms in the promoter region (2 kb) and one polymorphism in 3′-untranslated region (3′ UTR) were revealed (Fig. 4a).

In order to trace the distribution pattern of OsGluA2, we then sequenced the full-length OsGluA2 genes of 100 accessions from different countries, including 31 temperate japonica, 24 tropical japonica, 22 aus, 21 indica, and 2 aromatic accessions (Fig. 4b). Phylogenetic analysis showed that OsGluA2 diverges into five types (haplotypes) (Fig. 4c). To test the effect of these seven mutations on gene expression, we measured the transcript abundances of OsGluA2 in these five types in endosperms at 15 DAF and their corresponding glutelin contents in mature seeds. The result indicates that the cultivars possessing type 1 tend to show significant lower expression levels and lower glutelin content than other types (Fig. 4d, e). These results suggest that four nucleotide mutations (SNP-1191, SNP-1100, InDel-229 to
Fig. 3 The expression patterns of OsGluA2 and genes involved in storage materials. a The expression patterns of OsGluA2 was determined by qRT-PCR. 5E, 10E, 15E, 20E, and 25E: endosperms at 5, 10, 15, 20, and 25 DAF. n = 3. b GUS assays in transgenic rice carrying qGPC-10S::GUS and qGPC-10H::GUS at 15, 20, and 25 DAF. Scale bar = 1 mm. c GUS activity in maturing seeds at 15, 20, and 25 DAF. GUS activity is expressed as pmol 4MU min⁻¹ μg⁻¹ protein. n = 10. d Expression levels of key genes involved in synthesis and storage of endosperm components in T2 transgenic plants and NILs. Transgene-positive plants of OsGluA2 and NIL-qGPC-10H are shown relative to transgene-negative plants and NIL-qGPC-10S, respectively, set as 1. Error bars, s.d., n = 3. e Ultrastructure of cells in developing endosperm of NILs and transgenic plants in 15 DAF. PB I and PB II protein bodies I and II, SG starch granule, ECS extracellular space, CW cell wall. Scale bars = 2 μm. f Quantitative comparison of PB II area in NILs and transgenic plants. P-values were calculated by independent-samples t-test. Error bars, s.d., n = 3. Source data of Figs. 3a, 3c, 3d and 3f are provided as a Source Data file.
OsGluA2 contributes to the divergence of indica and japonica. In order to investigate the natural variation of OsGluA2 in rice germplasm, we further analyzed the genome sequences of this gene in 3005 cultivated accessions, constitutive of cultivars from seven groups. We noticed that the distribution of two functional types of OsGluA2 is unbalanced in seven cultivated rice groups (Supplementary Table 8, $X^2 = 1025.774, P < 0.0001$). The cultivars of temperate japonica possessed the highest proportion of OsGluA2$^{LET}$ (213 out of 317, 67.19%), while those of indica have the lowest level (70 out of 1755, 3.99%). In addition, the proportions of OsGluA2$^{LET}$ in unclassified japonica (74 out of 131, 56.49%) and tropical japonica (141 out of 387, 36.43%) are also much higher than that of indica. These results indicate that japonica cultivars generally have higher proportion of OsGluA2$^{LET}$ than indica cultivars. Moreover, we noticed that the proportion of OsGluA2$^{LET}$ has the trend of increasing along with the increase of latitude, suggesting that a regional differentiation in OsGluA2 (Fig. 5a), especially for japonica. We further investigated the parameters of genetic divergence for OsGluA2 and its flanking regions between indica and japonica subspecies, including the estimates of haplotype and nucleotide $F_{ST}$, Nei’s $G_{ST}$, and Hudson’s $G_{ST}$ and $H_{ST}$ (Fig. 5b). All the estimates of these five parameters in OsGluA2 locus are much higher than its flanking genomic regions. These results suggest there is genetic differentiation in OsGluA2 between the two subspecies, which might contribute to the divergence of indica and japonica. OsGluA2$^{LET}$ is derived from wild rice. In order to reveal whether artificial selection has contribution to the domestication of OsGluA2, the genomic information of 3005 cultivated rice and 27 O. rufipogon accessions were used to analyze the genetic diversity of this gene and its flanking regions (Fig. 5c and Table 2). The nucleotide diversity value ($\pi$) for this gene is much lower than its flanking regions in both cultivated and wild populations (O. rufipogon), suggesting that low nucleotide diversity in the locus of OsGluA2 might be the result of natural selection rather than artificial selection. Moreover, all the estimation values of Tajima’s $D$ for cultivated and wild populations are negative, indicating low-frequency polymorphisms in these regions. However, none of these values are statistically significant except for the wild population, suggesting that the locus of OsGluA2 did not escape from neutral evolution during domestication.

| Lines       | $n$ | Protein content (mg g$^{-1}$) | Albumins content (mg g$^{-1}$) | Globulins content (mg g$^{-1}$) | Prolamins content (mg g$^{-1}$) | Glutelins content (mg g$^{-1}$) | AC (%) | GC (mm) | Starch content (%) |
|-------------|-----|-------------------------------|-------------------------------|-------------------------------|-------------------------------|--------------------------------|--------|---------|-------------------|
| NIL-gGCP-10$^K$ | 20  | 131.3 ± 4.5                  | 19.5 ± 0.8                    | 20.6 ± 0.1                    | 10.1 ± 0.6                    | 68.4 ± 3.4                     | 15.9 ± 0.4 | 93.7 ± 4.5 | 88.5 ± 0.7            |
| NIL-gGCP-10$^H$ | 20  | 161.4 ± 6.0                  | 23.7 ± 0.9                    | 20.5 ± 0.1                    | 12.5 ± 0.4                    | 85.5 ± 3.2                     | 14.9 ± 0.3 | 70.0 ± 4.1 | 86.1 ± 0.5            |
| $P$-value    |     | 2.9 ± 10$^{-4}$              | 2.5 ± 10$^{-3}$               | 9.7 ± 10$^{-2}$               | 1.0 ± 10$^{-3}$               | 8.4 ± 10$^{-4}$                | 6.7 ± 10$^{-2}$ | 3.7 ± 10$^{-4}$ | 5.1 ± 10$^{-2}$          |
| HpHc (--)   | 20  | 129.8 ± 3.9                  | 13.2 ± 1.1                    | 19.2 ± 0.7                    | 12.4 ± 1.6                    | 69.9 ± 1.9                     | 15.7 ± 0.3 | 94.4 ± 2.2 | 87.9 ± 0.3            |
| HpHc (+)    | 20  | 148.1 ± 5.3                  | 21.8 ± 1.4                    | 20.8 ± 0.6                    | 19.2 ± 1.5                    | 80.1 ± 1.6                     | 14.8 ± 0.3 | 80.6 ± 3.3 | 85.6 ± 0.4            |
| $P$-value    |     | 7.9 ± 10$^{-3}$              | 2.4 ± 10$^{-5}$               | 9.3 ± 10$^{-2}$               | 4.0 ± 10$^{-3}$               | 2.5 ± 10$^{-4}$                | 5.2 ± 10$^{-2}$ | 4.8 ± 10$^{-3}$ | 1.2 ± 10$^{-2}$          |
| KO (--)     | 20  | 153.9 ± 3.7                  | 20.3 ± 0.8                    | 21.8 ± 0.2                    | 11.3 ± 0.7                    | 83.7 ± 2.8                     | 14.9 ± 0.4 | 72.3 ± 4.6 | 86.5 ± 0.3            |
| KO (+)      | 20  | 136.8 ± 3.0                  | 16.6 ± 1.1                    | 21.4 ± 0.3                    | 9.0 ± 0.3                     | 70.1 ± 2.3                     | 16.4 ± 0.3 | 91.2 ± 4.8 | 88.8 ± 0.3            |
| $P$-value    |     | 9.5 ± 10$^{-4}$              | 7.7 ± 10$^{-3}$               | 2.4 ± 10$^{-1}$               | 4.0 ± 10$^{-3}$               | 4.8 ± 10$^{-4}$                | 3.5 ± 10$^{-3}$ | 7.2 ± 10$^{-3}$ | 7.2 ± 10$^{-3}$          |

AC: amylose content, GC: gel consistency. At least 300 grains of each plant were harvested. All data are given as means ± s.e. ($n = 3$ in starch content). $P$-values were calculated by $t$-test. Source data of Table 1 are provided as a Source Data file.

Discussion
Rice grain quality is a complex character, of concern to a wide range of consumers and breeders. GPC serves as a critical factor in determining rice nutrition quality, as well as eating and cooking quality. However, the genetic mechanism of GPC remains largely unclear yet, mainly due to the difficulty in mapping and cloning GPC related genes/QTLs. In the present study, we reveal that GPC variation in rice germplasm is mainly attributed to the glutelin content variation. Subsequently, two stably inherited QTLs, qGPC-1 and qGPC-10, are identified in a cross of indica and japonica. qGPC-10, encoding the glutelin type-A2 precursor, is successfully cloned and functionally characterized. Our results signify a step toward unraveling the mechanism underlying natural variations of GPC between indica and japonica subspecies.

Rice is a widely distributed crop and there are many differences between indica and japonica subspecies. Like many other agronomic traits, GPC is also found to dramatically differentiate...
between the two subspecies, being generally higher in \textit{indica} ones\textsuperscript{14,15}. However, why is the GPC in \textit{indica} accessions higher than in \textit{japonica}? In the present study, we reveal that glutelin contents variation is the main reason causing GPC variation between \textit{indica} and \textit{japonica} germplasms, accounting for more than 70\% ($R^2 = 0.739$) of phenotypic variation (Fig. 1b, c). This result brings us some enlightenment that glutelin content should be regarded as a target for direct manipulation in rice quality improvement. Interestingly, in our study, qGPC-10, one of the environmentally stable QTL detected in the \textit{indica} and \textit{japonica} cross, is found to encode glutelin precursors and thus to regulate the glutelin synthesis and accumulation through the way of differentially transcriptional expression intensity. This stimulates us to explore that whether other glutelin synthesis-related genes are also responsible for the glutelin accumulation contribute to the GPC variation in the same way. To address this issue, 12 modern cultivars (six typical \textit{indica} and six \textit{japonica} cultivars) were employed to investigate the expression level of glutelin synthesis-related genes, it happens that there is a similar pattern, most of glutelin synthesis-related genes exhibit higher transcriptional...
Fig. 5 The geographic distribution, phylogenetic relationship and genetic difference of OsGluA2. a The geographic distributions among 3005 cultivated rice accessions. The types of OsGluA2LET and OsGluA2HET were indicated by green and red circles, respectively. b The parameters of genetic difference between indica and japonica ecotypes for OsGluA2 and its flanking genomic regions. c Nucleotide diversity and selection analyses in OsGluA2 and flanking regions (~80 kb). Genomic sequences of 3005 cultivated and 27 wild accessions were acquired from RFGB (http://www.rmbreeding.cn/Snp3k) and OryzaGenome (http://viewer.shigen.info/oryzagenome/), respectively. ADM admixed, ARO aromatic, IND indica, JAP unclassified japonica, TEJ temperate japonica, TRJ tropical japonica. d Phylogenetic relationship of OsGluA2 generated from 3005 cultivated and 27 wild rice accessions. e Haplotype network of OsGluA2. Circle size is proportional to the number of samples for a given haplotype. Black spots on the lines indicate mutational steps between haplotypes.

Table 2 The estimated parameters of nucleotide diversity and Tajima’s D of OsGluA2 and its flanking regions

| Parameter     | ADM  | AUS  | ARO  | IND  | TEJ  | TRJ  | JAP  | Cultivated | Wild  |
|---------------|------|------|------|------|------|------|------|------------|-------|
| Upstream 40 kb | \(\pi\) | 0.0070 | 0.0043 | 0.0076 | 0.0044 | 0.0051 | 0.0069 | 0.0067 | 0.0063 | 0.0125 |
|               | \(\theta\) | 0.0065 | 0.0066 | 0.0069 | 0.0083 | 0.0053 | 0.0065 | 0.0063 | 0.0095 | 0.0194 |
| OsGluA2       | \(\alpha\) | 0.0008 | 0.0005 | 0.0010 | 0.0004 | 0.0005 | 0.0006 | 0.0005 | 0.0008 | 0.0017 |
|               | \(\delta\) | 0.0011 | 0.0011 | 0.0013 | 0.0012 | 0.0007 | 0.0006 | 0.0008 | 0.0019 | 0.0034 |
| Tajima’s D    | \(-0.6403\) | \(-1.0996\) | \(0.3464\) | \(-1.3539\) | \(-0.1449\) | 0.1775 | 0.2363 | \(-0.9493\) | \(-1.4035\) |
| Downstream 40 kb | \(\pi\) | 0.0060 | 0.0046 | 0.0062 | 0.0035 | 0.0053 | 0.0060 | 0.0059 | 0.0053 | 0.0095 |
|               | \(\theta\) | 0.0055 | 0.0054 | 0.0055 | 0.0072 | 0.0048 | 0.0055 | 0.0047 | 0.0083 | 0.0150 |
| Tajima’s D    | 0.3125 | \(-0.5139\) | 0.4022 | \(-1.4720\) | 0.2731 | 0.2926 | 0.8691 | \(-1.0362\) | \(-1.4514\) |

*Single asterisk (*) indicates significance level at \(P < 0.05\) tested by computing 1000 coalescent simulations.*
expression levels in indica genetic background except for GluB2 (Supplementary Figs. 5 and 6). Based on this phenomenon, it is reasonable to deduce that the expression level differences in glutelin synthesis gene may be the main cause of glutelin content variations between indica and japonica subspecies. 

SSPs, accounting for ~90% of rice GPC, are storage nitrogen sources for the germinating seedlings and serve as a nutrient source for humans and livestock. SSPs genes constitute of multigenic families. Previous studies showed that several mutations in a few genes have little effect on seed protein content and amino acid composition. Kawakatsu et al. (2010) characterized glutelin, globulin, and prolamin knockdown lines in rice, and found that a reduction of one or a few SSPs could be compensated for by increasing in other SSPs at both the mRNA and protein levels. For example, Glu-less lines display marked reduction of Glu-2 and Glu-B-4, and bands corresponding to Glu-A1, Glu-B-2, and Glu-B-1 are invisible on a stained gel, whereas levels of globulin 1 (Glb-1) and 13-kDa prolamins significantly increased. Similarly, they also found the 13-kDa prolamin-less transgenic seeds have high levels of glutelin and Glb-1 (ref. 40). However, in the present study, we find that the glutelin and prolamin contents alter in the same directions, increase or decrease simultaneously in NILs, HpHc, and KO transgenic lines (Table 1). Moreover, this phenomenon is also observed in a similar case, in which reduced abundances of glutelin and prolamin are simultaneously observed in dry seeds of a globulin-deficient rice mutant, which is generated with RNA interference (RNAi)-induced suppression of globulin expression. Thus, further studies are needed to address whether there is any common mechanism in corporately regulating the glutelin content and prolamin content in rice.

In the past decade, although numerous QTLs for GPC variation have been detected in rice germplasms, few QTLs have been cloned except for qPCI/OsAAP6 (ref. 35). However, the OsAAP6 expression level is found to be associated with GPC variation only in indica accessions, and no correlation between OsAAP6 expression level and GPC variation was detected in the japonica genetic background. It seems that OsAAP6 could be only used as a target gene to regulate GPC in indica breeding programs. Unlike OsAAP6, two haplotypes of OsGluA2, OsGluA2\textsuperscript{LET} and OsGluA2\textsuperscript{HET}, are found to mainly reside in japonica and indica cultivars, respectively (Fig. 4b).

The OsGluA2\textsuperscript{LET} exhibits a lower transcriptional level than OsGluA2\textsuperscript{HET}, which corresponds with the glutelin content variation (Fig. 4c–e). So, our result suggests that the OsGluA2\textsuperscript{LET} allele from japonica accessions could be directly replaced with OsGluA2\textsuperscript{HET} allele to improve the nutrition quality in japonica cultivar development through marker-assisted selection. It should be noted that GPC is also found to be inversely correlated with qGPC-10. An additional 1085 F\textsubscript{3} individuals were then used for screening recombinants for fine mapping of qGPC-10. The genomic DNA fragments corresponding to the candidate gene in SL431 and Sasanishiki were sequenced and analyzed. A list of the markers used for QTL analysis and positional cloning is given in Supplementary Table 9.

**Trait measurement.** Rice grains of each sample were harvested at ~40 DAF and air-dried, then stored at room temperature for 3 months for measuring grain quality and yield traits. For each plant, 15–20 g of brown rice were prepared for the measurement of grain quality, including AC, GC, starch content, and GPC. The latter was determined by near-infrared spectroscopy on 1241 Grain Analyzer (Foss Tecator, Sweden) equipped with STM model. Briefly, approximately 15–20 g dry seeds per sample were dehulled into brown rice using a TR 200 dehuller (Kett, Tokyo, Japan) and loaded into the sample cup. Three scans per sample were used in data analysis. Data of seed protein contents were presented on a 0% moisture basis. These calibrations involved more than 500 rice samples that varied in GPC. Grains GPC and storage protein fractions in rice flour were also determined by the Kjeldahl method using a Kjeltec 2300 Autoanalyzer (Foss AB, Sweden). A nitrogen conversion factor of 6.25 was used to calculate the GPC.

**Protein extraction from rice grains and the SDS-PAGE assay.** Mature dry seeds of rice plants were harvested, dehulled and ground to a fine powder using a mortar and pestle at room temperature. The total seed proteins were extracted from the powder (100 mg) with 1 mL of SDS-sample buffer (4% SDS, 8 M urea, 0.25 M Tris-HCl (pH = 6.8), 20% glycerol, and 5% β-mercaptoethanol (β-ME)) by shaking overnight. The supernatant was removed after centrifugation at 1500 × g for 30 min. The total seed protein (10 μL) was separated by gradient SDS-PAGE gels (15%) and the densitometry of the gels was measured by using the ImageJ software.

**RNA extraction, CDNAs preparation, and quantitative reverse transcriptase PCR.** Total RNAs were extracted from rice various tissues using the Trizol reagent (Invitrogen). Total RNA was used for synthesizing first-strand cDNA with a reverse transcription kit (TOYOBO). Quantitative reverse transcription (qRT-PCR) was performed on an ABI 7500 instrument using SYBR Green PCR Mastermix followed the manufacturer’s instructions (Takara, RR20). All assays were performed with at least three biological replicates. Rice TUBULIN gene serves as an internal control to normalize gene expression. A list of primers used for qRT-PCR is given in Supplementary Tables 9 and 10.

**Vector constructions and plant transformation.** Various DNA fragments were amplified and inserted into different binary vectors for plant transformation. The information of relevant primer sequences was given in Supplementary Table 9. To further complement the SGC-Pi in SL431 (Hpla), we fused a 1 kb promoter fragment and the whole-genomic sequence of OsGluA2 by PCR amplification from SL431 and then inserted them into the plant binary vector pCAMBIA1301. To prepare the construct for CRISPR-Cas9 (KO-OsGluA2), a 20 bp fragment with NGG (PAM) at 3’ end from the second exon of OsGluA2 was introduced into the vector pCAMBIA1301. To confirm the片段of OsGluA2 were amplified from a cross between Sasanishiki and SL431, respectively, and then were inserted into the expression vector pCAMBIA1391Z-GUS, namely GPC-10\textsuperscript{L-GUS} and GPC-10\textsuperscript{H-GUS}. The

**Methods**

**Plant materials and cultivation.** The CSSL population derived from a cross of Sasanishiki (japonica) and Habataki (indica) with 39 lines in total was employed for identification of QTLs controlling GPC, which was introduced from National Institute of Agrobiological Sciences, Japan (http://www.rgrc.dna.affrc.go.jp/stock.html). The CSSLs population and two parents were planted in three environments for GPC measurements, e.g. Yangzhou (Jiangsu Province, China, 32°24’N) in 2014 and 2015, and Linghuai (Hainan Province, China, 18°30’N) in 2014. For fine mapping of qGPC-10, the F\textsubscript{2} and F\textsubscript{3} populations were generated from a cross between SL425 and Sasanishiki, and cultivated in Yangzhou in summer seasons for GPC investigation in 2014 and 2015, respectively, with an interplant spacing of 10 × 25 cm for transplanting. Other materials used in this study are listed in Supplementary Data 1 and 2, including a core germplasm mainly from China (205 indica and 197 japonica cultivars) for GPC investigation.

**Map-based cloning.** A total of 790 F\textsubscript{2} plants with extreme phenotypes derived from the cross of SL431 and Sasanishiki were used for genetic analysis and mapping of qGPC-10 on chromosome 10. An additional 1085 F\textsubscript{3} individuals were then used for screening recombinants for fine mapping of qGPC-10. The genomic DNA fragments corresponding to the candidate gene in SL431 and Sasanishiki were sequenced and analyzed. A list of the markers used for QTL analysis and positional cloning is given in Supplementary Table 9.

**RNA extraction, CDNAs preparation, and quantitative reverse transcriptase PCR.** Total RNAs were extracted from rice various tissues using the Trizol reagent (Invitrogen). Total RNA was used for synthesizing first-strand cDNA with a reverse transcription kit (TOYOBO). Quantitative reverse transcription (qRT-PCR) was performed on an ABI 7500 instrument using SYBR Green PCR Mastermix followed the manufacturer’s instructions (Takara, RR20). All assays were performed with at least three biological replicates. Rice TUBULIN gene serves as an internal control to normalize gene expression. A list of primers used for qRT-PCR is given in Supplementary Tables 9 and 10.

**Vector constructions and plant transformation.** Various DNA fragments were amplified and inserted into different binary vectors for plant transformation. The information of relevant primer sequences was given in Supplementary Table 9. To further complement the SGC-Pi in SL431 (Hpla), we fused a 1 kb promoter fragment and the whole-genomic sequence of OsGluA2 by PCR amplification from SL431 and then inserted them into the plant binary vector pCAMBIA1301. To prepare the construct for CRISPR-Cas9 (KO-OsGluA2), a 20 bp fragment with NGG (PAM) at 3’ end from the second exon of OsGluA2 was introduced into the vector pCAMBIA1301. To confirm the fragments of OsGluA2 were amplified from a cross between Sasanishiki and SL431, respectively, and then were inserted into the expression vector pCAMBIA1391Z-GUS, namely GPC-10\textsuperscript{L-GUS} and GPC-10\textsuperscript{H-GUS}. The

**ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-019-09919-y | www.nature.com/naturecommunications**
complementation and GUS constructs were introduced into the rice cultivar Sasanishiki and the KO-OsGluA2 construct was introduced into the CSSL line SL431. All the resulting constructs were introduced into Agrobacterium tumefaciens strain EHA105 and transferred by Agrobacterium-mediated transformation into receipts.

Analysis of GUS gene expression. For histochemical analysis, rootlets, leaves, leaf sheaths, and stalks cut into 5 mm sections and maturing seeds at 5, 10, 15, 20, and 25 DAF sectioned longitudinally with a razor blade were subject to incubation in a solution with 10% methanol, 0.5% Triton X-100, 50 mM NaPO4, and 1 mM X-gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid), at 37°C for 12 h under the dark condition. After GUS staining, chlorophyll was removed using 75% ethanol before taking pictures. Fluorometric assays of GUS activities were conducted according to Kawakatsu et al., with maturing seeds at 15, 20, and 25 DAF used as samples. For each construct, ten independent transformants were subjected to fluorometric GUS assays.

Transmission electron microscopy analyses. Transverse sections (1 mm thick) of immature 15 DAF seeds from wild-type and transformants were fixed for 12 h in 0.2 M phosphate buffer (pH 7.3) containing 2.5% glutaraldehyde at 4°C. The sections were post fixed on ice for 1 h in 1% osmium tetroxide in 0.1 M phosphate buffer, dehydrated in a graded ethanol series, and embedded in Epon 812. The specimens were sliced into ultrathin sections by using an ultramicrotome (Leica EM, Austria) with a diamond knife (Diatome, Switzerland), and stained with a solution of uranyl acetate and lead citrate. Microscopic observation was performed using a transmission electron microscope (CM100, Holland) at an accelerating voltage of 100 kV. The area of PBs in each sample was determined using ImageJ software (NIH).

Amino acid analysis. For total amino acid analysis, 10 mg of rice power of each sample was hydrolyzed with 1 mL of 6 N HCl (Sigma, USA) in a 2 mL screw-cap tube before adding 10 mmol l-(-)+norleucine (Wako Pure Chemicals, Japan). The samples were then heated at 110°C for 24 h, followed by the treatment of 6 h at 65°C in order to evaporate HCl completely. The residue was then dissolved in 1 mL Na2S2O3 buffer and centrifuged at 16000 g for 10 min at room temperature. The supernatant was filtered with a 0.45 μm nylon membrane syringe filter ( Pall Life Sciences, USA) and transferred to an autosampler bottle for amino acid analysis. HPLC data were normalized with the level of l-(-)+norleucine per sample. Three biological replicates were designed for each sample. Seventeen amino acids were measured, including alanine (Ala), arginine (Arg), aspartic acid (Asp), cysteine (Cys), glutamic acid + glutamine (Glu), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tyrosine (Tyr), and valine (Val).

Transient expression assays in rice protoplasts. A series of mutated promoters of OsGluA2 were cloned into the pGreenII 0800-LUC vector43, and a Renilla luciferase gene was used as an internal transformation control to provide an estimate of the extent of transient expression in the same construct. Previous studies showed that the RISBZ1 (rice rice b-Zipper 1) regulated, factor expressing with rice SSP genes coordinately, and activates transcription from SSP gene promoters in transient expression assays44–46. The full-length cDNAs of RISBZ1 was fused into pCambia1300-221-Flag vector. The different combinations of vectors were co-transfected into rice protoplasts47 by PEG-mediated transformation. Ratios of LUC to Ren activity were calculated to define relative promoter activity. Ten biological replicates were designed. Relevant PCR primer sequences were provided in Supplementary Table 9.

Nucleotide diversity and evolutionary analyses. The genomic sequences of 3005 cultivated and 27 wild accessions were obtained from Rice Functional Genomics and Breeding Database27,29 (RFG. http://www.rmbreeding.cn/Snp3k) and OryzaGenome68 (http://viewer.shigen.oryzagenome/), respectively. The distribution of two functional types of OsGluA2 in different cultivated rice groups was compared using chi-square (χ2) test. The geographical information of cultivated rice groups was acquired from RFGB, and was marked on map using R software to observe geographic distribution characteristics of two types. The parameters of genetic difference between indica and japonica, including haplotype and nucleotide diversity, Nei’s Dst, and Hudson’s DST were calculated within per 2 kb regions using PopGenome package in R software48. The estimates of nucleotide diversity (θ) and Tajima’s D for each rice group in OsGluA2 and flanking regions (±80 kb) were calculated using DnaSP V5 (ref. 50). In addition, nucleotide diversity curves were acquired using 2000 bp window and 100 bp step length. In order to evaluate the evolution of OsGluA2, a phylogenetic tree and haplotype network were constructed using all the polymorphic sites with minor allele frequency ≥0.005 in this locus. The phylogenetic tree for OsGluA2 was constructed using UPGMA method by Mega7.0 (ref. 51), while the haplotype network was built by pegas package52.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the findings of this study are available within the paper and its Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. The datasets generated and analyzed during the current study and plant materials are available from the corresponding author upon reasonable request. The source data underlying Figs. 2a-c, 2e–h, 3a, 3c, 3d, 3f, and 4d–f, Table 1, Supplementary Figures 1b, 1d, 2c, 3c, 4, 5, and 6, as well as Supplementary Table 7 are provided as a Source Data file.

Received: 11 October 2018 Accepted: 5 April 2019
Published online: 26 April 2019

References

1. Sasaki, T. & Burr, B. International rice genome sequencing project: the effort to completely sequence the rice genome. Curr. Opin. Plant Biol. 3, 138–141 (2000).
2. Umemoto, T. Genes affecting eating and processing qualities. Rice Genetics Genet. Breed. 21, 417–434 (2018).
3. Hori, K. Genetic dissection and breeding for grain appearance quality in rice. Rice Genomics Genet. Breed. 22, 435–451 (2018).
4. Martin, M. & Fitzgerald, M. A. Proteins in rice grains influence cooking properties. J. Cereal Sci. 36, 285–294 (2002).
5. Nakamura, Y. Towards a better understanding of the metabolic system for amylopectin biosynthesis in rice endosperm as a model tissue. Plant Cell Physiol. 43, 718–725 (2002).
6. Yan, C. J. et al. Genetic analysis of starch paste viscosity parameters in glutinous rice (Oryza sativa L.). Theor. Appl. Genet. 122, 63–76 (2011).
7. Tian, Z. et al. Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. Proc. Natl. Acad. Sci. USA 106, 21760–21765 (2009).
8. Fitzgerald, M. A. et al. Not just a grain of rice: the quest for quality. Trends Plant Sci. 14, 133–139 (2009).
9. Kawakatsu, T. et al. Characterization of a new rice glutelin gene GluD1 expressed in the starchy endosperm. J. Exp. Bot. 59, 4233–4245 (2008).
10. Saito, Y. et al. Formation mechanism of the internal structure of type I protein bodies in rice endosperm: relationship between the localization of prolamin species and the expression of individual genes. Plant J. 70, 1043–1055 (2012).
11. Usaba, M. et al. Low glutelin content1: a dominant mutant that suppress the glutelin multigene family via RNA silencing in rice. Plant Cell 15, 1445–1467 (2003).
12. Hamaker, B. R. & Griffin, V. K. Effect of diastase bond-containing protein on rice starch gelatinization and pasting. Cereal Chem. 70, 377–380 (1993).
13. Kawakatsu, T. & Takaiwa, F. Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. Plant Biotechnol. J. 8, 939–953 (2010).
14. Lin, R. et al. in Rice Germplasm Resources in China (ed. Ying, C.) 83–93 (Agricultural Science and Technology, Beijing, 1993).
15. Zhou, L. H. et al. Variation and distribution of seed storage protein content and composition among different rice varieties. Acta Agron. Sin. 35, 884–891 (2009).
16. Shenoy, V. V. et al. Inheritance of protein per grain in rice. Indian J. Genet. 51, 214–220 (1991).
17. Shi, C. et al. Genetic analysis for protein content in indica rice. Euphytica 107, 135–140 (1999).
18. Terao, T. & Hirose, T. Control of grain protein contents through SEMIDWARF1 mutant alleles: sd1 increases the grain protein content in Dee- geo-woo-gen but not in Reimei. Mol. Genet. Genomics 290, 939–954 (2015).
19. Ren, Y. et al. GLUTELIN PRECURSOR ACCUMULATIONS encodes a regulator of post-Golgi vesicular traffic essential for vacuolar protein sorting in rice endosperm. Plant Cell 26, 410–425 (2014).
20. Tian, L. et al. Small GTPase Sert is crucial for proglutelin and alpha-globulin export from the endoplasmic reticulum in rice endosperm. J. Exp. Bot. 64, 2831–2845 (2013).
21. Wang, G. et al. Opaque7 encodes an acyl-activating enzyme-like protein that affects storage protein synthesis in maize endosperm. Genetics 189, 1281–1295 (2011).
22. Wang, Y. et al. OsRal5a regulates endomembrane organization and storage protein trafficking in rice endosperm cells. Plant J. 64, 812–824 (2010).
23. She, K. C. et al. A novel factor FLOURY ENDSOMER2 is involved in regulation of rice grain size and starch quality. Plant Cell 22, 3280–3294 (2010).
24. Wang, Y. et al. The vacuolar processing enzyme OsVPE1 is required for efficient glutenin processing in rice. **Plant J.** 58, 606–617 (2009).

25. Takemoto, Y. et al. The rice mutant esp greatly accumulates the glutenin precursor and deletes the protein disulfide isomerase. **Plant Physiol.** 128, 1212–1222 (2002).

26. Yano, K. et al. Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. **Nat. Genet.** 48, 927–934 (2016).

27. Xu, F. et al. Genome-wide association study of eating and cooking qualities in different subpopulations of rice (*Oryza sativa L.*). **BMC Genomics** 17, 663 (2016).

28. Yang, Y. et al. Identification of quantitative trait loci responsible for rice grain protein content using chromosome segment substitution lines and fine mapping of qPC-1 in rice (*Oryza sativa L.*). **Mol. Breed.** 35, 1–9 (2015).

29. Cheng, L. et al. Identification of stably expressed quantitative trait loci for grain yield and protein content using recombinant inbred line and reciprocal introgression line populations in rice. **Crop Sci.** 53, 1436–1446 (2013).

30. Zheng, L. et al. Genetic relationship between grain chalkiness, protein content, and paste viscosity properties in a backcross inbred population of rice. **J. Cereal Sci.** 56, 153–160 (2012).

31. Zheng, L. et al. Dynamic QTL Analysis of rice protein content and protein index using recombinant inbred lines. **J. Plant Biol.** 54, 321–328 (2011).

32. Liu, X. et al. Dissecting the genetic basis for the effect of rice chalkiness, amylose content, protein content, and rapid viscosity analyzer profile characteristics on the eating quality of cooked rice using the chromosome segment substitution line population across eight environments. **Genome** 56, 64–80 (2011).

33. Ye, G. et al. QTL mapping of protein content in rice using single chromosome segment substitution lines. **Theor. Appl. Genet.** 121, 741–750 (2010).

34. Zhao, K. et al. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. **Nat. Commun.** 2, 467 (2011).

35. Peng, B. et al. OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. **Nat. Commun.** 5, 4847 (2014).

36. Zhang, H. et al. A core collection and mini core collection of *Oryza sativa L.*. **Mol. Breed.** 32, 49–61 (2011).

37. Wang, W. Sh. et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. **Nature** 557, 43–49 (2018).

38. Luo, H. et al. Up-regulation of OsBHD1, a rice gene encoding BELL homeodomain transcriptional factor, in disease resistance responses. **Plant Biol.** 7, 459–468 (2005).

39. Sun, C. et al. RPAN: Rice pan-genome browser for ~3000 rice genomes. **Nucleic Acids Res.** 45, 597–605 (2017).

40. Kawakatsu, T. et al. Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. **Plant Physiol.** 154, 1842–1854 (2010).

41. Lee, H. J. et al. Lack of Globulin synthesis during seed development alters accumulation of seed storage proteins in rice. **Int. J. Mol. Sci.** 16, 14717–14736 (2015).

42. Perbandt, D. et al. Assessment of mass flows and fuel quality during mechanical dehydration of silages using near infrared reflectance spectroscopy. **Bioenergy Res.** 3, 194–203 (2010).

43. Hellens, R. P. et al. Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. **Plant Methods** 1, 13–13 (2005).

44. Ondrera, Y. et al. An rice functional transcriptional activator, RISBZ1, responsible for endosperm-specific expression of storage protein genes through GCN4 motif. **J. Biol. Chem.** 276, 14139–14152 (2001).

45. Yamamoto, M. P. et al. Synergism between RPBF DoF and RISBZ1 BZIP activators in the regulation of rice seed expression genes. **Plant Physiol.** 141, 1694–1707 (2006).

46. Kawakatsu, T. et al. Compensation and interaction between RISBZ1 and RPBF during grain filling in rice. **Plant J.** 59, 908–920 (2009).