Indian J. Genet. Plant Breed., (2022) 82(1): 16-24
doi: 10.31742/IJGPB.82.1.3
ISSN: 0975-6906
www.isgpg.org

RESEARCH ARTICLE

Genetic analysis and gene mapping of *bgl,* a gene controlling grain length and quality in rice

Mustapha Sunusi#, Asadullah Khan, Yohannes Tsago, Ziyan Chen, Chunhai Shi and Xiaoli Jin*

Abstract

Grain size is one of the important agronomic traits that determine rice yield. Map-based cloning of rice grain size-related genes impacts the genetic regulation mechanism of rice yield formation. Here we identified *bgl* mutant in rice which exhibits increase in grain length and panicle length with the decrease in number of grains per panicle and no observable difference in grain thickness and number of tillers as compared to the wild type (WT). Scanning electron microscopy observation of *bgl* mutant spikelets hull indicates an upsurge in number of external epidermal cells in a longitudinal way than the WT. Hence, it is likely that the long-grain phenotype was due to increased cell number longitudinally. Starch granule in both *bgl* and WT were large, sporadically polyhedral and densely packed, indicating no differences, signifying that both *bgl* mutant and WT have less chalkiness content. Genetic analysis showed that the *bgl* mutant was controlled by a recessive single gene, and was finely located on chromosome 3 through map-based cloning, within a physical distance of 40kb. Sequencing analysis revealed A to T substitution in the seventeen exons of *LOC_Os03g44500* resulted in an amino acid change from Tryptophan (Try) to Phenylalanine (Phe)**.** Taken together *bgl* mutant has a pleiotropic effect on grain yield and grain quality.

Keywords: Gene mapping, grain shape, grain quality, map based cloning, rice

Introduction

Rice plays an important role in food security because it is the major staple food crop that serves nearly half of the world's population, particularly in Asia and Africa, and also it is a primary source of income and employment. Grain size/shape/weight is determined by its three geometrical dimensions (grain length, width and thickness) and the degree of grain filling (Xing and Zhang 2010). Grain size is an important factor and indicator for grain quality appearance. Among them, grain length has a significant positive correlation with the thousand-grain weight (Huang et al. 2013). Grain length is a quantitative trait controlled by multiple genes (Qian et al. 2016). So far, rice grain length genes have been finely mapped and cloned but the genetic and molecular mechanisms of grain size control remain largely unknown. among them *GS3* (Fan et al. 2006), *GL3.1* (Qi et al. 2012), *GL3.3* (Hu et al. 2018), *LG3* (Xiong et al. 2018), and *LG3b* (Yu et al. 2018), were all mapped in chromosome 3. In addition, multiple signaling pathways were reported to regulate grain size, such as G-protein, ubiquitin-proteasome, mitogen-activated protein, transcriptional factor and phytohormones (Li et al. 2019; Zuo and Li 2014). Moreover, grain size is one of the most important target traits of grain yield, appearance quality, domestication and breeding. Thus, understanding the regulatory mechanisms of gain size underlying natural variations has recently become an important research field in agricultural science. Numerous studies have been completed on grain size regulation in rice, and many pointed out that grain size is determined by cell proliferation or cell expansion (Horiguchi et al. 2006; Horvath et al. 2006). Among them, *qGL3/GL3.1* is a grain weight and length QTL located on chromosome 3. Two

Key Laboratory for Crop Germplasm Resource. College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China.

#Present address: Department of Crop Science, Faculty of Agriculture, Federal University Dutse, Nigeria.

Corresponding Author: Xiaoli Jin, Key Laboratory for Crop Germplasm Resource, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China, E-Mail: jinxl@zju.edu.cn

How to cite this article: Sunusi M., Khan A., Tsago Y., Chen Z., Shi C. and Jin X. 2022. Genetic analysis and gene mapping of *bgl*, a gene controlling grain length and quality in rice. Indian J. Genet. Plant Breed., **82**(1): 16-24.

Source of support: National Key Research and Development and the Science and Technology Office, China

Conflict of interest: None.

Received: July 2021 **Revised:** Dec. 2021 **Accepted:** Jan. 2022

research teams have completed the cloning of this gene (Qi et al. 2012; Zhang et al. 2012). These genes are located within the range of 20kb and 47kb, respectively. *GL3.1* encodes serine/threonine protein phosphatase *OsPPKL1*, which increase cell division by affecting protein phosphorylation in spikelets, resulting in longer grains (Zhang et al. 2012). *OsPPKL1* contains two Kelch domains, of which the rare allelic variation *qgl3* in AVLDT conserved region located in second Kelch domain results in a long-grain phenotype (Hu et al. 2012). By sequencing 94 varieties only 2 varieties carried long-grain alleles of *GL3.1/qGL3*, indicating that *GL3.1/qGL3* is a rare allele (Zhang et al. 2012). *TGW3/qTGW3/GL3.3* is located on chromosome 3, which mainly controls grain weight and grain length. *GL3.3* encodes a protein kinase *OsGSK5/OsSK41* of the *GSK3/SHAGGY*-Like family. Functional loss of this gene increase grain length and grain weight and is a negative regulator of grain length and grain weight. The protein encoded by *TGW3* is located in the nucleus and cytoplasm. The tissue-specific expression shows that *TGW3* is mostly expressed in an immature panicle and can synergistically change the size and number of cells in the glume. Further research found that this gene and *GS3* have a genetic interaction effect, and the superposition of the two can cause a significant increase in rice grain size. The *GS3-TGW3* genotype combination is commonly used in the breeding of large-grain *japonica* rice varieties but the application of this combination has not been found in indica rice varieties, therefore, this locus can be used in cultivated *indica* rice varieties (Hu et al. 2018; Xia et al. 2018; Ying et al. 2018).Here we reported new long-grain *bgl* gene derived from *japonica* 02428 mutated with EMS. The present study was aimed at the examination of the morphological difference between *bgl* mutant and wild type; analyze grain quality parameters between the *bgl* mutant and WT and to identify the candidate gene for grain length through map-based cloning approach.

Materials and methods

Plant materials and growth conditions and traits measurement

The *bgl* mutant was isolated from the *japonica* variety 02428 mutagenized with EMS. Rice plants were grown in the experimental fields of Zhejiang University Hangzhou, China, under natural growth conditions. Plants from the middle row of *bgl* mutant and 10 from the WT were selected at the plant maturity stage. The main agronomic and yield traits were measured after four months of air-dried grain length, and the grain width was measured with a vernier caliper. At the same time, thousand grain weight (TGW) was evaluated by weighing 200 grains and repeated three times. The degree of chalkiness, amylose content, protein content, starch content, and alkali spreading values was measured, respectively, according to the national standard (NY/T593-2013) procedure.

Rapid visco analyzer RVA determination

The viscosity change of rice flour was measured with RVA Viscosity Tester (Perten, Sweden), and the parameters were set according to the AACC61-01 and AACC61-02. The first-level parameters read by the instrument were peak, hot paste, cool paste, pasting temperature, and peak time; second-level parameters are breakdown, setback, and consistency viscosity. Each sample is measured 3 times and the average value was taken.

Determination of protein content and protein fractionation

Crude protein was evaluated with Kjeltec 2300 Auto analyzer (Foss AB, Sweden). protein value was computed based on 6.25 nitrogen conversion factor, while rice grain protein fractions were measured according to (Kumamaru et al. 1988), with slight modifications. Briefly, 1.0g flour was successively extracted with 50mM phosphate buffer (pH 6.8) same buffer was used to measure albumin content, while buffer with 0.5 M NaCl was used for measuring globulin, and 60% *n*-propanol was used for evaluating prolamin. Furthermore, 0.1 M NaOH was used for measuring glutelin content. The sample was homogenate by adding 10ml solvent and kept at room temperature for 30min and centrifugation at 12,000 rpm for 10min to separate extract from residues.

Fatty acid composition of bgl mutants and wild-type

Samples were measured based on the modified method outlined by Verma and Srivastav (2017). With a little modification, briefly, 0.5g rice flour was measured and put in a 10ml glass tube, then added 8ml chloroform/isopropanol, 2:1,v/v as an extracting solution vortex vigorously on a shaker at room temperature for 2 hours in dark condition. Centrifugation of the samples tube at 3000 rpm at least 5 minutes, transferred 5 ml supernatant into another 10mL glass tube, and dehydrated under $N_{2.}$ Add 500 mL extraction solvent so that it will re-dissolved and gently vortex. Add 2 ml of 1% H_2SO_4/MeOH (v/v) into the samples tube and prepare the fatty acid methyl ester (FAME). After that, the extract was placed in gas-liquid chromatography (GLC) (Varian CP 3800, USA). And set the flame ionization detector (FID) along with a fused silica capillary column (50 mm × 0.25 mm), which was covered by CP-SIL 88. The oven temperature was set to be 200°C for at least 13 min. While injector and FID were at 250°C following by analyzing the standard reference in the same setting to decide the peak identity. Each fatty acid's final concentration was multiplied with 1.6 (5 ml of extraction solvent used out of the 8mL and divided by 500 (0.5 g of the flour used).

Cytological observations

Mature grains were observed under a Leica microscope (LEICA S8APO; Leica microsystems, Wetzlar, Germany) and photographed using a Leica CCD (DFC420). Grain length and width were measured using Image J software. Grain weight was determined by weighing 1000 dry grains using an electronic analytical balance (Mettler Toledo AL104, China). The weights of three replicates were measured for each grain lot. The size of epidermal cells in spikelet hulls were investigated using a scanning electron microscope. The samples were fixed in FAA solution (glacial acetic acid: formalin: 50% ethanol; 1:1:18) at 4°C overnight, dehydrated in a graded ethanol series, and substituted with 100% ethanol. The critical-point drier (HITACHI HCP-2) was used to dry samples. The samples were dissected under a microscope (LEICA S8APO; Leica microsystems, Wetzlar, Germany), sputter-coated with platinum and observed using a scanning electron microscope (HITACHI S-3000 N; Hitachi High-Technologies Corporation, Tokyo, Japan). Cell size was measured using Image J software. All samples were processed and observed at Zhejiang University.

Gene mapping and candidate gene prediction

The total DNA of each F_2 plant together with the mutant and WT leaf at three week after planting was extracted with CTAB method respectively an equal amount was polled according to the principles of bulk segregant analysis (BSA). A set of 300 polymorphic SSR and 50 InDel markers evenly distributed on 12 rice chromosomes available in our laboratory were used for linkage analysis and preliminary mapping according to Gramene (http://www.gramene.org) BLAST (http://blast.ncbi.nlm.nih.gov), align *indica/japonica* whole genome sequence to find insertion/deletion site, use Primer Premiers 5.0 software to design primers to finely map the mutant gene listed in (Supplementary Table S1). The PCR reaction system is 10 μL, and the components include: 1-μL 10x PCR Buffer (Mg 2+ Plus), 0.1-μL dNTPs (2.5 mM), m1-μL primer (10 μM), 1-μL DNA template, 0.1-μL Taq enzyme (5U/μL), add ddH_2O to make up to 10 μL. And PCR program: pre-denaturation at 94°C for 5 minutes, 94°C for 30s, 55°C for 30s, 72°C for 30s, 35 cycles, and a final extension at 72°C for 4 minutes. The product after electrophoresis was detected on 8% polyacrylamide gel and silver staining.

Gene expression analysis

Total RNA from leaves and other tissues at different stages of growth were extracted using TRIZOL and RNA concentration was checked by agarose gel electrophoresis and micro spectrophotometer (NanoDrop 2000, Prime Script RT reagent kit with gDNA eraser and SYBR Premix Ex Taq II were used for gene expression. The reaction system was 20μL, and it was performed on a fluorescent quantitative PCR machine (Light Cycler 480 Roche, Switzerland). PCR amplification program: 95°C 5 min; (95°C 10s, 60°C 30 s)×45 cycles; dissolution curve reaction conditions: 95°C 30 s, 65°C 1-min, 95°C continuous, and cooling at the end of the reaction: 40°C for 30s. After the PCR is completed, *OsActin* was used as reference gene, the relative amount of gene expression was calculated using the 2–△△CT method. Primers are listed in (Supplementary Table S2).

Results

Agronomic and yield traits comparison of bgl and WT

The phenotypic characteristics of the bgl mutant plant architecture were upright with compact tillers, the grains are long and the grain color is sandy beige. The agronomic and yield traits display an increase in the yield traits such as grain length, grain length-width ratio, thousand grain weight and panicle length, with a decrease in grain per panicle as compared to the WT. In addition, there is no significant difference between the *bgl* mutant and WT in grain thickness, grain width, and in number of tillers (Fig. 1a-d; Table 1).

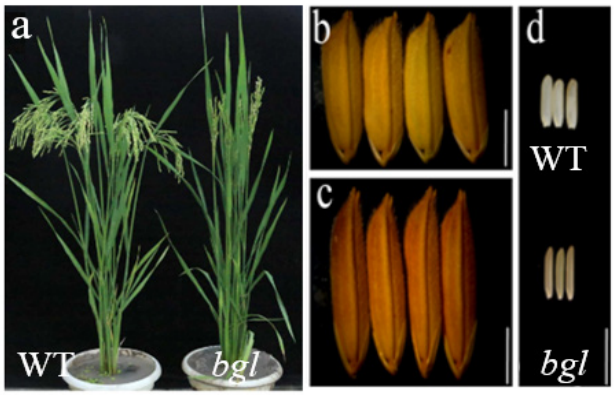

Fig. 1. Phenotypic characterization of bgl mutant with WT plants at anthesis, scale bar, 10 cm; (b) WT full-grown grains, scale bar, 1mm; (c) bgl full-grown grains, and (d) Fully de-husked grains

Table 1. Comparison of grain yield and other agronomic traits between *bgl* mutant and WT

	WT	*bgl*
Grain length (mm)	10.20 ± 0.12	12.2 ± 0.18**
Grain width (mm)	2.62 ± 0.05	2.64 ± 0.05
Grain thickness (mm)	2.17 ± 0.03	2.34 ± 0.02
Grain length/width ratio	3.89 ± 0.11	4.62 ± 0.32**
Thousand grain weight (g)	30.04 ± 1.28	39.06 ± 1.21**
Panicle length(cm)	24.51 ± 1.59	26.39 ± 2.27*
Grain per panicle	217.1 ± 14.2	211.3 ± 15.1**
Plant height (cm)	111.4 ± 4.50	113.1 ± 1.42*
Number of tillers	8.69 ± 1.54	8.71 ± 1.22
Yield per plot (kg)	12.87 ± 0.41	13.92 ± 0.11*

* and ** indicate values are significant at 0.05 and 0.01, respectively

Morphological and histological analysis of bgl mutant

To disclose the basis underlying cellular changes in grain size, we examined the cell size and cell number in the *bgl* mutant and WT spikelet by SEM. The number of external epidermal cells observed in grain longitudinal way was considerably higher in bgl mutant than in the WT. Hence, it is likely that the long-grain phenotype of the *bgl* mutant was due to increasing the cell numbers longitudinally, which is evident with the fact that *bgl* mutant grains are longer than the WT (Fig. 2a-j). We also measure the relative expression levels of genes known to regulate grain length due to cell cycle and cell expansion. We found that most of the genes that are known to control cell expansion such as *OsEXPA5*, *OsEXPA10*, *OsEXPB3*, *OsEXPB4*, and *OsEXPB7*, were down-regulated when compared with *bgl* mutant, which failed to offer comprehensible evidence about *bgl* mutant regulation due to cell expansion. However, cell cycle genes such as *CYCA2:1, CYCA2:2, CYCB2:2, CAK1, CDCD7.1* and *H1* examined were significantly up-regulated in *bgl* mutant panicle as compared with the WT (Fig. 3c). Supporting these results, expression of cell cycle genes was increased, particularly in young panicles, therefore, increased grain length might likely due to cell proliferation in the spikelet hull in both latitudinal and longitudinal directions.

Genetic analysis and map-based cloning of BGL

F_1 plants were obtained from crossing *bgl* mutant and WT, which exhibited the WT phenotypes. In F_2 population, 274 WT and 94 *bgl* mutant plants that were obtained which fit 3:1 Mendelian segregation ratio ($\chi^2 = 0.058 < \chi 02.05 = 3.84$) indicating single recessive gene controlled *bgl* mutant. To identify the location of *bgl*, map-based cloning in the F_2 population resulting from the cross between *bgl* mutant and WT was used. The *BGL* gene was mapped primarily on the long arm of chromosome 3 between simple sequence repeat (SSR) markers RM16 and M3B15. Based on *Indica-Japonica* genome sequences. Ten new *InDel* markers were designed for map-based cloning. The *bgl* gene was fine mapped to an interval of 54kb between *InDel* markers M3B16 and M3B12 using 456 F_2 mutant individuals. Eight putative genes were predicted within the region according to Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/). Afterward, genomic DNA for the *bgl* mutant and WT were used to sequence and analyzed all the genes in the predicted region (Supplementary Table S3). There was no diversity in all the genes except the *LOC_Os03g44500* gene. For this reason, we chose *LOC_Os03g44500* as a candidate gene for further sequence analysis.

DNA sequencing analysis of the candidate region reveals A to T substitution in the seventeen exons of *LOC_Os03g44500,* resulting in an amino acid change from tryptophan to phenylalanine. The genomic sequence of the *bgl* gene consists of twenty-one exons and one intron. The *BGL* encoded protein phosphatase with kelch-like repeat domain, serine/threonine phosphatase. The whole sequence length was 3631 bp and the CDS length was 3012bp and the amino acid was 1004.

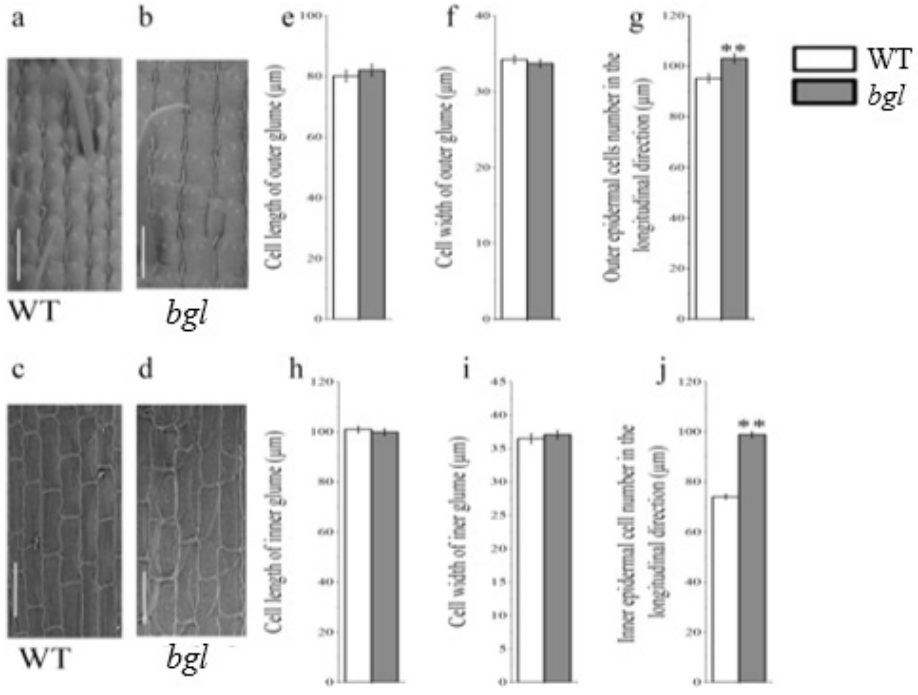

Fig. 2. SEM of *bgl* spikelet hull: (a-b) Outer glume cells of *bgl* mutant and WT.; (c-d) Inner glume cells of *bgl* mutant and WT. Scale bars, 5 mm; (e) Outer epidermal cells length; (f) Outer epidermal cells width; (g) Outer epidermal longitudinal cells; (h) Inner glume cells length; (i) Inner glume cells width, and (j) Longitudinal inner glume cell number. *$P < 0.05$, **$P < 0.01$, Student's *t*-test

BGL gene expression analysis

To verify the spatio-temporal expression of genes using quantitative real-time PCR was performed. The expression of *bgl* gene was observed relatively higher in developing panicle than other organs i.e., roots, stems and leaves. Predominantly, the expression was progressively high when the panicle size was 4–6 cm length stage which is subsequent to the anthesis

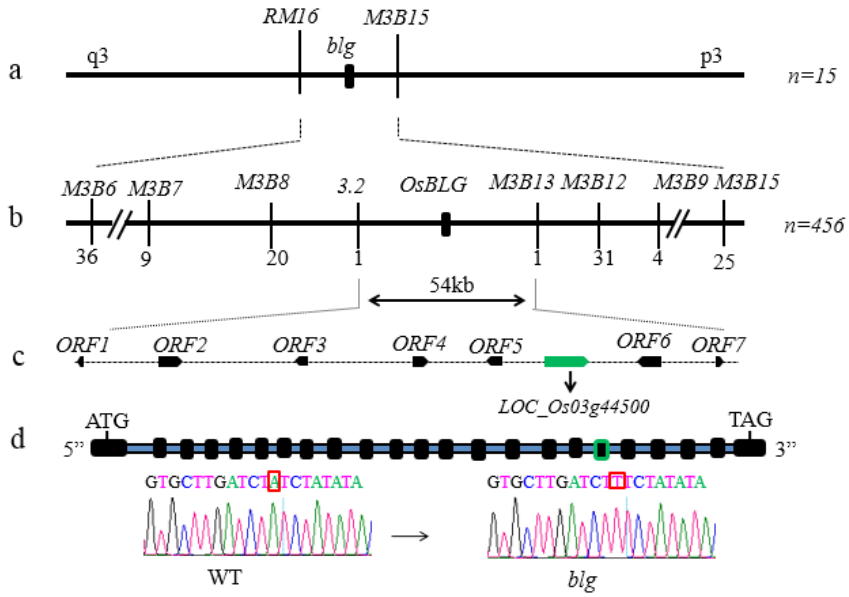

Fig. 3. Mapping of the *bgl* gene: (a) Primary mapping of region; (b) Mapped region distance (c) Distribution of eight ORF in the mapping region, and (d) Single based substitution of A-T of *LOC_Os03g44500*

Table 2. Rapid visco analyzer profile of *bgl* mutant and WT

RVA characteristics	*WT*	*bgl*	*t- value*
Peak viscosity (RVU)	210.00	212.42	1.479
Trough viscosity (RVU)	170.75	172.08	0.455
Breakdown (RVU)	39.25	40.33	0.469
Final viscosity (RVU)	304.33	305.71	0.957
Setback (RVU)	94.33	93.29	0.369
Consistence (RVU)	133.58	133.63	0.044
Peak viscosity time (min)	6.60	6.57	0.522
Pasting temperature (°C)	81.60	81.95	0.449

phase. Subsequently, it started to decline as the panicle matures, consistent with the gene biological function of regulating grain length (Fig. 4a-b.)

Comparison of grain quality evaluation of bgl mutant

Grain quality traits were compared between the *bgl* mutant and the WT. The amylose content and gel consistency in *bgl* mutant grains were significantly lower than WT grains. However, alkali value in *bgl* mutant was higher. However, there was no noticeable difference about chalky grain and total protein content between the *bgl* mutant and the WT (Fig. 5a-d) .

Morphological examination of the starch structure and evaluation of RVA viscosity between of bgl mutant and WT

Detailed SEM examination of the *bgl* mutants and WT was conducted and the result showed that both the *bgl* mutants and WT endosperm were large, sporadically polyhedral, and densely packed, which indicate no differences in the features signifying that both the WT and the mutant has less chalkiness content (Fig. 6 a-f). Comparative results of RVA show no significant difference between the *bgl* and the WT grains on the RVA indices measured such as trough viscosity, peak viscosity time, and pasting temperature. (Table 3).

Performance of four protein fraction components and the comparison of fatty acid composition of bgl mutant and WT

Performance of four protein fractions was observed and the result shows a slightly high content of glutelin, prolamin, and globulin contents were considerably higher in *bgl* mutant but the significance was not remarkable based on the t-test. Glutelin content in the *bgl* mutant was considerably decreased. Taken together the consequence, it was demonstrated that the chalkiness had little effect on prolamin, globulin, and albumin, contents, with a huge influence on glutelin contents. Signifying that the high content of chalkiness negatively affects the nutritional quality in *bgl* mutant and the WT to some extent. Comparative results for fatty acid revealed that the concentration of linoleic acid, palmitic acid, stearic acid and oleic was high in *bgl* mutants compared with WT while arachidic, and, docosanoic acid concentration was less and insignificant in both WT and the *bgl* mutant (Fig. 7a-b).

Table 3. Difference in protein fraction components of *bgl* mutant and WT

	Albumin (mg/g)	Globulin (mg/g)	Prolamin (mg/g)	Glutelin (mg/g)
WT	0.344 ± 0.053	0.638 ± 0.027	0.712 ± 0.101	5.481 ± 0.044
bgl mutant	0.339 ± 0.026	0.605 ± 0.037	0.648 ± 0.036	5.378 ± 0.031
t- value	0.143	1.249	1.032	3.335*

* Significant at 0.05 levels

Discussion

Although improving grain size and grain yield are the main goals of improvement. Several mutants are known to produce long-and slender grain, but their mechanism is still uncertain. Because of this, more planting materials from diverse genetic backgrounds need to be exploited, to enhance the insightful molecular mechanisms. Here, sandy beige grain length *bgl* mutant was isolated from *indica* 9311 and mapped on chromosome 3. The *bgl* gene

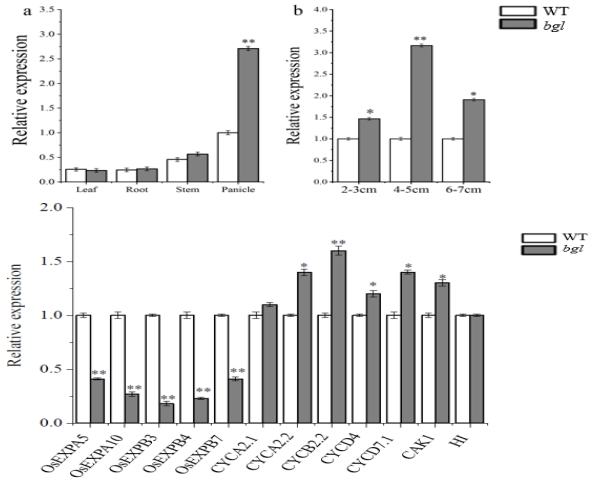

Fig. 4. Relative expression of BGL genes: (a) Relative expression of BGL genes in different tissues; (b) Relative expression of BGL genes at different stages of panicle growth, and (c) Relative expression of cell cycle and cell expansion gene. Data are presented as mean ± SD. Student's t-test was used to generate the P values; * P < 0.05, ** P < 0.01, respectively

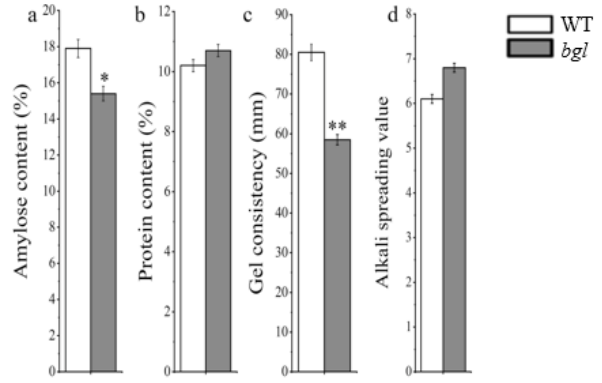

Fig. 5. Grain quality traits measurement: (a) amylose content; (b) protein content; (c) gel consistency and (d) alkali spreading value. * and ** show significance at P< 0.05 and P< 0.01 t-test

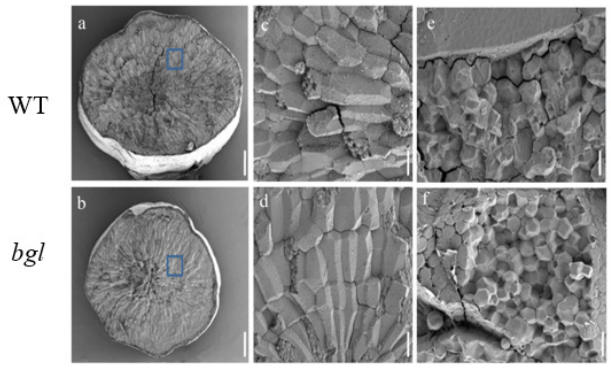

Fig. 6. SEM of starch granules bgl and WT: (a-b) Traverse piece of polished rice, scale bars, 100µm; (c-d) Starch granules within the inner endosperm cell, scale bars, 3µm. The rectangular box is the location of where the micrograph was taken; (e-f) The Starch granule contained in endosperm cells, scale bars, 3µm. The square box shows the position of imagery

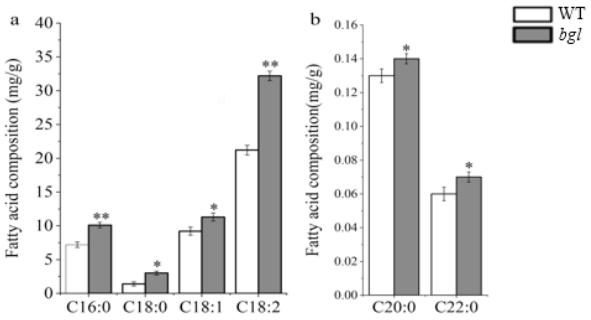

Fig. 7. Fatty acid composition of bgl mutant and WT: (a) Saturated fatty acid palmitic acid, stearic acid; oleic acid; linoleic acid, and (b) Unsaturated fatty acid arachidic acid; docosanoic acid. All the values are presented as means ± SD whereas the ** indicate significant difference

encodes protein phosphatase with Kelch-like repeat domain, serine/threonine phosphatase, which is grain length and yield regulator.

To date, almost all the genes recognized as well as the QTLs which persuade grain size regulation is either due to cell proliferation or cell expansion of spikelet hull consistency with these findings, the cytological extermination of grain spikelet hull, the number of external epidermal cells observed in grain longitudinal way was considerably high in *bgl* mutant than the WT. Hence, it is likely that the long-grain phenotype of the *bgl* mutant was due to increased cell number longitudinally, which is dependable because *bgl* mutant grains are longer than the WT. In other reports, for instance, *GS3*, *qGL3/GL3.1*, and *BG1* control grain size due to increased cell number in spikelets hull (Mao et al. 2010; Qi et al. 2012). *GL7* has similarity with *Arabidopsis*, high expression of *GL7* results in producing slender grains due to cell length increase longitudinally and decreasing cell size transversely (Wang et al. 2015a; Wang et al. 2015b).

Chalky rice grain differed with transparent grain in terms of morphological, physicochemical, textural, and thermal characteristics (Kim et al. 2000; Singh et al. 2003). As seen in the current study, the Microscope examination signifies that the shape of starch granule in the transparent part is a habitual polygon with few inter-granule gaps, whereas the chalky portion is irregular with the loose arrangement. The considerable disparity in starch granule structure, shape,

and arrangement consequences in chalkiness generation was recorded. The result shows less starch density in the chalky part of the grains than the transparent part. This will possibly explain the reason why chalky grains have less amylose content compared with transparent grains (Lisle et al. 2000; Singh et al. 2003).

Amylose content (AC) plays an imperative position in shaping cooking, eating and pasting characteristics like consistency, taste, stickiness, grain elongation, hardness, gel consistency, along with gelatinization temperature in rice (Adu-Kwarteng et al. 2003; Pooni et al. 1993). Besides amylose content, other factors were also reported to influence rice cooking quality viz., proteins and amylopectin (Cai et al. 2011; Pooni et al. 1993). In *bgl* mutant, low amylose content was observed compared with the WT (Fig. 5a.), which is consistent with the earlier reports (Fig. 5a-d.). These results signify that differences in rice grain texture occur in both the medium and long-grain cultivars. Thus, intermediary or low GT is preferred to be a good quality variety. Similar characteristics were observed in the *bgl* mutants identified by us with intermediary GT (Fig.5a-d). Understanding the relations between AAC, GT, and pasting properties will be essential to describe ECQ of many rice cultivars and have considerable influence on grain quality breeding (Bao et al. 2006; Chen et al. 2008; Xu et al. 2015).

Earlier reports elucidated that CEQ in rice is always related to the RVA characteristic (Dong et al. 2006). RVA breakdown is caused by the interruption of gelatinized starch granule composition, or due to the differentiation of gelatinized starch granules and the viscosity, the disruption can be complete or partial (BeMiller 2019; Han and Hamaker 2001). RVA profile is an indirect indicator of rice-eating quality and viscosity is used to find a particular feature in rice cultivar. Different RVA profiles represent diverse characteristics of rice cooking for instance, high setback viscosity with small breakdown viscosity is an indicator that the brown rice will be very hardened when cooked. While higher breakdown viscosity and small setback viscosity indicate that the cooked rice will be very soft and sticky (Shu et al. 1998). A similar result was obtained in *bgl* mutant with high peak viscosity and break down viscosity with low set back viscosity than the WT, which indicates that the *bgl* mutant is much softer than the WT.

High concentrations of protein fraction component, albumin, globulin, glutelin with prolamin separately, or in combination has a significant effect on RVA characteristics (Baxter et al. 2004; Baxter et al. 2014; Baxter et al. 2010). The globulin content is low in bgl mutant, which suggests no connection with RVA. This may be because globulins were synthesized in rice grain early development (Yamagata et al. 1982). likely globulin takes part in maintaining protein formation(Martin and Fitzgerald 2002), which is the reason why its contents are very small. Glutelin was highly correlated with paste than total protein content, and each rice grain protein component has a distinctive role in grain texture (Paula and Conti-Silva 2014). Prolamin is unsuccessfully digested by human beings (Kubota et al. 2010), therefore rice with low prolamin content will have soft texture and be more nutritious. Also, rice viscosity characteristics change due to ages or it can be a result of protein composition change (Zhou et al. 2010). All the protein fractions show different quantities of aqueous solubility, which might affect grain quality by influencing the rates of starch hydration at some point all through cooking time.

Fatty acids composition in the current study showed palmitic, oleic and linoleic acid to be present in *bgl* mutants and WT. While the composition of Arachidic and docosanoic acid was very low in *bgl* mutant and WT (Fig. 7). The present results are similar to the earlier findings (Cho et al. 2006; Hwang et al. 2002); both revealed that palmitic, oleic, and linoleic acids are the major fatty acids present in rice. It can be concluded that map-based cloning is a potential approach for gene discovery, which provides information for dissecting the genetic and molecular basis of grain size and improving the grain size. The *bgl* is a promising gene with pleiotropic effect on grain length and appearance quality and can be utilized in genomic-assisted breeding for *japonica* rice (*Oryza sativa* subsp. *japonica*) cultivar improvement.

Authors' contribution

Conceptualization of research (JXL); Designing of the experiments (JXL) ; Contribution of experimental materials (CHS); Execution of field/lab experiments and data collection (MS) ; Analysis of data and interpretation (AK, YS, CZ); Preparation ofmanuscript (MS).

Supplementary materials

Supplementary Tables S1 to S3 are presented.

Acknowledgment

This work was supported by the National Key Research and Development of China (Grant No. 2017YFD0100300-5) and the Science and Technology Office of Zhejiang Province, China (Grant Nos. 2016C02050-6 and 2012C12901-2).

References

Adu-Kwarteng E., W. O., Ellis I., Oduro and J. T. Manful. 2003. Rice grain quality: a comparison of local varieties with new varieties under study in Ghana. Food Control, **14**: 507-514.

Bao J., S. Shen, M. Sun and H. Corke. 2006. Analysis of genotypic diversity in the starch physicochemical properties of nonwaxy rice: Apparent amylose content, pasting viscosity and gel texture. Starch-Starke, **58**: 259-267.

Baxter G., C. Blanchard and J. Zhao. 2004. Effects of prolamin on the textural and pasting properties of rice flour and starch. J. Cereal Sci., **40**: 205-211.

Baxter G., C. Blanchard and J. Zhao. 2014. Effects of glutelin and globulin on the physicochemical properties of rice starch and flour. J. Cereal Sci., **60**: 414-420.

Baxter G., J. Zhao and C. Blanchard. 2010. Albumin significantly affects pasting and textural characteristics of rice flour. Cereal Chem., **87**: 250-255.

BeMiller J. N. 2019. "Carbohydrate chemistry for food scientists. Third edition." eBook, London: WP, AACC, International Press.

Cai Y., C. Liu, W. Wang and K. Cai. 2011. Differences in physicochemical properties of kernels of two rice cultivars during grain formation. J. Sci. Food Agri., **91**: 1977-1983.

Chen M. H., C. J. Bergman, S. R. M. Pinson and R. G. Fjellstrom. 2008. Waxy gene haplotypes: Associations with pasting properties in an international rice germplasm collection. J. Cereal Sci., **48**: 781-788.

Cho K. S., H. J. Kim, J. H. Lee, J. H. Kang and Y. S. Lee. 2006. Determination of fatty acid composition in 120 Korean native rice cultivars. Hortscience, **41**: 1082-1082.

Dong M. H., D. Z. Sang, P. Wang, X. M. Wang and J. C. Yang. 2006. Changes in cooking and nutritional qualities of grains at different positions within a rice panicle under different nitrogen levels. Zhongguo Shuidao Kexue, **20**: 389-395.

Fan C. H., Y. Z. Xing, H. L. Mao, T. T. Lu, B. Han, C. G. Xu, X. H. Li and Q. F. Zhang. 2006. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor. Appl. Genet., **112**: 1164-1171.

Han X. Z. and B. R. Hamaker. 2001. Amylopectin fine structure and rice starch paste breakdown. J. Cereal Sci., **34**: 279-284.

Horiguchi G., Ferjani A., Fujikura U. and Tsukaya H. 2006. Coordination of cell proliferation and cell expansion in the control of leaf size in Arabidopsis thaliana. J. Plant Res., **119**: 37-42.

Horvath B. M., Magyar Z., Zhang Y., Hamburger A. W., Bako L., Visser R. G. F., Bachem C. W. B. and Bogre L. 2006. EBP1 regulates organ size through cell growth and proliferation in plants. Embo. J., **25**: 4909-4920.

Hu Z., He H., Zhang S., Sun F., Xin X., Wang W., Qian X., Yang J. and Luo X. 2012. A kelch motif-containing serine/Threonine protein phosphatase determines the large grain QTL trait in rice. J. Integr. Plant Biol., **54**: 979-990.

Hu Z., S. J. Lu, M. J. Wang, H. He, L. Sun, H. Wang, X. H., Liu, L. Jiang, J. L. Sun, X. Xin, W. Kong, C. Chu, H. W., Xue and J. X. Liu. 2018. A novel QTL qTGW3 encodes the GSK3/SHAGGY-like kinase OsGSK5/OsSK41 that interacts with *OsARF4* to negatively regulate grain size and weight in rice. Mol. Plant, **11**: 736-749.

Huang R., Jiang L., Zheng J., Wang T., Wang H., Huang Y. and Hong Z. 2013. Genetic bases of rice grain shape: so many genes, so little known. Trends Plant Sci., **18**: 218-226.

Hwang Y. H., Jang Y. S., Kim M. K. and Lee H. S. 2002. Fatty acid composition of rice bran oil and growth-promoting effect of rice bran extract and rice bran oil on Bifidobacterium and Lactobacillus. Agricul. Chemistry Biotechnol., **45**: 121-124.

Kim S. S., Lee S. F., Kim O. W. and Kim D. C. 2000. Physicochemical characteristics of chalky kernels and their effects on sensory quality of cooked rice. Cereal Chem., **77**: 376-379.

Kubota M., Saito Y., Masumura T., Kumagai T., Watanabe R., Fujimura S. and Kadowaki M. 2010. Improvement in the In vivo digestibility, of rice protein by alkali extraction is due to structural changes in prolamin/protein body-I particle. Biosci. Biotechnol. Biochemis., **74**: 614-619.

Kumamaru T., Satoh H., Iwata N., Omura T., Ogawa M. and Tanaka K. 1988. Mutants for rice storage proteins .1. screening of mutants for rice storage proteins of protein bodies in the starchy endosperm. Theor. Appl. Genet., **76**: 11-16.

Li N., Xu R. and Li Y. 2019. Molecular networks of seed size control in plants. *In* "Ann. Review Plant Biology, **70**: 435-463.

Lisle A. J., Martin M. and Fitzgerald M. A. 2000. Chalky and translucent rice grains differ in starch composition and structure and cooking properties. Cereal Chem., **77**: 627-632.

Mao H., Sun S., Yao J., Wang C., Yu S., Xu C., Li X. and Zhang Q. 2010. Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc. National Academy of Sciences of the United States of America, **107**: 19579-19584.

Martin M. and Fitzgerald M. A. 2002. Proteins in rice grains influence cooking properties. J. Cereal Sci., **36**: 285-294.

Paula A. M. and Conti-Silva A. C. 2014. Texture profile and correlation between sensory and instrumental analyses on extruded snacks. J. Food Eng., **121**: 9-14.

Pooni H. S., Kumar I. and Khush G. S. 1993. Genetic-control of amylose content in selected crosses of indica rice. Heredity, **70**: 269-280.

Qi P., Lin Y. S., Song X. J., Shen J. B., Huang W., Shan J. X., Zhu M. Z., Jiang L., Gao J. P. and Lin H. X. 2012. The novel quantitative trait locus GL3.1 controls rice grain size and yield by regulating Cyclin-T1;3. Cell Res., **22**: 1666-1680.

Qian Q., Guo L., Smith S. M. and Li J. 2016. Breeding high-yield superior quality hybrid super rice by rational design. National Sci. Review, **3**: 283-294.

Shu Q., Wu D., Xia Y., Gao M. and McClung A. 1998. Relationship between RVA profile character and eating quality in *Oryza sativa* L. Scientia Agricultura Sinica, **31**: 25-29.

Singh N., Sodhi N. S., Kaur M. and Saxena S. K. 2003. Physico-chemical, morphological, thermal, cooking and textural properties of chalky and translucent rice kernels. Food Chem., **82**: 433-439.

Verma D. K. and Srivastav P. P. 2017. Proximate composition, mineral content and fatty acids analyses of aromatic and non-aromatic indian rice. Rice Sci., **24**: 21-31.

Wang S., Li S., Liu Q., Wu K., Zhang J., Wang S., Wang Y., Chen X., Zhang Y., Gao C., Wang F., Huang H. and Fu X. 2015a. The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. Nature Genet., **47**: 949-954.

Wang Y., Xiong G., Hu J., Jiang L., Yu H., Xu J., Fang Y., Zeng L., Xu E., Xu J., Ye W., Meng X., Liu R., Chen H., Jing Y., Wang Y., Zhu X., Li J. and Qian Q. 2015b. Copy number variation at the GL7 locus contributes to grain size diversity in rice. Nature Genet., **47**: 944-957.

Xia D., Zhou H., Liu R., Dan W., Li P., Wu B., Chen J., Wang L., Gao G., Zhang Q. and He Y. 2018. GL3.3, a novel QTL encoding a GSK3/SHAGGY-like kinase, epistatically interacts with GS3 to produce extra-long grains in rice. Mol. Plant, **11**: 754-756.

Xing Y. and Zhang Q. 2010. Genetic and molecular bases of rice yield. *In* "Ann. Review Plant Biology, **61**: 421-442.

Xiong H., Yu J., Miao J., Li J., Zhang H., Wang X., Liu P., Zhao Y., Jiang C., Yin Z., Li Y., Guo Y., Fu B., Wang W., Li Z., Ali J. and Li Z. 2018. Natural variation in OsLG3 increases drought tolerance in rice by inducing ROS scavenging. Plant Physiol., **178**: 451-467.

Xu F., Sun C., Huang Y., Chen Y., Tong C. and Bao J. 2015. QTL mapping for rice grain quality: a strategy to detect more QTLs within sub-populations. Mol. Breed., 35.

Yamagata H., Sugimoto T., Tanaka K. and Kasai Z. 1982. Biosynthesis of storage proteins in developing rice seeds. Plant Physiol., **70**: 1094-1100.

Ying J. Z., Ma M., Bai C., Huang X. H., Liu J. L., Fan Y. Y. and Song X. J. 2018. TGW3, a major QTL that negatively modulates grain length and weight in rice. Mol. Plant, **11**: 750-753.

Yu J., Miao J., Zhang Z., Xiong H., Zhu X., Sun X., Pan Y., Liang Y., Zhang Q., Rehman R. M. A., Li J., Zhang H. and Li Z. 2018. Alternative splicing of OsLG3b controls grain length and yield in japonica rice. Plant Biotechnol. J., **16**: 1667-1678.

Zeng D., Tian Z., Rao Y., Dong G., Yang Y., Huang L., Leng Y., Xu J., Sun C., Zhang G., Hu J., Zhu L., Gao Z., Hu X., Guo L., Xiong G., Wang Y., Li J. and Qian Q. 2017. Rational design of high-yield and superior-quality rice. Nature Plants, 3.

Zhang X., Wang J., Huang J., Lan H., Wang C., Yin C., Wu Y., Tang H., Qian Q., Li J. and Zhang H. 2012. Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. Proc. National Academy of Sciences of the U. S. A., **109**: 21534-21539.

Zhou Z., K. Robards S., Helliwell and C. Blanchard. 2010. Effect of storage temperature on rice thermal properties. Food Res. Int., **43**: 709-715.

Zuo J. and J. Li. 2014. Molecular genetic dissection of quantitative trait loci regulating rice grain size. In "Ann. Review of Genet., **48**: 99-118.

Supplementary Table S1. A list of markers used for mapping BGL locus

Marker	Forward primer (5′→3′)	Reverse primer (5′→3′)
RM16	GATCGCCCAAATCTTCCT	TGAGCCACCTTCCAACAC
RM32	AAGCCCAGTTACAAACCCAA	CGAGTCAGTTCTACTCCCAAGA
M3B15	CCAAACCACCCGACGAAT	GGGTATGCTGCGACGAAAG
M3B6	GGTTACTACCGCCTCGGATCCC	CCTTGTGACTGGTAAAGAGATTTCCGT
M3B7	TCCAGGAGCGGAACTTGA	GAGGCAGAAACGGAAAGG
M3B8	CAACGAATGCTAGAAACTGA	TATTTAAAACAATGTCCCCC
M3B13	TCAAACTAACTCTAACCCTT	CAGAAACTAAATACACCCT
M3B12	GCAAGGGGCTAAGATGTAAA	GCAGATGAAAACTTGTGGAAT
M3B9	GCTCCTCGTCGTCAATGC	CAGGCTCCCTCTGGTCAC
M3B15	AGGAAGTGACCAGAGGGA	AAATTGGAAGACCGAAGC
M3B18	CGAGTCAGTTCTACTCCCAAGA	AAGCCCAGTTACAAACCCAA
M3B19	CGAGTCAGTTCTACTCCCAAGA	CGCTAGGGCAGCATCTAAA
M3B21	GATCGCCCAAATCTT	TCCACCTTCCAACAC
M3B22	AAGCCCAGTTACAAACC	GTCAGTTCTACTCCCAAGA
M3B1	TGCTAGTGGAAAAGGAAA	TATTATACCGTGGATGC
M3B2	ATTTGTACTAACAATCGTCTCAA	CATACTCTATAGTCTTTCCGTTA
M3B11	CTTTGTGCGTGATTTATC	GGACTGGACTCTGTTCT
M3B20	AAATCCAACGGTGAGAA	GGTCTACGAGGAGGTAA

Supplementary Table S2. A list of the primers used for qRT PCR BGL gene

Marker	Forward primer (5′→3′)	Reverse primer (5′→3′)
CYCA2.1	AGGTTGTCAAGATGGAGAGCGA	CGCTTTTTGTCTTCCTGGCA
CYCA2.2	AGGTTGTCAAGATGGAGAGCGA	CGCTTTTTGTCTTCCTGGCA
CYCB2.2	CTCAAGGCTGCACAATCTGACA	GCATTGACGGCTGGAATTTG
CYCD4	GCCATGGAGTTGATACATCCAA	CCAGTAGGGCTCCGTGGAAT
H1	GCAAGGCACCTGCAGCTT	AGGCAGCCTTTGTACAGATCCT
CYCD7.1	CCTTCCACACTGACGGTACAGTT	TGCCGCTGCCAAATAGACA
CAK1	GACGGTCAGATTAGACGCAAGA	TCCAAAGGATGTCCACA
EXPA5	AAGGCTGTGGCTTGATTGACA	TTAGGCCCAATTTTGCTATTTTG
EXPA10	TGACCAACTACAACGTGGTCCC	GCCAGTGTATGTTTTGCCGAAG
EXPB3	CTTTGAGTGGTTGGAGTGGTGG	GCAGCCTTCTTGGAGATGGAA
EXPB4	GTCGGTCTGTGTTGCGATTTG	CCTCCATTTCCCACACAGCTT
EXPB7	ACGGTGATCATCACGGACAT	TCGAAGTGGTACAGCGACACT
OsActin	ACCATTGGTGCTGAGCGTTT	CGCAGCTTCCATTCCTATGAA

Supplementary Table S3. Annotated genes in the mapping region

Gene locus	Putative function	Sequence difference
LOC_Os03g44460	Retrotransposon protein, putative, unclassified, expressed	No difference
LOC_Os03g44470	Retrotransposon protein, putative, unclassified, expressed	No difference
LOC_Os03g44484	DNA-directed RNA polymerase II subunit RPB2, putative, expressed	No difference
LOC_Os03g44500	Serine/threonine-protein phosphatase, putative, expressed	Single base change of G to A in the 17th exon
LOC_Os03g44510	Transposon protein, putative, CACTA, En/Spm sub-class, expressed	No difference
LOC_Os03g44530	GTPase of unknown function domain-containing protein, putative, expressed	No difference
LOC_Os03g44520	Expressed protein	No difference
LOC_Os03g44540	Nuclear transcription factor γ subunit, putative, expressed	No difference