A large-effect QTL for grain weight in rice on chromosome 10

Nivedita Singh1,2, Shahana Majumder1, O. N. Singh3, Prashant Vikram4, A. K. Singh4 and Sanjay Singh5

1School of Engineering and technology, Sharda University, New Delhi, India
2National Bureau of Plant Genetic Resources, Pusa, New Delhi
3Central Rice research Institute, Cuttack, India
4College of Agriculture, N. D. University of Agriculture &Technology, Faizabad, UP, India
5National Research on Plant Biotechnology, Pusa, New Delhi, India

*Corresponding author: sanjay_singh777@yahoo.com

Abstract

Grain yield of the rice can be enhanced through genetic manipulation of the yield components. Thousand grain weight is an important component which can be improved through conventional as well as molecular breeding approach. A major effect novel genomic region, $TGW_{10}$, for high grain weight was identified through a bulked segregant analysis approach in Azucena × IR64 population. Markers associated with grain weight in this region were RM25719 and RM5352. These markers were located on chromosome 10 at 19.8 Mb and 21.1 Mb, respectively. This QTL explained phenotypic variance of more than 20% indicating this region to be a major effect locus. Markers linked with the QTL identified in present study should be used for enhancement in the grain yield potential of rice.

Keywords: Bulked segregant analysis; Grain weight; Grain yield; Quantitative trait loci; Rice.  
Abbreviations: BSA_Bulked segregant analysis; QTL_Quantitative trait loci; RIL_Recombinant inbred lines; TGW_Thousand grain weight.

Introduction

Rice accounts for around 20% of the global calorie intake and is a staple food in South and Southeast Asia. The region shares more than 60% of the world’s caloric intake (Fitzgerald et al., 2009; Miura et al., 2011). Rice production and productivity require a continuous increase for ensuring food security. The development and application of superior rice varieties is one of the most effective, sustainable, and economical approaches for improving rice yield (Rosegrant et al., 2003). An ideal superior rice cultivar should have high grain-yield potential with improved grain shape, nutritional value, disease resistance, and stress tolerance. Grain yield in rice is determined by three major components: number of panicles per plant, number of grains per panicle, and grain weight. Among these, the most important trait is grain weight, which is measured as the 1000-grain weight ($TGW$). Grain shape is characterized by the combination of grain length, grain width, grain length-to-width ratio, and grain thickness. These four parameters are positively correlated with grain weight (Tan et al., 2000). Most of these agronomically important traits are known to be genetically controlled by multiple genes (Tan et al., 2000) which are referred as quantitative trait loci (QTLs). Rice genetic improvement is currently being done by conventional and the molecular breeding approaches. Identification and deployment of the QTLs for $TGW$ will improve the trait with minimum linkage drag in popular rice cultivar. A number of QTLs have been mapped for grain weight and grain shape on 12 chromosomes of rice but only $GS_{3}$, $GW_{2}$, and $qSW_{1}$ have been cloned till date. The $GS_{3}$ gene, which is located near the centromere of chromosome 3, was cloned by Fan et al. (2006). This gene is considered as a key gene for grain length because it existed in all long-grain varieties used by Fan et al. (2006). The $GW_{2}$ gene, controlling grain width, is located on chromosome 2 (Song et al., 2007). This gene encodes a protein with the function of ring-type E3 ubiquitin ligase. $qSW_{1}$ gene, responsible for seed width, was mapped on chromosome 5 and cloned by Shomura et al. (2008). A 1.2 kb deletion was found in the coding sequence of $qSW_{1}$ (Shomura et al., 2008). Besides, QTLs $GW_{6,1}$ and $GW_{6,2}$ for grain width were fine-mapped in a 306.4 kb region on chromosome 8 and a 37.4 kb on chromosome 9 respectively (Xie et al., 2006; Xie et al., 2008). Two QTLs for grain weight, $GW_{1,1}$ and $GW_{1,2}$, have been mapped on chromosome 1 (Yu et al., 2008). Total number of fine-mapped QTL for grain weights are less than 10 in number (Huang et al., 2013). Majorly characterized QTLs are $GW_{3}$ and $GW_{9}$ on chromosomes 3 and 6 respectively (Guo et al. 2008). Other major QTLs characterized are on chromosome 8 ($GW_{8,3}$), chromosome 9 ($GW_{9,3}$), chromosome 11 ($GW_{11,3}$) and chromosome 1 ($GW_{1,1}$, $GW_{1,2}$) (Xu et al., 2006; Xue et al., 2008; Oh et al., 2011; Yu et al., 2008). Not only the grain weight, but also other rice traits are quantitative in nature, governed by QTLs and show continuous genetic variation in natural populations and inbred lines. Characterization of QTLs is an
important step for utilizing them in rice breeding programs. The QTL mapping basically relies on detecting correlations between genetic markers and phenotypic traits in a segregating population (Weir et al., 1987; Tanksley et al., 1993; Falconer et al., 1996; Lynch et al., 1998). An efficient QTL mapping requires a large segregating population (bi-parental mapping population) such as an F2 population or Recombinant Inbred Lines (or RILs). It is a well-known fact that maximum variation is observed in the F2 generation followed by F3. Chances of identifying major QTLs is more in these initial segregating generations compared to advance RILs. Bulk segregant analysis (BSA) has been proven to be an efficient method of detecting large effect QTLs for grain yields (Vikram et al., 2012). This approach involves genotyping of the pooled DNA samples of the phenotypic extremes to identify the linked markers. Further, linked markers are run on whole population to determine QTL effects (Vikram et al., 2012). Major QTLs can be detected reliably following BSA in early segregating generations. Present study aims to identify a large effect QTL for TGW using a F2 derived F3 population through BSA.

**Results**

**Phenotypic variation in the population**

Evaluation of the Azucena × IR64 population RILs for TGW at two environments (New Delhi, India: Environment-1 and Faizabad, India: Environment-2) revealed a normal distribution pattern (Fig 1). At environment-1 number of entries ranging 20 – 25, 26 – 30, 31 – 35, 36 – 40 and <2 grams were 3, 75, 148, 10 and 2 respectively. Similarly, at environment-2 number of entries ranging 20 – 25, 26 – 30, 31 – 35 and 36 – 40 grams were 4, 96, 125 and 13 respectively (Table 1). Mean TGW of Azucena × IR64 population at environment-1 and environment-2 were 31.19 and 30.60 grams respectively. Both parents differed significantly for the TGW and transgressive segregants having grain weight more than the high grain weight parent Azucena were also observed (Table 3).

**BSA for grain weight analysis**

Two markers showed association with grain weight in the BSA. These SSR markers were RM25719 and RM5352. RM25719 and RM5352 were located on chromosome 10 at 19.8Mb and 21.1 Mb respectively. These markers were polymorphic in parents as well as two bulks as revealed in Fig 2. The band pattern of bulks (bulk high and bulk low) for both markers was similar. Band of bulk high correspond to Azucena whereas the band of bulk low correspond to IR64 for both the markers. Further, polymorphic markers adjacent to RM25719 and RM5352 were run on whole population to define the QTL boundary.

**Analysis of the QTL region identified through BSA**

Markers linked with TGW were as well as adjacent polymorphic markers were analyzed to determine QTL position and effect. Both linked markers identified in BSA were found to be significantly associated with grain weight. Two QTL peaks were identified in the analysis using two environment data. These two loci were named TGW10.1 and TGW10.2.

### Table 1. Distribution of genotypes in population for different grain weight ranges.

<table>
<thead>
<tr>
<th>Grain weight range (g)</th>
<th>Environment-1</th>
<th>Environment-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-25</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>26-30</td>
<td>75</td>
<td>96</td>
</tr>
<tr>
<td>31-35</td>
<td>148</td>
<td>125</td>
</tr>
<tr>
<td>36-40</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>&gt;40</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2. Phenotypic variation of the trait.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>DF</th>
<th>Mean sum of square</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Varieties</td>
<td>71</td>
<td>14.32</td>
<td>12.15**</td>
</tr>
<tr>
<td>Replication</td>
<td>2</td>
<td>3.37</td>
<td>2.86*</td>
</tr>
<tr>
<td>Error</td>
<td>142</td>
<td>1.17</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>215</td>
<td>5.54</td>
<td></td>
</tr>
</tbody>
</table>

Flanking marker of TGW10.1 were RM5352 and RM5666. Flanking marker of TGW10.2 were RM25719 and RM58202. The phenotypic variance explained by TGW10.1 and TGW10.2 were 21.67 and 24.49. Additive effects contributed by TGW10.1 and TGW10.2 were 2.09 and 1.99. F-values confirming significance of QTL were 27.76 and 37.06 respectively for TGW10.1 and TGW10.2 (Table 4). These two loci shared common markers indicating towards a common region (Fig 3).

**Discussion**

There is a need to double the food production in world to meet the growing food demands of human population. Increase in genetic yield potential of rice is merely 1% in spite of expected 2.4% to meet the global food demands (Rays et al., 2013). Significant efforts have been made in past to increase grain yield of rice and further increase requires focus on grain yield parameters. TGW is the most important grain parameter associated with grain yield potential of rice. Identification and characterization of major QTLs for grain weight in rice will contribute significantly towards enhancing genetic yield potential of rice. Our present analysis was performed using a RIL population developed with the cross of a japonica cultivar ‘Azucena’ known for high grain weight and well known indica rice variety ‘IR64’. TGW was scored from experiments conducted at two different environments. Distribution of genotypes for TGW was more or less normal rendering population to a fit candidate for genetic analysis (Fig 1). Genotyping of the population was carried out through BSA method. Earlier numerous QTLs have been identified through this approach (Sandhu et al., 2014, Venuprasad et al., 2009). In this approach extreme/tail lines are used for analysis due to which smaller effect QTL regions are ignored which makes it cost effective. The cost effective genotyping strategy enables breeders to analyze multiple populations simultaneously to identify major effect genomic regions associated with trait of interests (Vikram et al., 2012). We have identified two different markers on chromosome 10 associated with TGW (Table 4, Fig 2). These markers were more than 1 Mb apart. These two loci may or may not belong to the same QTL region (Fig 3). Further analysis with population of recombinants will provide a clear picture of grain weight genomics associated with this region.
Fig 1. Distribution of population lines for 1000-grain weight. On ‘Y’ and ‘X’ axes number of genotypes and TGW were shown respectively. Figure also shows a normal distribution of population for TGW.

Fig 2. BSA results for grain weight in the Azucena × IR64 population. P1 = Azucena, P2 = IR64, BL = Bulk low & BH = Bulk high. BSA results with the two markers (RM25719 and RI5352) depicted a pattern. Band pattern of P1 and P2 correspond BH and BL, respectively.

Similar to our findings two different genomic regions have been identified on chromosome 1 for grain weight (Yu et al., 2008). Similarly multiple QTLs were identified on same chromosome arms for grain weight and many other traits. As an example for grain yield under drought stress in rice sub-QTLs were reported with a large effect QTL region (Dixit et al., 2012). On chromosome 10, QTLs associated with grain weight have been reported on different arms (Ishimaru et al., 2003). The nearest marker to QTL in this study was R1629 which was positioned around 12.6 Mb. However the TGW10-1 and TGW10-2 markers which we identified in the study are more than 6 Mb apart from R1629. The two QTL loci TGW10-1 and TGW10-2 explained phenotypic variance of 21.67 and 24.49% respectively. Positive alleles were contributed by the high grain weight parent ‘Azucena’ ruling out any possibility of epistatic effect associated with these regions and strengthening the view that they are additive. For an additive trait like TGW, QTL explaining phenotypic variance of more than 10% will be good enough to be utilized in pyramiding. Further the identified QTL alleles were characterized in a set of 50 random varieties with an aim of identifying multiple alleles, if available. Two markers RM5352 and RM25719 were run on all of the 50 genotypes and revealed that only two alleles were prominent. These two alleles were observed in the population analysis too (data not presented). Fine-mapping of these two QTL regions through high density mapping in a recombinant population can provide more insights. From breeding point of view, identified QTLs-TGW10-1 and TGW10-2 if pyramided with other large effect TGW and/or grain yield QTLs will significantly increase genetic yield potential of rice. Markers linked with high TGW for effect of TGW10-1 and TGW10-2, are suitable for enhancing the rice genetic yield improvement through pyramiding approach.

Materials and Methods

Plant material

The F2:3 RIL population was developed from a cross of a japonica cultivar ‘Azucena’ and an indica variety ‘IR64’. F1 seeds were selfed to get F2 seeds. The F2 plants were grown and
PCR reaction was performed using a 15 µl reaction mixture containing 1 µl of 25 ng DNA template, 2 µl of 10X PCR buffer, 2 µl of 1mM dNTPs, 0.6 µl of MgCl2, 1 µl each of 5 µM forward and reverse primer, 1 µl of 5 U/µl Taq DNA polymerase, and 6.4 µl sterile H2O. The PCR amplification profile included initial denaturation at 94°C for 5 minutes; then 35 cycles of denaturation at 94°C for 45 seconds, annealing at 58°C for 45 seconds, and extension at 72°C for 45 seconds; and final extension of 72°C for 7 minutes and storage at 4°C. PCR products were resolved in 4% metaphor agrose gel system (Sambrook et al., 1989).

**Bulk segregant analysis**

Parental polymorphism was carried with a total of 600 simple sequence repeat (SSRs) markers distributed among all 12 chromosomes of rice from available rice genetic maps (Tenmykh et al., 2001; McCouch et al., 2002; IRGSP, 2005). Bulks were prepared using phenotypic data of the two experiments. DNA of each sample was isolated and two bulks (bulk high and bulk low) were prepared. Bulk high constituted the bulk of DNA of 10 highest grain weight lines. Equal amount of each sample was mixed to prepare the bulk so that there was equal representation of each sample in the bulk. Similarly bulk low was prepared. Further 140 polymorphic SSR markers distributed on all 12 chromosomes were run on four DNA samples- (1) Azucena (2) IR64 (3) Bulk high and (4) Bulk low. Two markers identified as polymorphic between high and low tails were run on the whole population. Two additional polymorphic markers were further run on the whole population for determination of the confidence interval.

**Statistical and QTL analysis**

Statistical analysis was performed using CROPSTAT software v7.2.3 (available at www.irri.org). To calculate entry means within a year, replications were taken as random while entries were taken as fixed. During estimation of the entry means across years, season effects were also taken as random. Single-marker regression analysis was carried out with QTL cartographer v 2.5. Marker orders were taken from published physical map of rice (IRGSP, 2005). An assumption of one million bases ~ 4 cM was made while estimating genetic distances (IRGSP, 2005). QTL analysis was carried out using QTL Network v2.1 software (Yang et al., 2008). Composite interval mapping was performed. A total of 1000 permutations were used to calculate the F-value and to control the genome-wide type I error. For QTL detection as well as QTL effects an experiment-wise significance of P = <0.01 was followed. The window size and walk speed were fixed as 1 and 0.1 cM, respectively for genome scan.

**Conclusions**

TGW is one of the most important grain yield related parameter which contributes to improve the grain yield potential of rice. A major effect genomic region associated with high grain weight was identified on chromosome 10. This region explained phenotypic variation up to 24%. QTLs identified in this region if pyramided with other major grain yield related QTLs should help in enhancing rice yield potential.
Table 4. QTLs identified in the Azuena × IR64 population for grain weight in rice.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Environment</th>
<th>QTL</th>
<th>Marker interval</th>
<th>Position interval (Mb)</th>
<th>$R^2$ (%)</th>
<th>$F$-value</th>
<th>Additive effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>GY</td>
<td>Environment 1</td>
<td>$TGW_{10}$</td>
<td>RM35352-RM4477</td>
<td>21.11 – 21.12</td>
<td>21.67</td>
<td>27.76</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>Environment 2</td>
<td>$TGW_{10}$</td>
<td>RM25719-RM3773</td>
<td>19.82 – 19.7</td>
<td>24.49</td>
<td>36.06</td>
<td>1.99</td>
</tr>
</tbody>
</table>

Table 5. List of the foreground markers used for QTL characterization.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Marker name</th>
<th>Chromosome</th>
<th>Position</th>
<th>Sequence of forward primer (5' - 3')</th>
<th>Sequence of reverse primer (5' - 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RM25719</td>
<td>10</td>
<td>19.82</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
<tr>
<td>2</td>
<td>RM3773</td>
<td>10</td>
<td>19.89</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
<tr>
<td>3</td>
<td>RM25766</td>
<td>10</td>
<td>20.54</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
<tr>
<td>4</td>
<td>RM25775</td>
<td>10</td>
<td>20.72</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
<tr>
<td>5</td>
<td>RM4477</td>
<td>10</td>
<td>21.11</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
<tr>
<td>6</td>
<td>RM5352</td>
<td>10</td>
<td>21.12</td>
<td>ACCAGATCACATGAAGAGGA</td>
<td>ACCAGATCACATGAAGAGGA</td>
</tr>
</tbody>
</table>

Acknowledgements

The authors thank the seed production division, Indian agricultural research institute (IARI), Pusa, New Delhi for providing phenotyping facility and the crop improvement division, Central rice research institute (CRRI), Cuttack, India for providing materials and other resources.

References


