Full Length Research Paper

Mapping QTLs for submergence tolerance during germination in rice

S. Abdolhamid Angaji

Tarbiat Moallem University, Postal Code 19446-84831, Tehran, Iran. E-mail: Ershad110@yahoo.com. Tel: +98-21-22587556/ +989123058891.

Accepted 10 July, 2008

To understand the genetic bases of tolerance and to identify relevant quantitative trait loci (QTLs), KHAIYAN (Aus type) was crossed with IR64, a semi-dwarf, modern variety moderately sensitive to anaerobic conditions during germination. Results of screening of BC$_1$F$_2$ lines showed that survival percentage in IR64/KHAIYAN//IR64 population ranges from 0 to 68% with an average of 30%. A linkage map was constructed with 155 polymorphic SSR markers, resulting in a map of 1483.5 cM with a mean inter-marker distance of 9.6. Four putative QTLs were detected in this population, one each on chromosomes 1 (qAG-1), 2 (qAG-2-1), 11 (qAG-11), and 12 (qAG-12). The LOD value of these QTLs ranged from 3.66 to 5.71 with phenotypic variance in the range of 12 to 29.24%. Total phenotypic variation explained by the four QTLs was about 51.4%. The additive value of the QTL on chromosome 1 was negative, indicating that the allele from KHAIYAN increased tolerance to anaerobic conditions during germination. However, the QTLs on chromosome 2, 11, and 12 were positive indicating contribution of the alleles from IR64.

Key words: Oryza sativa L, anaerobic germination, QTL analysis.

INTRODUCTION

Two main methods are commonly used for crop establishment in rice: direct seeding and transplanting. The recent water and labor shortages and the new economic trends are encouraging farmers to shift from traditional transplanting to direct seeding. In spite of the advantages of the method, direct seeding also has its disadvantages. From sowing to seedling establishment, potential damage can occur because of birds, rats, snails, and other biological factors. Physical damage can also be caused by flooding, water stress, lodging at maturity, or weeds. Weed control through hand weeding or herbicides is particularly critical in direct-seeded rice and this is aggravated by the farmers’ practice of withdrawing irrigation water 7-10 d after seeding to promote seedling establishment. Consequently, farmers who practice direct seeding use more herbicides or more labor-intensive hand weeding than farmers using transplanting. Early flooding due to rain or uneven fields can also damage germinating seeds through impaired seedling root and shoot growth. Moreover, the seeds in standing water cannot form roots and both poor anchorage and lower initial growth rate contribute to poor crop establishment and encourage growth of weeds. It is therefore important to develop rice varieties capable of germinating and growing vigorously under flooded conditions to alleviate these constraints.

Breeding for tolerance to anaerobic condition during germination has been slow due to limited knowledge on the genetics of tolerance, involvement of several complex mechanisms (Seshu et al., 1988) and complex genetic background and lack of an effective method to evaluate this trait (Ling et al., 2006). Miyoshi and Sato (1997) established anoxia conditions by flushing the germinating seeds with N$_2$ gas, but this method is difficult to use to screen for tolerance to anaerobic condition during germination on a large scale. Frantz and Bugbee (2002) used a 5 cm of static water to impose anoxia, and found that traceable dissolved O$_2$ in the water had no effect on the scoring of seed germination under anaerobic condition.

Progress in the generation of a molecular genetic map and marker for rice has made possible a new phase of mapping individual genes associated with complex traits (Yano and Sasaki, 1997) including those controlling tolerance to anaerobic condition during germination (Ling et al., 2004, 2006). Recent efforts at IRRI identified few
genotypes with greater ability to germinate under flooding (Holt and Ismail, 2002). In this study, we selected one of the most tolerant genotypes, KHAICYAN from Bangladesh, based on its consistent superior performance in subsequent replicated experiments. It is Aus type and highly resistant to bacterial blight. This variety can be used as donor parents to introduce tolerance to flooding during germination into elite cultivars and popular varieties. In order to find out the genetic bases of tolerance in this genotype, and to facilitate the rapid development of tolerant high yielding breeding lines, the above mentioned variety was crossed with IR64 as a recurrent parent (moderately sensitive) and BC2F2 population was constructed and used for discovery of QTLs associated with tolerance.

MATERIALS AND METHODS

The study was conducted at NG-01 greenhouse, Genome and Mapping (GML), the Gene Array and Molecular Marker Applications (GAMMA) and blue laboratories of IRRI from March 2004 to Jun 2007.

Plant materials

KHAICYAN as the female parent was crossed with IR64 as a male parent. The resultant F1 progeny were backcrossed to IR64 and selfed, resulting in 104 BC2F2 lines. Each plant was harvested and kept in a separate prelabeled bag. The panicles were threshed, the seeds cleaned, counted and put into the oven at 50°C for 5 d to break dormancy. From which 10 BC2F2 lines were selected based on phenotype to eliminate non-desirable characteristics, including tall and sterile plants. These selected lines were then backcrossed to IR64 to produce BC3F1. BC3F1 was then selfed to produce BC4F2 with approximately 150 lines. From each of these plants, the “best” panicle was harvested and kept in a separate bag for genotyping, while the remaining panicles were bulked and used for phenotyping.

Seed germination test

In order to test seed vigor and quality, 40 seeds were chosen at random and the seeds were sown in trays and covered with little garden soil in two replications, with 20 seeds per replication. Germination rate was assessed based on the average number of seeds that germinated over 10 days. The lines with germination rate of less than 80% were discarded.

Trait evaluation

Screening was carried out by direct dry seeding in seeding trays containing several cells and each cell holds a seed. Seeds of individual lines were sown in rows and covered by a thin layer of soil and then immediately submerged to a water depth of 10 cm. This depth was maintained for 3 weeks when percentage germination was determined.

Marker analysis

Seeds from each of the selected best panicles were grown in a single row and one leaf was harvested from each plant within row and bulked for DNA extraction. Total genomic DNA was extracted after crushing in liquid nitrogen in microfuge tubes using a Tris/SDS extraction buffer (100 mM Tris-HCl pH 8, 50 mM EDTA pH 8, 500 mM NaCl, 1.25% SDS (w/v), and 0.38 g sodium bisulfite per 100 ml of buffer) and chloroform extraction followed by ethanol precipitation. Agarose gel electrophoresis was used to estimate DNA concentration, and each sample was then diluted to approximately 5-10 ng/µL.

QTL results in the current study were compared with previously detected rice QTLs for the trait by employing RICE QTL in TRAITS option in http://www.gramene.org/. The QTLs that overlap with other studies fall into two categories: i) QTLs that share similar map and linkage groups with other QTLs and ii) QTLs with different genetic backgrounds that also fit in the same general location on the genome. The final set of marker data and QTL results were used to create a sequence-based map. The linkage map was constructed using QTL CARTOGRAPHER (version 2.5, 2007) and QGene 3.4 software (Nelson, 1997). The QTLs detected by SPA and IM corresponded well, therefore the results from IM were presented. For the IM analysis, the following parameters were selected: the BC2 population structure, 1-cM intervals, and the Kosambi function. To identify an accurate significance threshold for the trait, an empirical threshold was determined for IM using 10,000 permutations across all 12 chromosomes. For IM, the experiment-wise significance level of P < 0.01 corresponded to an average LOD > 4.3 across traits, while the level of P < 0.05 corresponded to a LOD > 3.59.

The CIM method (Basten et al., 1994) that can reduce the effects of genetic backgrounds (Zeng, 1994) was used to confirm the QTLs and identify additional QTLs that may have been masked by larger QTLs. The cofactors were selected automatically using Forward-Backward stepwise regression with F-in = 0.01 and F-out = 0.01 using QTL CARTOGRAPHER (version 2.5, 2007). Significance threshold for CIM were determined using 1,000 permutations for the population. For CIM, the experiment-wise significance level of P < 0.01 corresponded to an average LOD > 3.19, whereas the level of P < 0.05 corresponded to a LOD > 3.01.

The CIM method (Basten et al., 1994) that can reduce the effects of genetic backgrounds (Zeng, 1994) was used to confirm the QTLs and identify additional QTLs that may have been masked by larger QTLs. The cofactors were selected automatically using Forward-Backward stepwise regression with F-in = 0.01 and F-out = 0.01 using QTL CARTOGRAPHER (version 2.5, 2007). Significance threshold for CIM were determined using 1,000 permutations for the population. For CIM, the experiment-wise significance level of P < 0.01 corresponded to an average LOD > 3.19, whereas the level of P < 0.05 corresponded to a LOD > 3.01.
QTL analysis

Six hundred and ten SSR markers were used for the parental survey. Out of which 170 (27.8%) markers showed polymorphism between the two parents. Polymorphism is a measure of genetic diversity and varies with the parental combinations used. The lower percentage of polymorphism may be due to the higher degree of genetic similarity between KHAIYAN and IR64. Polymorphic markers were then applied in 120 BC$_2$F$_2$ lines in this population. Out of the 170 polymorphic markers, 155 exhibited clear bands and the rest showed no clear banding patterns. The amplified fragments ranged in size from 83 to 440 bp. A molecular map was constructed with this microsatellite mapping data using the Nipponbare/Kasalath genetic map (Harushima et al., 1998) and the physical rice map (IRGSP, 2005). Figure 2 shows the linkage map for 155 SSR markers and their interval distances. The total size of the linkage map was 1458.5 cM and the average interval size was 9.6 cM. There were few gaps with large distances mainly located on chromosomes 3, 4, 6, 9 and 10. These gaps may indicate that the genetically related parents caused the low turn in polymorphism for microsatellites in these regions.

The results of single point analysis (SPA) and interval mapping were consistent in this population and three QTLs were identified using these mapping methods; on the other hand, composite interval mapping (CIM) detected fewer QTLs (2 QTLs) and only one QTL was recognized by the three mapping methods. The QTLs conferring tolerance to flooding during germination were located on chromosomes 1 (qAG-1), 2 (qAG-2), 11 (qAG-11), and 12 (qAG-12). QTLs were named based on the nomenclature suggested by McCouch et al. (1997). In the proposed nomenclature, “q” stands for QTL and “AG” for tolerance to anaerobic conditions during germination; the number at the end of the locus name designates the position and mapped to same trait and ii) QTLs that share a similar map position, but mapped to a different trait.

In order to find an association between the identified QTLs controlling tolerance to anaerobic conditions during germination and the genes known to be regulated by anoxia, initially primer sequences of adjacent markers in the intervals within which QTLs most likely occur were aligned using Sequences-Blast in http://www.gramene.org/. By means of a Contig view icon in the same site, TIGR Gene IDs of nearby putative candidate genes were obtained and the IDs were searched out in Rice Annotation (release 5) available in the TIGR Rice Genome Annotation (http://www.tigr.org/tdb/e2k1/osa1/) to find out the function of respective candidate genes. PubMed (http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed) and Basic Query in rice mpss (http://mpss.udel.edu/rice/) were sought to investigate more on the function of candidate genes.

RESULTS

Initially, the germination rates of all lines were assessed and those with rates lower than 80% were discarded before phenotyping. Ten lines with low germinability were removed from this population in this process. Survival of the remaining lines ranged from 0 to 68% and was categorized into seven classes with the second class (10-20%) having the highest frequency (Figure 1). The overall mean was around 30%, with a variance of about 0.017 using SAS software. The character seems to vary continuously among the backcross lines and the distribution of the trait cannot easily be distinct into discrete phenotypic classes, suggesting that the trait is polygenic or quantitatively inherited.

In Figure 1, elongated tail is at the right side of the graph and more data are in the right tail than would be expected in a normal distribution. That is the data are right or positively skewed, with skewness value of around 0.53 which is considered moderately skewed since it is between 0.5 and 1 (Risk Analysis Using Crystal Ball, Training CD).
Figure 2. Molecular linkage map of microsatellite markers with LOD graphs of the QTLs for tolerance to anaerobic conditions during germination. The green bars represent putative regions of QTLs with allele effect from the tolerant parent and the brown bars indicate putative regions of QTLs with allelic effects from sensitive parent. The vertical dotted lines show the LOD threshold of 3. The QTL intervals were composed of an inner and an outer interval. The inner interval was shown as a rectangle (LOD > 3), and the outer interval as line segments (LOD < 3). The dark blocks signify the position of the centromere of the respective chromosome.
Table 1. QTLs related to tolerance to anaerobic condition during germination detected in this population.

<table>
<thead>
<tr>
<th>QTL</th>
<th>CHR.</th>
<th>Peak markerA</th>
<th>Increased effectB</th>
<th>IMC</th>
<th>LOD</th>
<th>R²</th>
<th>Additive effect</th>
<th>CIMD</th>
<th>LOD</th>
<th>R²</th>
<th>Additive effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>qAG-1</td>
<td>1</td>
<td>RM312</td>
<td>KHAIIYAN</td>
<td></td>
<td>4.23</td>
<td>25.23</td>
<td>-1.06</td>
<td>3.66</td>
<td>11</td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td>qAG-2</td>
<td>2</td>
<td>RM341</td>
<td>IR64</td>
<td></td>
<td>3.92</td>
<td>21.1</td>
<td>2.46</td>
<td>4.44</td>
<td>14.5</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>qAG-11</td>
<td>11</td>
<td>RM206</td>
<td>IR64</td>
<td></td>
<td>5.71</td>
<td>29.24</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>qAG-12</td>
<td>12</td>
<td>RM28759</td>
<td>IR64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Putative candidate genes annotated near the QTL regions associated with tolerance to anaerobic conditions during germination.

<table>
<thead>
<tr>
<th>Gene function</th>
<th>TIGR LOCUS ID</th>
<th>CHR.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycosyl transferase family 1 protein</td>
<td>LOC_Os01g46430</td>
<td>1</td>
</tr>
<tr>
<td>CTP synthase</td>
<td>LOC_Os01g46570</td>
<td>1</td>
</tr>
<tr>
<td>Acyltransferase family protein</td>
<td>LOC_Os01g19390</td>
<td>1</td>
</tr>
<tr>
<td>Patatin</td>
<td>LOC_Os01g55650</td>
<td>1</td>
</tr>
<tr>
<td>dnaK protein</td>
<td>LOC_Os01g08560</td>
<td>1</td>
</tr>
<tr>
<td>ABC transporter family protein</td>
<td>LOC_Os02g32690</td>
<td>2</td>
</tr>
<tr>
<td>Acetyl-coenzyme A synthetase</td>
<td>LOC_Os02g32490</td>
<td>2</td>
</tr>
<tr>
<td>Glutathione S-transferase, N-terminal domain containing protein</td>
<td>LOC_Os11g37730</td>
<td>11</td>
</tr>
<tr>
<td>ABC transporter</td>
<td>LOC_Os11g37700</td>
<td>11</td>
</tr>
<tr>
<td>Glycosyltransferase</td>
<td>LOC_Os11g36700</td>
<td>11</td>
</tr>
<tr>
<td>RING-H2 finger protein ATL2I</td>
<td>LOC_Os12g42540</td>
<td>12</td>
</tr>
<tr>
<td>Cysteine synthase</td>
<td>LOC_Os12g42980</td>
<td>12</td>
</tr>
<tr>
<td>Auxin responsive protein</td>
<td>LOC_Os12g43110</td>
<td>12</td>
</tr>
<tr>
<td>Actin-depolymerizing factor</td>
<td>LOC_Os12g43340</td>
<td>12</td>
</tr>
<tr>
<td>P21 protein</td>
<td>LOC_Os12g43450</td>
<td>12</td>
</tr>
<tr>
<td>Alpha-amylase/trypsin inhibitor</td>
<td>LOC_Os12g43490</td>
<td>12</td>
</tr>
<tr>
<td>Amine oxidase</td>
<td>LOC_Os12g43590</td>
<td>12</td>
</tr>
</tbody>
</table>

chromosome number on which the QTL has been mapped.

The LOD value of each QTL ranged from 3.66 to 5.71 and the percentage of the variance explained by individual QTLs varied from 12 to 29.2%. Total phenotypic variation explained by the 4 putative QTLs was 51.4% based on the PROC GLM in SAS. The additive value of the QTL on chromosome 1 was negative. The result indicates that the allele from KHAIIYAN increased tolerance to flooding during germination. But QTLs on chromosomes 2, 11, and 12 were positive, indicating that the alleles from IR64 boosted up tolerance to flooding during germination (Figure 2; Table 1).

Candidate loci in the vicinity of the QTLs associated with tolerance of anaerobic conditions during germination

A large number of putative candidate genes were annotated in the intervals within which QTLs associated with tolerance to flooding during germination were detected. The putative candidate genes were short-listed based on their relevance to the intermediates and products of alcoholic fermentation and glycolysis as well as other enzymes and pathways known to be associated with anaerobic metabolism. The short list is presented in Table 2.

DISCUSSION

In this study, the backcross population was used to detect QTLs conferring tolerance to flooding during germination. Screening for tolerance in the populations showed that the distribution of the data followed a continuous trend, suggesting quantitative inheritance. Previous studies (Gibbs et al., 2000; Fukao et al., 2003) showed that numerous enzymes changed their activity when shifting from aerobic to anaerobic metabolism, such as α-amylase, alcohol dehydrogenase (ADH), pyruvate decarboxylase (PDC), phosphofructokinase (PFK), pyrophosphate: fructose-6-phosphate 1-phosphotransferase (Kato-Noguchi, 2002). The intermediates or products of alcoholic fermentation also seem to affect seed
germinability under anoxia. This complexity of processes involved probably explains the quantitative nature of this trait, making progress in breeding more difficult to pursue through conventional methods.

One hundred and fifty five polymorphic SSR markers were applied to construct the genetic map and QTL identification. Since genetic maps play a crucial role in linkage mapping and gene discovery, it is important to consider their reconcilability with physical maps. So in this study linkage distances between SSR markers were inferred from the Nipponbare/Kasalath genetic map (Harushima et al., 1998) and the rice physical map (IRGSP, 2005). There are issues that compromise the reliability of genetic maps. As well recognized and emphasized in the description of the construction of the genetic map by Kong et al. (2002), factors such as the number of meioses studied and the informativity of the DNA markers used can impact the reliability of a genetic map. In addition, it is also known that the construction and ultimate use of genetic maps can be compromised because of missing data and simple genotyping errors (Goring and Terwilliger, 2000). As a result of problems associated with the construction of physical and genetic maps, differences with respect to the positions of loci exist not only between physical and genetic maps, but also within different physical or genetic maps. Because genetic maps play a crucial role in pedigree-based meiotic (or “linkage”) mapping and gene discovery strategies, it is important to consider their reconcilability with physical maps whenever available. In addition, it has been well documented that misspecification of genetic maps can have negative effects on linkage analyses (Goring and Terwilliger, 2000; Hackett and Broadfoot, 2003). Three types of genetic map misspecification, listed from the most severe form to the least severe, plague linkage analyses, especially in the context of “multipoint” analyses that depend crucially on reliable DNA marker positions: (1) Loci are erroneously assumed to be on a particular chromosome. (2) Loci are out of order but on the correct chromosome. (3) Loci are on the correct chromosome and in the correct order but their interlocus distances, in terms of the amount of recombinations that occurs between them, are misspecified (Nievergelt et al., 2004). To avoid these genetic map misspecifications in this study, particularly types 1 and 2, the map position for each SSR marker was identified through its position on the physical map using the available high quality sequence based map of the rice genome (IRGSP, 2005). The genetic map and physical map should give the same SSR marker order, although the distances could be different.

While different QTLs for the same trait generally do not interact with each other, the total genotypic variance explained by all QTLs for the same trait is generally much smaller than the sum of genotypic variances explained by different QTLs. In the current study, this came true. Total phenotypic variation explained by the three QTLs is around 51.4% while the sum of $R^2$ was about 90%. This contradicts the statistical prediction of additivity of independent variables. Factors such as environmental influences, physiological pleiotropy, QTL with effects that are too small effects to be detected, covariance between closely linked QTLs, and epistasis may partially explain these discrepancies (Li et al., 1997).

Thus far, there have been only two studies on QTL mapping for tolerance to anaerobic conditions during germination in rice. In one of the studies, Ling et al. (2006) found two QTLs associated with tolerance to hypoxia during germination on chromosomes 5 and 11. The QTL on chromosome 11 is consistent with $qAG-11$ identified in the population the peak SSR marker is RM21.

Several studies have indicated that many genetic factors controlling important rice traits are clustered in chromosomal blocks (Cai and Morishima, 2002). An example is the association between seed dormancy and seed longevity that has been reported before (Ota and Takemura, 1970; Siddique et al., 1988). Ikehashi (1973) observed that some indica varieties had long seed longevity in spite of having weak seed dormancy. Miura et al. (2002) identified QTLs for seed dormancy using BILs (BC$_{F_3}$) derived from Nipponbare/Kasalath//Nipponbare. The location of $qAG-11$ in IR64/KHAIYAN//IR64 BC population will most likely correspond to one of the QTLs ($qSD-11$) detected by Miura et al. (2002) on chromosome 11 using RFLP markers, and C189 is the nearest marker locus of the putative QTL.

To establish MAS to incorporate tolerance of anaerobic conditions during germination in rice breeding programs, it is necessary to find additional and more diagnostic DNA markers that are tightly linked with the QTLs of interest. We annotated the regions in the vicinity of the identified QTLs and short-listed the genes that are more likely associated with tolerance to hypoxia or anoxia based on published functional evidences (Table 2). This list will form the entry point for future studies to further shorten the list after fine-mapping of these QTLs and subsequently validate the functional roles of the genes involved. The isolation of the genes at the QTLs for tolerance to anaerobic conditions will help design more precise gene-specific diagnostic markers for MAB and also advance our understanding of the genetic and physiological mechanisms of tolerance. Besides, this can also facilitate searching for new and stronger alleles for tolerance.

Efficiency of marker-aided selection in breeding programs depends on the strength of linkage between molecular markers and the target trait. Traditionally, anonymous molecular markers are used to establish linkage with a phenotype. However, even for tightly linked markers, the effectiveness of marker aided selection is greatly diminished by the occasional uncoupling of the marker from the trait during many cycles of meiosis in a breeding program. With the availability of large genome databases, it is now possible to predict putative function
of a gene based on sequence information, thus enabling the identification of candidate genes involved in a particular biochemical pathway. These candidate genes, or DNA sequences with predicted function, are used as molecular markers to associate with phenotypes expressed in segregating populations or genetic stocks (Huh et al., 2001; Thorup et al., 2000). The idea is to propose previously sequenced genes of known function that could correspond to major loci (Mendelian trait loci, MTLs, or QTLs). The CGs may be structural genes or genes involved in the regulation of a metabolic pathway. The working hypothesis assumes that a molecular polymorphism within the CG is related to phenotypic variation. The CG approach has been used with success in human and animal genetics (Rothschild and Soller, 1997) and, since 1990s, in plant genetics (Byrne and Mc-Mullen, 1996). Many traits of agricultural significance exhibit quantitative inheritance, which is often the result of multiple genes influenced by the environment. Because of the multiplicity of genes defining a complex trait, their partial effects on phenotypic variation and their imprecise localization on genetic maps, the CG approach is more adapted to QTL characterization than positional cloning or insertion mutagenesis. However, fine-mapping coupled with map-based cloning recently resulted in the first cloned QTLs in tomato (Frary et al., 2000; Fridman et al., 2000) and more recently, in rice (Xu et al., 2006). The definition of a CG differs by discipline. Physiologists consider CG as all genes involved in the expression of a given trait, whereas geneticists consider as only polymorphic genes putatively involved in the trait variation. Thus, for geneticists, CGs differ based on the crosses being studied. For plant geneticists, a CG is any gene putatively involved in trait variation, based on its biological function and/or its map localization. In this study we aim to adapt the definition of plant geneticists: CGs are either genes with molecular polymorphisms genetically linked to major loci or QTLs, or genes with molecular polymorphisms statistically associated with variation of the trait being studied (Rothschild and Soller, 1997).

REFERENCES


