GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Fine mapping of a grain weight quantitative trait locus, *qGW6*, using near isogenic lines derived from *Oryza rufipogon* IRGC105491 and *Oryza sativa* cultivar MR219

M.S. Ngu[1], M.J. Thomson[2], M.A.R. Bhuiyan[1], C. Ho[3] and R. Wickneswari[1]

[1]School of Environmental and Natural Resource Sciences, Faculty Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
[2]Plant Breeding, Genetics, and Biotechnology Division, International Rice Research Institute, Metro Manila, Philippines
[3]Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, Selangor, Malaysia

Corresponding author: M.S. Ngu
E-mail: meesiing@yahoo.com

Genet. Mol. Res. 13 (4): 9477-9488 (2014)
Received October 16, 2013
Accepted May 26, 2014
Published November 11, 2014
DOI http://dx.doi.org/10.4238/2014.November.11.13

ABSTRACT. Grain weight is a major component of rice grain yield and is controlled by quantitative trait loci. Previously, a rice grain weight quantitative trait locus (*qGW6*) was detected near marker RM587 on chromosome 6 in a backcross population (BC_2F_2) derived from a cross between *Oryza rufipogon* IRGC105491 and *O. sativa* cv. MR219. Using a BC_2F_5 population, *qGW6* was validated and mapped to a region of 4.8 cM (1.2 Mb) in the interval between RM508 and RM588. Fine mapping using a series of BC_4F_3 near isogenic lines further narrowed the interval containing *qGW6* to 88 kb between markers RM19268

and RM19271.1. According to the Duncan multiple range test, 8 BC_4F_4 near isogenic lines had significantly higher 100-grain weight (4.8 to 7.5% over MR219) than their recurrent parent, MR219 ($P < 0.05$). According to the rice genome automated annotation database, there are 20 predicted genes in the 88-kb target region, and 9 of them have known functions. Among the genes with known functions in the target region, *in silico* gene expression analysis showed that 9 were differentially expressed during the seed development stage(s) from gene expression series GSE6893; however, only 3 of them have known functions. These candidates provide targets for further characterization of *qGW6*, which will assist in understanding the genetic control of grain weight in rice.

Key words: Fine mapping; Grain weight; *Oryza rufipogon*; Quantitative trait locus; Near isogenic lines

INTRODUCTION

Grain weight (GW) is of great importance to rice yield because it influences grain yield directly. Together with grain number, GW is the main determinant of sink size for photosynthesis products in the rice plant. It is commonly indicated by thousand-grain weight, which represents a combined index of the components of grain length, width, and thickness. Rice grain length and shape are important to consumers because they determine the physical appearance, and they affect the cooking and eating quality of rice (Webb, 1991). GW is also important to farmers because it is one of the most stable components of yield in rice. Furthermore, larger grains have been constantly selected in most cultivated species during the domestication of crops (Doganlar et al., 2000). GW is also a highly heritable characteristic, making it useful for genetic analysis (Chauhan, 1998).

With the advances in molecular marker technology, nearly 300 quantitative trait loci (QTLs) for GW have been identified (www.gramene.org). Among these GW QTLs, many were detected in populations derived from crosses between the wild rice *Oryza rufipogon* (IRGC105491) and diverse cultivars (McCouch et al., 2007). It is interesting to note that most of the QTLs reported in these interspecific crosses identified *O. rufipogon* alleles as contributing to an increase in GW despite the fact that the *O. rufipogon* grain is small in size (Xiao et al., 1998; Moncada et al., 2001; Thomson et al., 2003; Septiningsih et al., 2003a,b; Wickneswari et al., 2012). Remarkable progress in fine mapping and cloning of QTLs associated with GW and shape has also been achieved (Huang et al., 2013). Xie et al. (2006, 2008) identified two QTLs for GW, *gw8.1* and *gw9.1*, in a 308- and 37.4-kb region on rice chromosomes 8 and 9, respectively, and the *O. rufipogon* alleles contributed to an increase in GW. Meanwhile, Li et al. (2004) fine mapped *gw3.1* to a pericentromeric region of 93.8 kb on chromosome 3, with the *O. rufipogon* allele being dominant for small grain size. Subsequently, a major gene for grain length and weight in this same region, *GS3*, was cloned (Fan et al., 2006). Moreover, the QTLs *GW2* on chromosome 2 for rice grain width and weight (Song et al., 2007) and *qSW5* for seed width (Shomura et al., 2008) were isolated via a map-based cloning strategy. However, only a few GW QTLs detected on chromosome 6 have been reported (Li et al., 1997; Xing et al., 2002; Ishimaru, 2003; Guo et al., 2009).

Previously a backcross population, BC_2F_2, between *O. rufipogon* (IRGC105491) and *O. sativa* cultivar MR219, a high-yielding variety in Malaysia, was used to identify 47 QTLs, including *qGW6*, which has an interval size of 25 cM (Sabu et al., 2006; Lim, 2007; Wickneswari et al., 2012). In this study, a BC_2F_5 family from the population was used to develop a BC_4F_3 fine mapping population to refine the position of *qGW6* to identify underlying candidate genes and to provide closely flanking markers to assist in the introgression of the QTL to increase GW for rice improvement.

MATERIAL AND METHODS

Plant materials and population development

A BC_2F_2 population was constructed for QTL mapping as described by Sabu et al. (2006) using *O. sativa* cv. MR219 as the recurrent parent and a wild accession of *O. rufipogon* (IRGC105491 from Malaysia) as the donor parent. From this population, one BC_2F_5 family (R2-10-18-2) was selected as the starting material for fine mapping and near isogenic line (NIL) development of *qGW6* on the basis that it contained an *O. rufipogon* IRGC105491 introgression in the target region, a high GW value, and relatively few nontarget background introgressions.

Backcrossing and field trials were carried out at the Malaysian Agricultural Research and Development Institute (MARDI) field experimental station in Seberang Perai, Penang, Malaysia (5°32'N, 100°27'E). The NILs and fine mapping population were developed by backcrossing twice the selected individuals from the BC_2F_5 family to the MR219 parent to eliminate nontarget genomic regions followed by selfing to fix the target introgression (Figure 1). Dormancy breaking was carried out at every generation by treating the seed in an oven at 55°C for 5 days. The seeds were then germinated in humidity-controlled germination chambers. Three-day-old seedlings were sown in plastic trays. At 21 days old, seedlings were transplanted to cement troughs in the greenhouse. Initial planting was performed in the greenhouse at MARDI research station, in Seberang Perai, Pulau Pinang. In the case of the BC_4F_3 and BC_4F_4 populations, the seedlings were transplanted to the farmer's field in Bumbung Lima, Seberang Perai, Pulau Pinang. The 37 BC_4F_4 families and MR219 were planted in a randomized complete block design with two replications. Thirty-six plants were planted in three rows (12 plants per row), and 10 plants in the center of the second row were used for phenotypic trait evaluation. Uniformed spacing (30 cm between rows and 25 cm between plants) was used in all plots. The space between two adjacent plots was 60 cm.

Phenotypic evaluation

Phenotypic trait evaluation was performed for the selected individuals of the BC_2F_5, BC_4F_2, BC_4F_3, and BC_4F_4 populations. In this study, we used hundred-grain weight (100GW) as a measure of GW to facilitate the phenotypic evaluation of large sample sizes. At the BC_2F_5, BC_4F_2, and BC_4F_3 generations, 100 mature and fully filled grains were selected randomly from each selected individual plant. At BC_4F_4, each sub-NIL was represented by 10 plants in the middle row of a randomized complete block design with three replications. The moisture content was determined by drying the seeds at 130°C for 2 h in an oven, followed by storage

in a desiccator for 20 min before recording the weight. The weight of 100 grains was adjusted to 14% moisture content by multiplying the dry GW with an adjustment coefficient. Data for GW assessment were arranged in an Excel sheet for statistical analysis using the Student *t*-test and the Duncan multiple comparison test by the SAS software v.9.1.3.

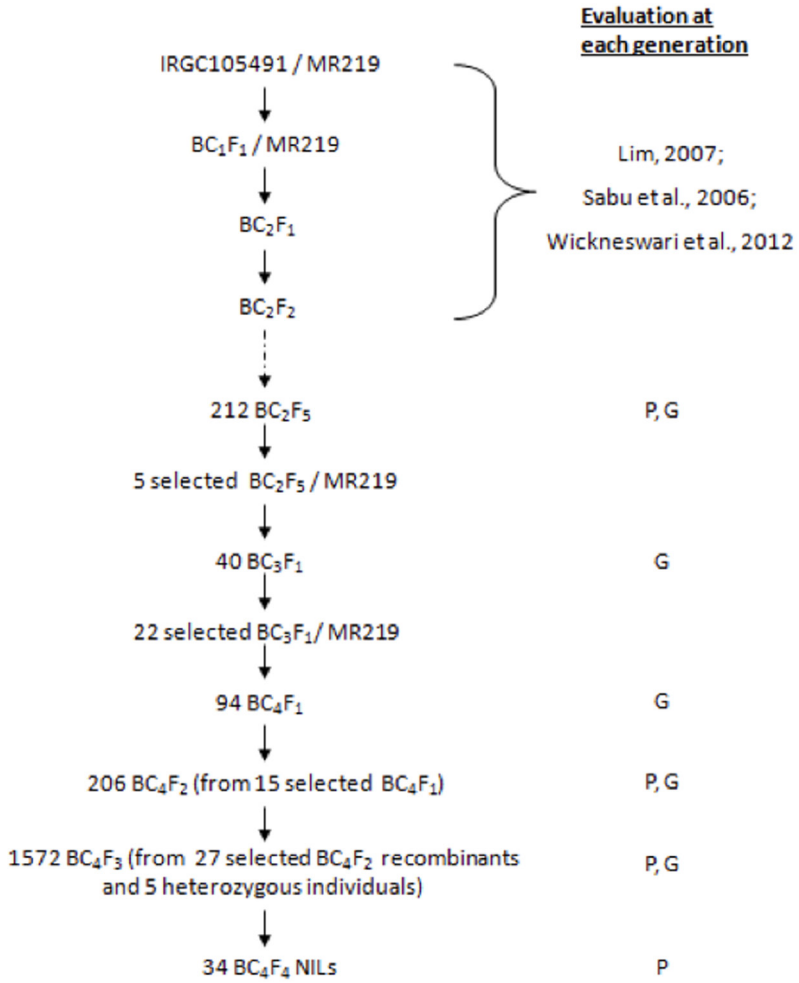

Figure 1. Procedures of near isogenic line (NIL) development and fine mapping. P = phenotype; G = genotype.

Molecular marker analysis

Genomic DNA of each plant was extracted and purified using the DNeasy Plant Mini Kit (QIAGEN, Germany). For high-throughput DNA extraction, the DNeasy 96 Plant Kit (QIAGEN) and TissueLyser (Retch, Germany) were used. The required density of molecular markers for the QTL region was obtained from microsatellite markers generated in previous studies (Lim, 2007; Wickneswari et al., 2012), the published rice simple sequence repeat

(SSR) framework map by Chen et al. (1997) and Temnykh et al. (2001), and the consolidated marker data by the International Rice Microsatellite Initiative (McCouch et al., 2002). Primers of some SSRs were redesigned using the Primer3 software (Rozen and Skaletsky, 2000) to get specific polymerase chain reaction (PCR) amplification. Redesigned primers were designated by the suffix ".1." Primer sequences and the map position of primers used in this study are listed in Table 1.

PCR conditions were as described in Panaud et al. (1996). The forward primers were fluorescent-labeled at the 5'-end with 6-FAM, VIC/HEX, NED, or PET dye and incorporated into PCR for fragment detection using capillary separation on the ABI3100 Genetic Analyzer (Applied Biosystems, USA). Fragment lengths were estimated and scored using the GeneMapper v.4.0 (Applied Biosystems).

Data analysis

QTL analysis was carried out by single-marker regression using QGene v4.3.2 (Nelson, 1997). A QTL was declared if the phenotype was associated with a marker locus at logarithm of odds (LOD) >3.0. The proportion of observed phenotypic variation attributable to a particular QTL was estimated by the coefficient of determination (R^2). The proportion of the genome corresponding to *O. rufipogon* introgressions for the final fine mapping populations was calculated using the GGT 2.0 software (van Berloo, 2008).

In silico gene expression analysis for candidate genes

The physical position of the refined *qGW6* QTL region was obtained based on the position of flanking markers from the rice genome automated annotation database (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/). The genomic region underlying *qGW6* was also searched for putative candidate genes by scanning all the annotated genes present in the target region from the rice genome automated annotation database (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/) and confirmed by protein Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Gene IDs of the obtained putative candidate genes were used to extract the cDNA microarray data from the rice gene expression database (http://www.ncbi.nlm.nih.gov/geo/). The results were in the form of $\log_2$ (fold-change of gene expression, M). The gene expression was considered significantly up-regulated when $\log_2(M) > 1$ or 2-fold induced, and it was significantly down regulated when $\log_2(M) < 1$ or 2-fold suppressed. In this study, only the gene expression series GSE6893 (Jain et al., 2007), which profiled the temporal gene expression of rice seed during various stages of reproductive development, was analyzed in detail.

RESULTS

Fine mapping of *qGW6*

In earlier studies (Lim, 2007; Wickneswari et al., 2012), *qGW6* was detected near marker RM587 and was flanked by RM585 and RM540 on chromosome 6. A total of 45 pairs of new SSR primers were selected and synthesized for the target region of *qGW6*. Among them, 20 (44.4%) showed polymorphism between the parents, *O. rufipogon* and

MR219. Subsequently, 15 markers that gave consistent and repeatable amplification were used to screen the progenies and map populations (Table 1). Among them, 4 primers (RM19266.1, RM19269.1, RM19271.1, and RM19272.1) were redesigned using Primer 3 software.

Table 1. Details of simple sequence repeat (SSR) primers used in the fine mapping of the *qGW6* quantitative trait locus region.

SSR loci	Forward primer (5'-3')	Reverse primer (5'-3')	Repeat motif	Ta (°C)	Dye
RM6775/RM19231	AATTGATGCAGGTTCAGCAAGC	GGAAATGTGGTTGAGAGTTGAGAGC	$(AGG)_8$	55.0	HEX
RM540*	GCCTTCTGGCTCATTTATGC	CTAGGCCTGCCAGATTGAAC	$(AG)_{16}$	52.7	6-FAM
RM508/RM19244	AGAAGCCGGTTCATAGTTCATGC	ACCCGTGAACCACAAAGAACG	$(AG)_{17}$	55.0	HEX
RM435/RM19256	CTGGTTAATTACGTGCATGTCTGG	GGCATGTCATGTCTTGGTCTCC	$(ATG)_7$	55.0	HEX
RM3132/RM19266.1	GAGGAGAGTTCTTGGGTTTGG	GCCCTCCACTTCTCTCTTCA	$(AC)_{14}$	55.0	6-FAM
RM19268	CTGATTTGCACGATGAGAACTACC	CTTCATCTCCTTTGTGCAGAAGC	$(AAG)_7$	55.0	6-FAM
RM19269.1	CGAGAGGAGACGAGACGAAG	ATCAGCCCAAAATTGAATCG	$(AGCG)_5$	61.6	HEX
RM19271.1	GGTGGGGTTCAGATTAGACG	GCCTCTGTACCGCAAGTACC	$(AGG)_7$	63.0	6-FAM
RM19272.1	GTCCCTTCTCCCTTCACTCC	CAATCATCTGCTGCTTGGTG	$(CCG)_7$	59.0	6-FAM
RM19285	TAATCAAGCAGCACAGCATCAGG	TCTGAACTTCCACTCGAGATTGC	$(AG)_{18}$	55.0	6-FAM
RM19296	CTAGCTTGACGCCAAGGACACC	GCACAGACGCACACTGATCTCC	$(AC)_{15}$	55.0	6-FAM
RM19304	TCCCTTTCAGTCAAGACAACTACTGG	TCTCCTCTTCTCTGCTTCTCAAGG	$(AG)_{11}$	55.0	HEX
RM8074/RM19306	GCTGCTCTAGCAATTTGGTTGG	TTCATCATCCACAGTCCAAGTCC	$(TTTC)_5(AG)_{13}$	55.0	6-FAM
RM588/RM19322	TCTTGCTGTGCTGTTAGTGTACG	GCAGGACATAAATACTAGGCATGG	$(TGC)_9$	55.0	6-FAM
RM190/RM19341	GCTACAAATAGCCACCCACACC	CAACACAAGCAGAGAAGTGAAGC	$(AG)_{18}$	58.0	HEX
RM587*/RM19364	TTCCCATCTGCACTACCATAATCC	GAGCAGAGATGTGCTTTGCTACC	$(AAG)_{18}$	55.0	NED
RM585*/RM19422	CTAGCTAGCCATGCTCTCGTACC	CTGTGACTGACTTGGTCATAGGG	$(AG)_{45}$	55.0	VIC

*Markers from Lim (2007).

To narrow the target region of *qGW6*, 212 individuals from the BC_2F_5 population of family R2-10-18-2, were evaluated both phenotypically and genotypically using 9 markers (RM6775, RM540, RM508, RM435, RM8074, RM588, RM190, RM587, and RM585). Using QTL analysis, the 25-cM target region of *qGW6* was narrowed to a region of 4.8 cM (1.2 Mb) in the interval between RM508 and RM588 (Figure 2). Two rounds of backcrossing were performed to remove background introgressions, and 94 BC_4F_1 lines were genotyped to select 15 recombinant individuals for the first round of fine mapping. A population of 206 BC_4F_2 recombinants generated from the selected BC_4F_1 lines was evaluated both phenotypically and genotypically using 4 markers in the target region (RM508, RM435, RM8074, and RM588). Twenty-seven BC_4F_2 recombinants, five BC_4F_2 heterozygous lines with high GW, and three controls (*O. rufipogon*-like, MR219-like with high GW, and MR219-like with low GW) were identified and genotyped using 96 background markers that were evenly distributed across all 12 chromosomes. All the recombinants and heterozygous lines were advanced to BC_4F_3 with the three controls and recurrent parent, MR219.

While developing the fine mapping population, 99 BC_4F_2 individuals were also used to confirm the allelic effect at *qGW6*. To do this, phenotypic means were compared among the three genotypic classes that were defined by the allele constitution of 4 markers (RM508, RM435, RM8074, and RM588) (Figure 3). Based on the Student *t*-test, the mean GW of homozygous *O. rufipogon* class (2.518 g) was significantly different than that of the homozygous MR219 class (2.470 g) ($P < 0.05$), but the heterozygous class (2.492 g) was not significantly different from the homozygous classes. This demonstrated that the *O. rufipogon* allele at *qGW6* contributed to the increase of GW in the MR219 background.

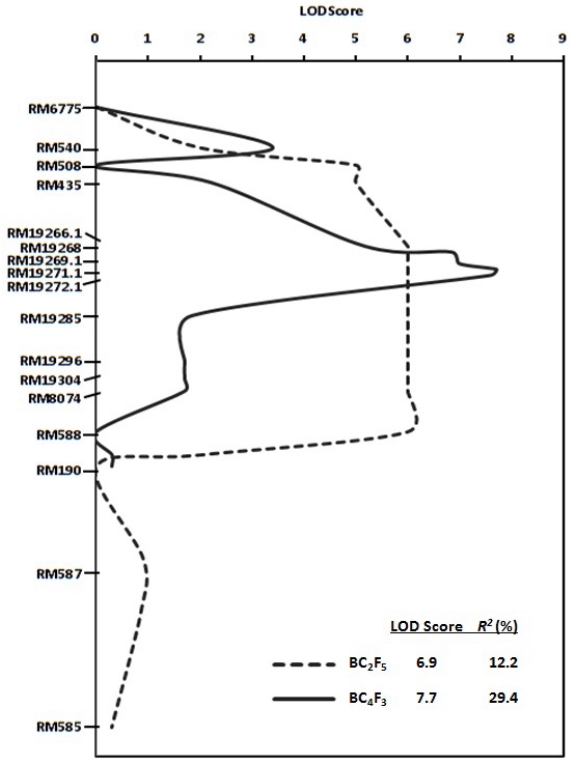

Figure 2. Quantitative trait locus analysis of *qGW6* in the BC_2F_5 and BC_4F_3 populations.

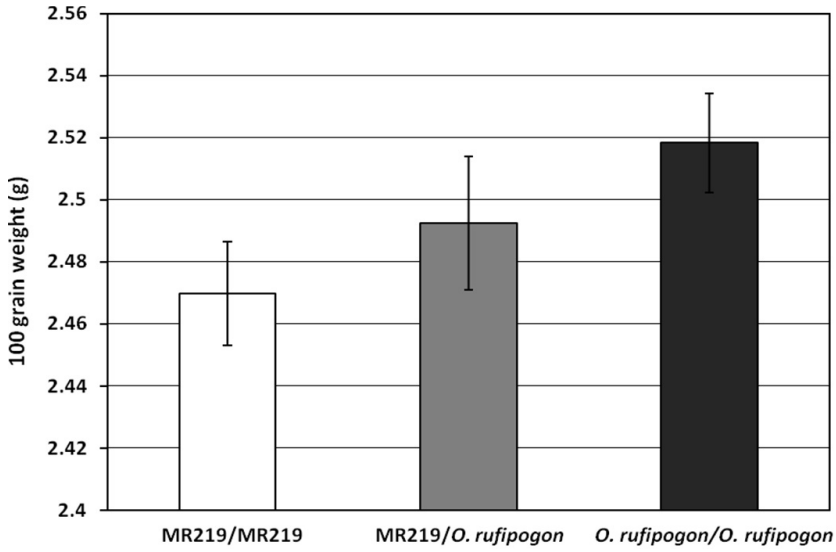

Figure 3. Allelic effect of *qGW6* based on the genotypic classes at the QTL (RM508-RM588) using 99 BC_4F_2 individuals.

At BC_4F_3, a total of 1572 individuals were first screened using the 4 markers flanking the 4.8-cM target region. From the screening, 108 individuals with fixed introgression and 72 recombinants were selected for phenotyping and genotyping with markers RM6775, RM540, and RM190, as well as 4 additional markers (RM19268, RM19285, RM19296, and RM19304) in the target region (Table 1). A new QTL peak was detected between markers RM435 and RM19285, with the LOD score of 7.7 and R^2 of 29.4% (Figure 2). This new QTL peak was confirmed when the phenotypes and genotypes of the informative lines were compared (Figure 4A). Comparing the overlap between the high GW groups 2/3 and 4/5 delimited the QTL between RM435 and RM19296, while excluding the regions of the low GW group 6 and group 8 narrowed *qGW6* between RM19268 and RM19285 (a 276-kb region). To more precisely determine the location of *qGW6*, 4 new markers (RM19266.1, RM19269.1, RM19271.1, and RM19272.1) were added to further subdivide the RM435-RM19285 interval (Figure 4B). All the BC_4F_3 lines with recombination breakpoints between RM435 and RM19285 were genotyped with the new markers. On the basis of these data, 41 informative lines were identified and grouped into five genotypic classes (Figure 4B). When comparing the phenotypic means for each class with the overlapping *O. rufipogon* introgressions, the high GW group 3 and the low GW group 4 indicate that the right boundary is limited to RM19271.1, which further delimits *qGW6* to the 88-kb region between RM19268 and RM19271.1.

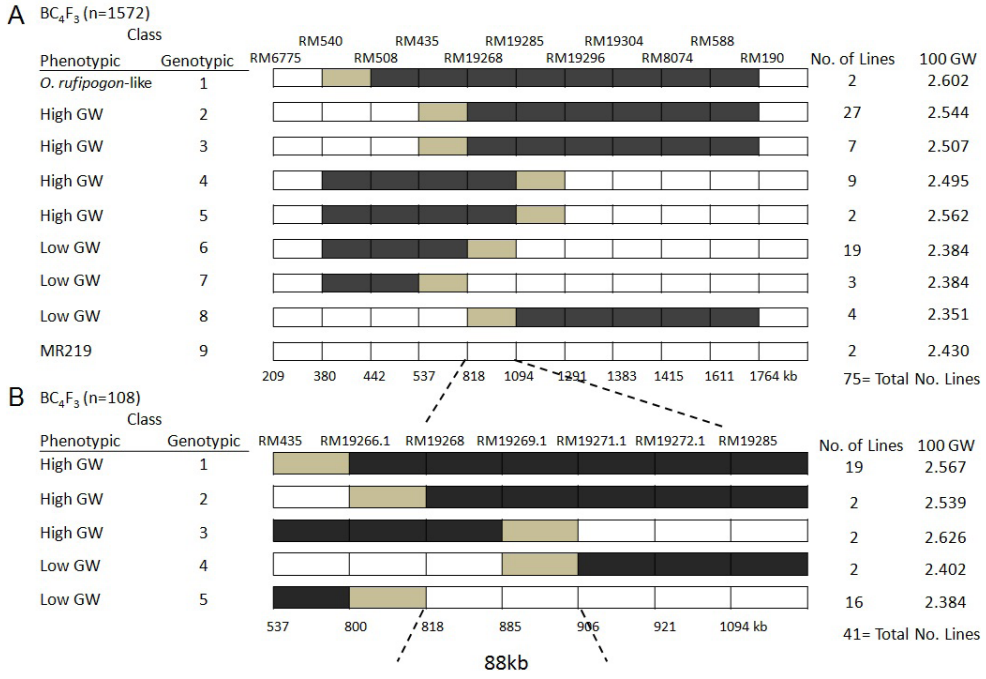

Figure 4. Graphical genotypes of lines used in the fine mapping of *qGW6*. Darkly shaded rectangles = *Oryza rufipogon* genotype; open rectangles = homozygous *O. sativa* cv. MR219 genotype; lightly shaded rectangles = marker intervals containing recombination breakpoints. **A.** BC_4F_3 population, population size, N = 1572; dotted lines = location of *qGW6* fine mapped to a 276-kb region between RM19268 and RM19285. **B.** One hundred and eight BC_4F_3 lines were genotyped using four new markers between RM435 and RM19285; dotted lines = location of *qGW6* fine mapped to an 88-kb region between RM19268 and RM19271.1.

To eliminate the remaining background introgression(s), the 108 fixed lines were screened again using 32 background markers. A total of 34 BC_4F_3 lines were selected and advanced to BC_4F_4 for phenotypic trials with three controls and MR219. According to the Duncan multiple range test, 8 BC_4F_4 NILs had a significantly higher 100GW than the recurrent parent, MR219 ($P < 0.05$). Phenotypic trait evaluation between the 8 NILs and MR219 showed that the GW was increased by 4.8 to 7.5% over MR219, with an average of 6.3% ($P < 0.05$). the GGT32 software calculated that the proportion of genome corresponding to *O. rufipogon* introgressions per individual ranged from 1.4 to 7%, with an average of 3.8%.

In silico gene expression analysis for candidate genes underlying *qGW6*

According to the rice genome automated annotation database (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/) and protein BLAST results (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), there were 20 predicted genes in the 88-kb target region (RM19268-RM19271.1) with 9 having known functional annotations, 9 being expressed proteins, and 2 being hypothetical proteins. *In silico* gene expression analysis revealed that 9 genes were differentially expressed during seed development stage(s) of series GSE6893, and only 3 of them had known functions (**Table S1**). The genes with known functions were 1) LOC_Os06g02390, a gene classified as a putative ADP-ribosylation factor 1 with 5 exons and a transcript length of 570 bp; 2) LOC_Os06g02470, a gene with a transcript length of 234 bp containing 2 exons, which was classified as ATOZI1; and 3) LOC_Os06g02560, a gene of 1071 bp consisting of 3 exons, which was classified as a putative growth-regulating factor.

DISCUSSION

Previous GW QTLs on chromosome 6 were identified in the region between 57.1 and 116.9 cM (Li et al., 1997; Xing et al., 2002; Ishimaru, 2003; Guo et al., 2009; Ishimaru et al., 2013). However, the position of *qGW6* that was detected in this study was not near any of these previously published QTLs. In this study, *qGW6*, which is located on the short arm of chromosome 6, was fine mapped to an 88-kb interval between RM19268 and RM19271.1, and phenotypic evaluation demonstrated that *qGW6* contributed to the increase of GW in the MR219 background. Shan et al. (2009) revealed that *spd6*, which is located about 1277 kb from *qGW6*, is responsible for small panicles, small grains, and dwarfness in the background of the *indica* cultivar Teqing. This implies the importance of fine mapping for the successful introgression of desired QTLs with a minimum of undesired effects or linkage drag. Even though *qGW6* is a minor QTL, the study of this QTL may contribute to a better understanding of the genetic regulation of rice GW. It is also important to note that *qGW6* is located about 800 kb from RM190. RM190 is the marker anchoring the waxy gene (*Wx*), which could cause changes in amylose content. Amylose content analysis using nine BC_4F_3 lines with different GWs compared to their parents, MR219 and *O. rufipogon*, showed no increase in amylose content, and there was no *Wx* gene polymorphism observed. Although carrying an introgression from *O. rufipogon* (which has high amylose content), these lines with a high GW have an intermediate level of amylose content, which is preferred by local rice consumers (Ong et al., 2012).

O. rufipogon matures earlier than MR219. According to Lim (2007) and Wickneswari et al. (2012), the QTL controlling date to maturity (*qDTM-6*) and *qGW6* were clustered on the same short arm of chromosome 6 (RM587-RM585). Even though *qDTM-6* was not

studied in this project, we observed the early maturity of the fine mapping population. These lines that combine earliness and increased GW can be deployed as pre-breeding materials to develop varieties that can be cultivated in climatic regions where long-growth duration varieties cannot be grown and offering flexibility in planting dates while allowing double or triple cropping. The increase of GW in the eight BC_4F_4 NILs may appear small (4.8 to 7.5% over the MR219 parent); however, this is still valuable for increasing the grain yield of the elite MR219 parent without changing the grain size so much as to affect consumer acceptance of the grain quality.

During the background screening of the BC_4F_2 and BC_4F_3 individuals, another GW QTL (*qGW1*), which was contributed by *O. rufipogon*, was identified and mapped to a 4.8-Mb region on chromosome 1. *qGW1* was not detected or mapped in the BC_2F_2 population that was developed by Sabu et al. (2006). This could have been contributed by the mapping population size, the low degree of saturation with molecular markers in the previous linkage map, or possible interactions between the QTLs (Lim, 2007; Wickneswari et al., 2012). *qGW1* may have a higher phenotypic effect than *qGW6*. NIL development and increasing the resolution for this QTL are needed to reduce the size of the *O. rufipogon* introgression at *qGW1*. By crossing two NILs harboring *gw8.1* and *gw9.1*, Jin et al. (2011) revealed an epistatic interaction between QTLs for GW using a two-way analysis of variance. Hence, further studies with larger mapping populations are recommended to understand the epistatic interaction between *qGW1* and *qGW6*, if any.

In silico analysis has been utilized in rice to identify candidate genes (Ma et al., 2006; Liu et al., 2007; Chandel et al., 2011). In this study, *in silico* gene expression analysis was conducted to identify putative candidate genes underlying *qGW6* using the microarray data in the rice gene expression database. A total of nine genes were found to be differentially expressed during seed development in rice (Jain et al., 2007), suggesting that they may have a role in grain formation, which contributes to GW. Jain et al. (2007) revealed specific and/or overlapping expression of F-box proteins during various seed development stages, suggesting the involvement of F-box protein-encoding genes in rice seed development. Because gene expression analyses were not conducted using the NILs in this study, the information obtained from the *in silico* analysis can serve as a reference for further studies. Transcript profiling to validate the *in silico* analysis will provide valuable data to further narrow the list of candidate genes underlying *qGW6*.

ACKNOWLEDGMENTS

Research supported by the Ministry of Science, Technology and Innovation, Malaysia (grant #UKM-ABI-NBD0001-2007). In addition, the Ministry of Higher Education provided some assistance to this research through the awarded grant #LRGS/TD/2011/UPM-UKM/KM/01. We would like to thank the Malaysian Agricultural Research and Development Institute staff for their assistance in the development of the fine mapping population. We would also like to express our gratitude to the Malaysian Nuclear Agency for allowing us to use the ABI3100 Genetic Analyzer. We also want to acknowledge Ankitta Mishra for her help in grain weight evaluation.

Supplementary material

REFERENCES

Chandel G, Dubey M, Samual P and Meena R (2011). *In silico* mining and characterization of novel SSRs and candidate genes within QTLs controlling grain protein contents using MPSS signatures and micro array analysis in rice (*Oryza sativa* L.). *Plant Omics J.* 4: 126-135.

Chauhan JS (1998). Inheritance of grain weight, size and shape in rainfed rice (*Oryza sativa*). *Indian J. Agric. Sci.* 68: 9-12.

Chen X, Temnykh S, Xu Y, Cho YG, et al. (1997). Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 95: 553-567.

Doganlar S, Frary A and Tanksley SD (2000). The genetic basis of seed-weight variation: tomato as a model system. *Theor. Appl. Genet.* 100: 1267-1273.

Fan C, Xing Y, Mao H, Lu T, et al. (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.

Guo L, Ma L, Jiang H, Zeng D, et al. (2009). Genetic analysis and fine mapping of two genes for grain shape and weight in rice. *J. Integr. Plant Biol.* 51: 45-51.

Huang R, Jiang L, Zheng J, Wang T, et al. (2013). Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci.* 18: 218-226.

Ishimaru K (2003). Identification of a locus increasing rice yield and physiological analysis of its function. *Plant Physiol.* 133: 1083-1090.

Ishimaru K, Hirotsu N, Madoka Y, Murakami N, et al. (2013). Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat. Genet.* 45: 707-711.

Jain M, Nijhawan A, Arora R, Agarwal P, et al. (2007). F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 143: 1467-1483.

Jin FX, Ji SD, Xie XB and Kang JW (2011). Detection of epistatic interaction of two QTLs, *gw8.1* and *gw9.1*, underlying grain weight using near isogenic lines in rice. *Breed. Sci.* 61: 69-75.

Li J, Thomson MJ and McCouch SR (2004). Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. *Genetics* 168: 2187-2195.

Li ZK, Pinson SR, Park WD and Paterson AH (1997). Epistasis for three grain yield components in rice. *Genetics* 145: 453-465.

Lim LS (2007). Identification of Quantitative Trait Loci for Agronomic Traits in an Advanced Backcross Population Between *Oryza rufipogon* Griff. and the *O. sativa* L. cultivar MR219. Doctoral thesis, Universiti Kebangsaan Malaysia, Selangor.

Liu XQ, Lin F, Wang L and Pan QH (2007). The *in silico* map-based cloning of Pi36, a rice coiled-coil-nucleotide-binding site-leucine-rich repeat gene that confers race-specific resistance to the blast fungus. *Genetics* 176: 2541-2549.

Ma H, Zhang S, Ji L, Zhu H, et al. (2006). Fine mapping and *in silico* isolation of the EUI1 gene controlling upper internode elongation in rice. *Plant Mol. Biol.* 60: 87-94.

McCouch SR, Teytelman L, Xu Y, Lobos KB, et al. (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* 9: 199-207.

McCouch SR, Sweeney M, Li J and Jiang H (2007). Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. *Euphytica* 154: 317-339.

Moncada P, Martínez CP, Borrero J and Châtel M (2001). Quantitative trait loci for yield and yield components in an *Oryza sativa* x *Oryza rufipogon* BC2F2 population evaluated in an upland environment. *Theor. Appl. Genet.* 102: 41-52.

Nelson JC (1997). QGENE: software for marker-based genomic analysis and breeding. *Mol. Breed.* 3: 239-245.

Ong Y, Ngu MS, Parviz F and Tilakavati K (2012). Study on waxy gene polymorphism and amylose content of breeding lines derived from *Oryza rufipogon* x *Oryza sativa* CV. MR219. *Pertanika J. Trop. Agric. Sci.* 35: 833-842.

Panaud O, Chen X and McCouch SR (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 252: 597-607.

Rozen S and Skaletsky HJ (2000). Primer3 on the WWW for General Users and for Biologist Programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (Krawetz S and Misener S, eds.). Humana Press, Totowa, 365-386.

Sabu KK, Abdullah MZ, Lim LS and Wickneswari R (2006). Development and evaluation of advanced backcross families of rice for agronomically important traits. *Commun. Biometry Crop Sci.* 1: 111-123.

Septiningsih EM, Prasetiyono J, Lubis E, Tai TH, et al. (2003a). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theor. Appl. Genet.* 107: 1419-1432.

Septiningsih EM, Trijatmiko KR, Moeljopawiro S and McCouch SR (2003b). Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theor. Appl. Genet.* 107: 1433-1441.

Shan JX, Zhu MZ, Shi M, Gao JP, et al. (2009). Fine mapping and candidate gene analysis of spd6, responsible for small panicle and dwarfness in wild rice (*Oryza rufipogon* Griff.). *Theor. Appl. Genet.* 119: 827-836.

Shomura A, Izawa T, Ebana K, Ebitani T, et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40: 1023-1028.

Song XJ, Huang W, Shi M, Zhu MZ, et al. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39: 623-630.

Temnykh S, DeClerck G, Lukashova A, Lipovich L, et al. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Res.* 11: 1441-1452.

Thomson MJ, Tai TH, McClung AM, Lai XH, et al. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theor. Appl. Genet.* 107: 479-493.

van Berloo R (2008). GGT 2.0: versatile software for visualization and analysis of genetic data. *J. Hered.* 99: 232-236.

Webb BB (1991). Rice Quality and Grades. In: Rice. Vol. 2. Utilization. 2nd edn. (Lu BS, ed.). An AVI Book, New York, 89-119.

Wickneswari R, Bhuiyan MAR and Sabu KK (2012). Identification and validation of quantitative trait loci for agronomic traits in advanced backcross breeding lines derived from *Oryza rufipogon* x *Oryza sativa* cultivar MR219. *Plant Mol. Biol. Rep.* 30: 929-939.

Xiao J, Li J, Grandillo S, Ahn SN, et al. (1998). Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150: 899-909.

Xie X, Song MH, Jin F, Ahn SN, et al. (2006). Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived from a cross between *Oryza sativa* and *Oryza rufipogon*. *Theor. Appl. Genet.* 113: 885-894.

Xie X, Jin F, Song MH, Suh JP, et al. (2008). Fine mapping of a yield-enhancing QTL cluster associated with transgressive variation in an *Oryza sativa* x *O. rufipogon* cross. *Theor. Appl. Genet.* 116: 613-622.

Xing Z, Tan F, Hua P, Sun L, et al. (2002). Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. *Theor. Appl. Genet.* 105: 248-257.