Fine mapping of GS2, a dominant gene for big grain rice

Wuhan Zhang, Pingyong Sun, Qiang He, Fu Shu, Jie Wang, Huafeng Deng *

State Key Laboratory of Hybrid Rice, Hunan Hybrid Rice Research Center, Changsha, Hunan 410125, China

ARTICLE INFO

Article history:
Received 26 July 2013
Received in revised form
13 August 2013
Accepted 12 October 2013
Available online 24 October 2013

Keywords:
Rice
Big-grain
Fine mapping
GS2
Grain shape

ABSTRACT

Grain shape as a major determinant of rice yield and quality is widely believed to be controlled by quantitative trait loci (QTL). We have identified a novel gene “GS2” to largely regulate grain length and width in rice. The GS2 allele in the big-grain rice line ‘CDL’ functioned in a dominant manner. In the present study, we employed a chromosome walking strategy in the residual heterozygous lines from recombinant inbred population between cultivar “R1126” and CDL, and located the GS2 gene in an interval of ~33.2 kb flanked by marker GL2-35-1 and GL2-12 in the long arm of rice chromosome 2. According to genome annotations, three putative gene loci, LOC_Os02g47280, LOC_Os02g47290 and LOC_Os02g47300, exist in this candidate region. In addition, allelic analysis with previously reported genes demonstrated that GS2 was novel for regulating rice grain shape. These results will help promote the cloning and functional characterization of the GS2 gene and further develop linked markers to be used in marker-assisted breeding.

© 2013 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

Rice is one of the most important grain crops and staple foods for more than half of the global population [1]. Improving rice yield is an important means to fight hunger caused in part by a rapidly growing population along with reduced arable land area and occurring climate change and disease. Grain weight is a key component of rice grain yield, which is primarily defined by grain size that is determined by length, width and thickness. It is well known that all these three gain traits are controlled by quantitative trait loci (QTLs), affected by the environment [2,3]. High-precision genome sequences and various molecular markers have allowed mapping and identification of hundreds of QTLs associated with grain traits in rice (http://www.gramene.org). Many QTLs have been identified from different rice germplasms by a map-based cloning approach and these accomplishments imply the promise to help understand the molecular mechanisms underlying seed development and find ways to improve rice yield.

GS3, a major QTL for grain weight and length with a minor role in grain width and thickness, was recently fine mapped to a genomic region of 7.9 kb on chromosome 3 using 5,740 BC3F2 plants [4]. GS3 encodes a putative transmembrane protein composed of four domains, and each functions differently in regulating grain size [5]. Sequence analyses showed that large grains are due to an early stop codon from a substitution in the second exon and, suggesting that GS3 functions as a negative regulator of grain size. Similarly, loss of GS3 leads to

* Corresponding author. Tel.: +86 731 82872959; fax: +86 731 82873050.
E-mail address: dhf@hhrrc.ac.cn (H. Deng).
Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

2214-5141/$ – see front matter © 2013 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cj.2013.10.003
grain enlargement, which is true for GW2 [6], qSW5 [7] and TGW6 [8]. The major QTL for thousand-grain weight, TGW6, is mapped on chromosome 6 and encodes a novel protein with indole-3-acetic acid (IAA)-glucose hydrolase activity. Deletion of 1-bp in TGW6 exon results in a premature stop codon to prevent the production of the mature protein. It has been shown that function loss of the TGW6 allele results in simultaneous increase of grain weight and yield [8].

Furthermore, GS5 is a recently cloned QTL, which variation is associated with grain size diversity in rice, thus may be useful in improving yield in rice and, potentially, other crops [9]. Its spatial expression patterns demonstrate that higher expression of GS5 results in larger grains, suggesting that GS5 is a positive regulator of grain size [9]. Another QTL affecting grain width and yield, GW8, encodes a protein to positively regulate grain size. The GW8 function on grain is attributable to a critical deletion polymorphism in the promoter region. In contrast, a loss-of-function mutation brings about a better quality of appearance [10].

In this study, we report the identification and fine mapping of GS2 candidate gene. Our results demonstrated that GS2 was a novel gene involved in the regulation of grain length and width in rice. The identification and functional characterization of GS2 will help breed high-yield rice varieties and understand the underlying molecular mechanisms to control grain shape in rice and other crops.

2. Materials and methods

2.1. Construction of mapping population and phenotypic characterization of grain traits

Big-grain rice line CDL was crossed with a medium-grain line R1126. The resultant F1 plants were selfed to yield a F2 population of 1000 individuals, and the following recombined inbred lines (RIL). A differentiation of grain shape was observed in RIL28 line of F0, indicating heterozygous. The individual plants of RIL28 were selfed to generate a F2 population. 800 F2 RIL28 population segregated medium-grain (length ≤11 mm, width ≤3 mm) and big-grain (length >11 mm, width >3 mm) progenies in a ratio of 3:1, and these progenies were used to preliminarily map the grain-shape QTL using bulked segregant analysis (BSA) and recessive-class analysis (RCA) method. A preliminary map the grain-shape QTL using bulked segregant progenies in a ratio of 3:1, and these progenies were used to preliminarily map the grain-shape QTL using bulked segregant analysis (BSA) and recessive-class analysis (RCA) method. A locus between markers RM13819 and RM13863 on chromosome 2, designated as GS2, was clearly associated with the variation of grain phenotypes. The segregating populations were developed for fine-mapping of GS2 from the 10th plant in F2 of RIL28 line, named RIL28-10 which was heterozygous in the GS2 region flanked by RM13819 and RM13863. The selfing progenies of the selected residual heterozygous line (RHL) RIL28-10 produced RHL-F2 (3000 individuals) and RHL-F3 (30,000 individuals) population. Grain length and width were averaged from randomly chosen ten mature, filled and grains of 100 RHL-F2 individuals. The ten grains were lined up end to end along Vernier calipers to measure the length, and then arranged by breadth to measure grain width. Genetic analyses were conducted according to the frequency distribution maps of grain length and the \( \chi^2 \)-test. All rice materials were provided by the Hunan Hybrid Rice Research Center and planted in a field in Chunhua, Changsha City (summer) or Sanya, Hainan (winter).

2.2. Marker development and PCR analysis

Simple sequence repeat (SSR) markers distributed in whole genome were identified using publicly available rice genomic sequences (http://www.gramene.org). Single feature polymorphism (SFP) and intron length polymorphism (ILP) markers were obtained from Edwards et al., [11] and Wang et al., [12]. Other additional primers were designed and evaluated using Primer Premier 5. Sequence comparisons of the japonica cultivar Nipponbare (http://www.ncbi.nlm.nih.gov/guide) and the indica cultivar R1126 (the whole genome of R1126 was resequenced by BGI) within the target region of the chromosome were first analyzed online to obtain information on potential insertions/deletions (InDel). Primers were designed and evaluated for potential InDels containing the R1126 sequence using Primer Premier 5. Eight newly developed InDel markers are listed in Table 1. Total genomic DNA was extracted from fresh leaves using the CTAB method [13], and PCR analysis was performed according to Sun et al. [14].

2.3. Molecular mapping

Molecular marker analysis was carried out according to the method described by Zuo et al., [15] with minor modifications. Briefly, polymorphic markers between the two parents were first analyzed in a small population including the two parents, ten F1, medium-grain plants, and ten F1, big-grain plants. Then, the markers selected from this small population were further utilized to screen a part of big-grain individuals and all medium-grain individuals in the same segregating population for linkage analysis. Finally, data were collected and transformed according to the requirement of MAPMAKER 3.0 [16] to construct the linkage map.

3. Results

3.1. GS2 functions as a dominant gene in control of grain shape

Random evaluation of grain length and width of 100 RHL-F2 individuals revealed a continuous bimodal distribution (Fig. 1). Individuals of RHL-F2 population were classified into medium-grain (length ≤11 mm, width ≤3 mm, 24 individuals) and big-grain (length >11 mm, width >3 mm, 76 individuals) groups with a segregation ratio for big versus medium grain fit to a ratio of 3:1 \( (\chi^2 = 0.05 < \chi^2_{0.05,1} = 3.84) \). Large populations were investigated in F2, RHL-F2 and RHL-F1, with 179, 720 and 7400 medium grain individuals found in the total populations of 800, 3000, 30,000 individuals, respectively. Likewise, the segregation ratios of big versus medium grain fit to a ratio of 3:1 \( (\chi^2 = 2.80, 1.55, 1.76 < \chi^2_{0.05,1} = 3.84) \).

3.2. Fine mapping of GS2 gene

A total of 129 polymorphic markers were detected between R1126 and CDL from 400 SSR, SFP and ILP markers, and 113 well-distributed polymorphic markers were used to survey the ten medium-grain plants, ten big-grain plants of F2 population and parents. The GS2 gene was roughly mapped
to the interval between RM13819 and RM13863 on the long arm of chromosome 2. We found that six SSR markers, namely RM3289, RM1342, RM5305, RM13819, RM3212 and RM13863, located on chromosome 2 were clearly associated with the medium-grain phenotype. After further studying 179 F7 medium-grain plants using these six markers, the GS2 gene was located between RM13819 and RM13863 with genetic distances of 0.84 cM and 0.28 cM, respectively. Furthermore, 0 recombinant was detected by marker RM3212. These data were derived according to the recombinants revealed by each marker, covering a ~553-kb physical segment on the region of rice chromosome 2 (Fig. 2-A).

To fine-map the GS2 locus, 29 polymorphic InDels were selected from 142 InDels developed according to the information on the sequence (R1126 and Nipponbare) between RM3212 and RM13863. Further genotyping 2576 medium grain plants of the RHL-F3 revealed one recombinant in the proximity of GL2-35-1 and GL2-12. In addition, RM3212 and GL2-11 were verified to be linked to the GS2 gene. The GS2 locus was therefore finally narrowed down to the genomic region flanked by GL2-35-1 and GL2-12, a fragment of approximately 33.2 kb in length (Fig. 2-B).

Table 1 – InDel markers developed for fine mapping of the GS2 locus.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Size (bp)</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>GL2-26</td>
<td>186</td>
<td>AATAGTCGTAAGGCTCTTAG</td>
<td>CGGCCTATTATCCCATAT</td>
<td>29688978</td>
</tr>
<tr>
<td>GL2-31</td>
<td>119</td>
<td>TTCAGTCCCTTCTGTCGTC</td>
<td>CCAATCCCAATCTGACCC</td>
<td>29705167</td>
</tr>
<tr>
<td>GL2-35</td>
<td>144</td>
<td>GTGGCAACCATATTTCCTG</td>
<td>CTCTGCCCTGTCATTAA</td>
<td>29737769</td>
</tr>
<tr>
<td>GL2-35-1</td>
<td>84</td>
<td>GAGTATAGTATGTGGCAGGTCG</td>
<td>GATGGGAGGGTGGAGAC</td>
<td>29741153</td>
</tr>
<tr>
<td>GL2-11</td>
<td>657</td>
<td>GAGAAGCCATATGGCA</td>
<td>TCTACAAACTACAAAACAA</td>
<td>29767941</td>
</tr>
<tr>
<td>GL2-12</td>
<td>180</td>
<td>GGTTGCGCCCTGTAGTGAAT</td>
<td>ACCGCTGCCTGAGAGGA</td>
<td>29774335</td>
</tr>
<tr>
<td>GL2-17</td>
<td>112</td>
<td>TTGATTACTAACCGAGGAAAG</td>
<td>CATCTGCAATAGTGAAT</td>
<td>29800233</td>
</tr>
<tr>
<td>GL2-20</td>
<td>199</td>
<td>AACCTTTCCCGTAATTTGTGC</td>
<td>CGCTGAGTGATACATGTC</td>
<td>29822237</td>
</tr>
</tbody>
</table>

* The size of amplicon in Nipponbare; ** The marker position in NCBI database.

3.3. Predicted genes at the GS2 locus

In the 33.2-kb genomic interval of the Nipponbare genome, a total of three putative genes including LOC_Os02g47280, LOC_Os02g47290 and LOC_Os02g47300 were predicted by TIGR rice annotation (http://rice.plantbiology.msu.edu/cgi-
bin/gbrowse/rice/) (Fig. 2-B). LOC_Os02g47280 encoded a putative growth-regulating factor; LOC_Os02g47290 and LOC_Os02g47300 encoded hypothetical proteins with no further evidence such as expressed sequence tag (EST) or RNA.

4. Discussion

Because of the recent developments in bioinformatics and genome sequencing to yield an impressive number of molecular markers, many major QTLs responsible for grain shape and yield have been fine mapped and cloned in the past 20 years. In this paper, we fine mapped GS2 using RHL population developed from a big-grain rice line CDL and a medium-grain line R1126. GS2, which controls grain length and width, was narrowed down to a candidate genomic region of 33.2 kb defined by the InDel markers GL2-35-1 and GL2-12. Three annotated genes (LOC_Os02g47280, LOC_Os02g47290 and LOC_Os02g47300) were identified within the critical 33.2-kb genomic region of Nipponbare (japonica) genome (http://rice.plantbiology.msu.edu). LOC_Os02g47290 and LOC_Os02g47300 encoded hypothetical proteins with no gene ontology annotation; thus those two genes might have no or marginal direct relevance to the grain shape development according to their putative functions.

The LOC_Os02g47280 encodes a growth-regulating factor protein, which belongs to the GRF family of proteins consisting of twelve members. The protein of LOC_Os02g47280 has two putative alternative splice forms, and both contain a WRC domain and a QLQ domain. Interestingly, the protein of LOC_Os02g47280 shares homology with a protein in Brachypodium, Zea mays L., Populus L. and Sorghum vulgare Pers. (http://rice.plantbiology.msu.edu/). The WRC domain contains a putative nuclear localization signal and a zinc-finger motif (C3H). The WRC
domain was suggested to be involved in DNA binding while QLQ domain was shown to affect protein–protein interactions [17]. Recently it was demonstrated that LOC_Os02g47280 is down-regulated by mir396 during grain development in rice [18]. Therefore, LOC_Os02g47280 should be considered the most likely candidate for GS2. We are currently investigating a genetic complementation of the candidate gene by transformation and other functional analyses.

To date, more than 40 QTLs related to grain shape and yield have been primarily mapped on chromosome 2 of rice (http://www.gramene.org/). Some of these are located in the proximity of GS2. For example, the QTL qGL-2a, which affects grain length, was mapped in an interval between the restriction fragment length polymorphism (RFLP) marker C560 and C1408, accounting for 11.7% of total phenotypic variations [19]. Another QTL C560 and C1408, accounting for 11.7% of total phenotypic variations [19]. Another QTL

GL2-35-1 and GL2-12 on chromosome 2. Three annotated genes were identified within the GS2 locus from Nipponbare genome, and the LOC_Os02g47280 was considered the most likely candidate for GS2. No QTL responsible for grain shape and yield has been fine mapped and cloned at GS2 locus.

5. Conclusions

The research presented in this study identified a novel gene GS2 responsible for grain length and width in rice. GS2 was localized to an interval of −33.2 kb between the markers.

Acknowledgments

This work is financially supported by the National High Technology Research and Development Program of China (2011AA10A101) and the Hunan Provincial Natural Science Foundation of China (10J2025).

References


