Analysis of Genome Sequence Variations Among Three U.S. Rice Varieties Showing Differential Quantitative Disease Resistance to Bacterial Panicle Blight and Sheath Blight

Bishnu Shrestha¹, Dong-Ha Oh², Maheshi Dassanayake², Jong Hyun Ham¹*

¹Department of Plant Pathology and Crop Physiology, Louisiana State University Agricultural Center, Baton Rouge, USA
²Department of Biological Sciences, Louisiana State University, Baton Rouge, USA

*Corresponding author: Jong Hyun Ham, Department of Plant Pathology and Crop Physiology, Louisiana State University Agricultural Center, Baton Rouge, LA 70803, USA. Tel: +12255786798; Email: jham@agcenter.lsu.edu


Received Date: 25 April, 2018; Accepted Date: 01 May, 2018; Published Date: 09 May, 2018

Abstract

Bacterial Panicle Blight (BPB) and Sheath Blight (SB) are major rice diseases in the southeastern United States, and only quantitative disease resistance is known for these diseases. We analyzed draft genome sequence data for three U.S. rice varieties showing differential disease resistance traits for BPB and SB; Trenasse (long-grain, susceptible to BPB and SB), Bengal (medium-grain, susceptible to BPB and SB), and Jupiter (medium-grain, partial resistance to BPB and SB). Comparative genome sequence analysis along with 50 rice accessions revealed that the three US varieties are genetically close and clustered together, separated from other 50 accessions. According to this analysis, the long-grain semi-dwarf variety Trenasse is a tropical japonica type carrying a fraction of indica genome, while the medium-grain varieties Bengal and Jupiter are admixtures of tropical and temperate japonica types. Consistent with the breeding history and the phenotypic trait in grain-shape (but not with the phenotype in disease resistance), more variations were found in the Jupiter/Trenasse and Bengal/Trenasse pairs compared to the Jupiter/Bengal pair. The whole genome sequence information of these US rice varieties will be a useful resource for genetic studies of disease resistance to BPB and SB as well as development of new disease-resistant lines.

Introduction

Bacterial Panicle Blight (BPB) and Sheath Blight (SB) are important chronic diseases of rice in the southeastern United States, as well as other parts of rice-growing regions around the world [1-4]. BPB is caused by two bacterial pathogens, Burkholderia glumae and B. gladioli, and the phytotoxin toxoflavin is known as the major virulence factor of the pathogens [5]. Oxolinic acid is somewhat effective to control BPB, but this chemical is not registered to use for agricultural purpose in the U.S., and resistant strains of B. glumae have been reported indicating the limitation of this chemical as a reliable control measure [5]. SB is caused by the fungal pathogen, Rhizoctonia solani, and fungicide application is the primary way to control this disease. However, recent reports of R. solani isolates resistant to Strobilurin-type fungicides [6,7] indicate the urgent need of developing reliable alternative management measures for this disease, including the cultivation of disease-resistant varieties. Quantitative (or partial) disease resistance, which is usually conferred by multiple Quantitative Trait Loci (QTLs), is thought to be the primary disease resistance mechanism of rice to BPB and SB. Qualitative (or complete) disease resistance, which involves specific interactions between a resistance gene of the host and its cognate a virulence gene of the pathogen [8], has not been found in BPB or SB of rice. Major QTLs associated with the disease resistance to BPB have been identified from several rice varieties. qBPB3-1 was identified on the short arm of chromosome 3 from the resistant variety, Teqing [9]. Later, qRBS1 (renamed later as RBG1) was mapped on the short arm of chromosome 10 with the 393-kb interval from the resistance variety, Nona Bokra [10]; and RBG2 was identified on the long arm of chromosome 1 with the 502 kb interval from the resistant traditional lowland variety, Kele [11]. Genetics of SB resistance
has been studied more intensively and widely compared to that of BPB resistance. More than 50 QTLs for SB resistance have been identified, and candidate genes responsible for SB resistance have also been found within some QTL regions [12-19]. Among the QTLs identified, qSB9-2 on chromosome 9 and qSBR11-1 on chromosome 11 are known to be major QTLs identified from multiple rice varieties [17,20]. Nevertheless, our knowledge of rice disease resistance to BPB and SB is still fractional and rudimentary.

Three U.S. rice varieties, Trenasse, Jupiter and Bengal, have been used for our genetic studies of the disease resistance to BPB and SB. These rice varieties cultivated in the southeastern United States show different phenotypes in terms of major agronomic traits, including grain shape and quantitative disease resistance to BPB and SB; in that Jupiter (a medium-grain variety) shows quantitative resistance to BPB and SB [21], while Bengal (a medium-grain variety) and Trenasse (a long-grain variety) are highly susceptible to both diseases [22,23]. Nevertheless, genetic studies of the disease resistance to BPB and SB with these materials have been hindered due to the lack of polymorphic markers to be used for linkage mapping of QTLs. Whole genome sequencing of rice accessions using a High-Throughput DNA Sequencing (HTS) platform provides excellent opportunities to determine genome-wide sequence variations associated with various traits of rice, such as disease resistance, and to develop more reliable molecular markers in a high-throughput and cost-efficient way [24-26]. Especially, HTS data are very useful for identification of DNA polymorphisms between genetically close genotypes and for fine mapping. For example, comparative analyses have been conducted for the identification of DNA polymorphisms within japonica or indica rice varieties [27-31], and the whole genomes of 13 rice inbred lines derived from US varieties were analyzed for the identification of candidate genes for sheath blight resistance [32]. In this study, we sequenced and analyzed the whole genome sequences of three US rice varieties, Trenasse, Jupiter and Bengal, which represents differential phenotypes in disease resistance/susceptibility to BPB and SB, using an HTS platform (Illumina HiSeq1000) in an attempt to develop new sequence-based molecular markers and to find a genome information basis for future genomic and genetic studies of rice disease resistance.

Materials and Methods

Rice Plants and DNA Extraction

One-week-old seedlings of the rice varieties, Jupiter (medium-grain, and moderately resistant to BPB and SB), Trenasse (long-grain, and susceptible to BPB and SB) and Bengal (medium-grain, susceptible to BPB and SB) were used to extract genomic DNA for whole genome sequencing. DNeasy Plant Mini Kit (Qiagen, Valencia, CA) was used for DNA extraction following manufacturer’s instructions. The genomic DNA library for sequencing was prepared using a Nextera DNA Library Preparation Kit (Illumina Inc., San Diego, CA), and 100-bp paired-end sequencing was processed using the Illumina HiSeq1000 platform (Illumina Inc., San Diego, CA) at Virginia Bioinformatics Institute (VBI) Genomics Lab at Virginia-Tech (Blacksburg, VA).

Mapping and Identification of Variants in Genome Sequences

The quality of the sequence reads were examined using Fast QC [33], and cleaned high quality reads were aligned to the rice reference genome version 7 released by the International Rice Genome Sequence Project (IRGSP) for the japonica rice variety Nipponbare, using Bowtie 2 [34]. Genome-wide variants including Single Nucleotide Polymorphisms (SNPs) and small insertions and deletions (indels) between the reference genome and three rice varieties were identified and processed using SAMtools [35], and annotated using SnpEff v3.5e [36].

Population Structure Analysis

Genetic relatedness of the three rice varieties with 50 rice accessions, including temperate and tropical japonica, aromatic and indica types, were analyzed by using FRAPPE [37]. SNP data of the 50 rice accessions for comparison were from the study by Xu, et al. [38], and the three US varieties for this study were analyzed in terms of admixture proportions with increasing value of K (number of clusters) from 3 to 7.

Pairwise Comparison

Pairwise comparisons between two varieties (Jupiter vs. Trenasse, Jupiter vs. Bengal, and Bengal vs. Trenasse) were performed with the help of vcftools, using the vcf files obtained from SnpEff analysis [39]. The output file from each comparison was filtered for common variants present in the varieties, which were identified from the comparison with the Nipponbare reference genome. Those variants between the varieties were again annotated and classified based on their effect on various regions in the genome and their functional type, using SnpEff v3.5e [36].

Statement of Reagent and Data Availability

All the rice DNA sequence data used for this study were deposited to the NCBI SRA (accession numbers: SRX4017380, SRX4017381, and SRX4017382). DNA samples of the rice varieties, Bengal, Jupiter and Trenasse, were sent to the researchers upon request.

Results and Discussion

High-Throughput Sequencing (HTS) Data Obtained in This Study

In this study, the genomes of three US rice varieties, Jupiter, Trenasse and Bengal, were sequenced for comparative analysis of genome sequence variations. Fifty to 84 million of 100-bp paired-
end reads were obtained from each variety, resulting in 12X to 18X coverage based on the reference genome of ‘Nipponbare’ (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>Jupiter</th>
<th>Trenasse</th>
<th>Bengal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total reads (in millions)</td>
<td>79</td>
<td>84</td>
<td>50</td>
</tr>
<tr>
<td>Coverage</td>
<td>18X</td>
<td>19X</td>
<td>12X</td>
</tr>
<tr>
<td>Mapped with chromosomal genome (%)</td>
<td>95.78</td>
<td>91.02</td>
<td>96.33</td>
</tr>
<tr>
<td>Mapped with organelle genome (%)</td>
<td>13.21</td>
<td>11.05</td>
<td>9.43</td>
</tr>
</tbody>
</table>

Table 1: Total sequence reads (100-bp paired end) obtained from the high-throughput sequencing in this study, and percentage of the sequence reads mapped to the reference genome of *Oryza sativa* subsp. *japonica* (Nipponbare, IRGSP pseudomolecule version 7).

An average of more than 91% were aligned with the reference genome of the International Rice Genome Sequencing Project (IRGSP) pseudomolecule version 7 (Table 1). The reads from Bengal has the highest alignment percentage with 96.33% when mapped with the chromosomal reference sequence, followed by the reads from Jupiter (95.78%) and Trenasse (91.02%). In case of the HTS data mapped to the organelle sequences, the total sequence data of Bengal contained least portion of organelle sequences (9.43%), while 11.05% and 13.21% of the total sequence data were mapped to the organelle sequences with Trenasse and Jupiter, respectively (Table 1). Combined proportions of chromosomal and organelle sequences exceeded 100% in all varieties, indicating that some portions of sequence data were mapped to both chromosomal and organelle genomes.

### Population Structure Analysis Using SNPs Identified From HTS

When compared to the reference genome (Nipponbare), total 1,007,294, 817,884, and 2,139,891 SNPs were identified in Bengal, Jupiter, and Trenasse genome sequences, respectively, indicating that the tropical *japonica* medium-grain varieties, Bengal and Jupiter, are genetically more closely related to Nipponbare (a temperate *japonica* variety) than the tropical *japonica* long-grain variety, Trenasse (Table 2).

<table>
<thead>
<tr>
<th></th>
<th>Bengal</th>
<th>Jupiter</th>
<th>Trenasse</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNPs</td>
<td># of SNPs</td>
<td># of SNPs per Mb</td>
<td># of SNPs</td>
</tr>
<tr>
<td>Chromosome 1</td>
<td>135,180</td>
<td>3,124</td>
<td>88,362</td>
</tr>
<tr>
<td>Chromosome 2</td>
<td>78,248</td>
<td>2,177</td>
<td>73,313</td>
</tr>
<tr>
<td>Chromosome 3</td>
<td>41,286</td>
<td>1,134</td>
<td>39,733</td>
</tr>
<tr>
<td>Chromosome 4</td>
<td>147,305</td>
<td>4,149</td>
<td>98,462</td>
</tr>
<tr>
<td>Chromosome 5</td>
<td>52,613</td>
<td>1,756</td>
<td>62,823</td>
</tr>
<tr>
<td>Chromosome 6</td>
<td>68,284</td>
<td>2,185</td>
<td>57,508</td>
</tr>
<tr>
<td>Chromosome 7</td>
<td>65,303</td>
<td>2,199</td>
<td>70,094</td>
</tr>
<tr>
<td>Chromosome 8</td>
<td>63,626</td>
<td>2,237</td>
<td>62,114</td>
</tr>
<tr>
<td>Chromosome 9</td>
<td>27,344</td>
<td>1,188</td>
<td>16,854</td>
</tr>
<tr>
<td>Chromosome 10</td>
<td>130,505</td>
<td>5,623</td>
<td>135,193</td>
</tr>
<tr>
<td>Chromosome 11</td>
<td>150,632</td>
<td>5,190</td>
<td>68,711</td>
</tr>
<tr>
<td>Chromosome 12</td>
<td>46,968</td>
<td>1,706</td>
<td>44,717</td>
</tr>
<tr>
<td>Total: 1,007,294</td>
<td>Average: 2,699</td>
<td></td>
<td>Total: 817,884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th># of insertions</th>
<th># of insertions per Mb</th>
<th># of insertions</th>
<th># of insertions per Mb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome 1</td>
<td>8,630</td>
<td>199</td>
<td>5,765</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10,276</td>
<td>237</td>
</tr>
</tbody>
</table>
Table 2: Number of variants on individual chromosomes identified between the reference genome, Nipponbare and the three rice cultivars.

Same patterns were also observed with the numbers of insertions and deletions (Table 2). Among the SNPs identified, 1,188,460 non-ambiguous and biallelic SNP positions shared with at least one of the 50 rice accessions analyzed in the previous study by Xu, et al. (2012) were used to reconstruct the population structure including the three varieties along with the 50 previously sequenced rice accessions, using the program FRAPPE [37]. Among the number of populations (K) from 3 to 7 to set a population structure of the 53 accessions tested, FRAPPE produced a population structure with the highest likelihood at K=7 (Figure 1).
In this analysis, it was revealed that Trenasse is mostly *tropical japonica* type with a trace of *indica* genome, while Jupiter and Bengal are admixtures of *tropical* and *temperate japonica* types (Figure 1). This result is congruent with the previous work by Lu, et al. [40], which studied the population structure and breeding patterns of 145 U.S. rice varieties based on genotypes of 169 Simple Sequence Repeat (SSR) markers. In that study, it was revealed that US varieties of *tropical japonica* medium-grain were positioned between those of *temperate japonica* and *tropical japonica* long-grain in terms of the genetic distance determined based on the SSR genotypic data [40]. The small portion of the Trenasse genome attributed to the *indica* type is seemingly derived from the *indica* germplasms, which were utilized for introducing beneficial agronomic traits (e.g. semi-dwarf stature and high yield) to US rice varieties.

**Pairwise Comparisons of Single Nucleotide Polymorphisms (SNPs) and Small Insertions and Deletions (InDels)**

SNPs and InDels were first identified between the reference genome and each of three rice varieties, in which Trenasse and Jupiter showed the highest and lowest number of variations, respectively, when compared to the reference genome (Table 2). By comparing the polymorphism profiles compared to the reference genome, we also discovered variations between Jupiter...
and Trenasse, Jupiter and Bengal, and Trenasse and Bengal (Tables 1-3).

<table>
<thead>
<tr>
<th></th>
<th>Jupiter vs Trenasse</th>
<th>Jupiter vs Bengal</th>
<th>Trenasse vs Bengal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transitions (Ts)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/T</td>
<td>708,909</td>
<td>264,152</td>
<td>678,615</td>
</tr>
<tr>
<td>G/A</td>
<td>708,530</td>
<td>264,109</td>
<td>678,632</td>
</tr>
<tr>
<td>Total</td>
<td>1,417,439</td>
<td>528,261</td>
<td>1,357,247</td>
</tr>
<tr>
<td><strong>Transversions (Tv)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C/G</td>
<td>102,800</td>
<td>38,989</td>
<td>97,671</td>
</tr>
<tr>
<td>A/T</td>
<td>167,197</td>
<td>64,324</td>
<td>161,360</td>
</tr>
<tr>
<td>A/C</td>
<td>142,652</td>
<td>54,174</td>
<td>136,051</td>
</tr>
<tr>
<td>G/T</td>
<td>142,445</td>
<td>53,568</td>
<td>136,464</td>
</tr>
<tr>
<td>Total</td>
<td>555,094</td>
<td>211,055</td>
<td>531,546</td>
</tr>
<tr>
<td>Ts/Tv</td>
<td>2.55</td>
<td>2.50</td>
<td>2.55</td>
</tr>
</tbody>
</table>

Table 3: Number of transitions and transversions in SNPs identified from the comparisons, Jupiter vs. Trenasse, Jupiter vs. Bengal, and Trenasse vs. Bengal.

The distribution of SNPs at the chromosome level was observed by calculating the density of identified variations in 1-Mb intervals in each comparison (Figure 2A). Among the three pair-wise comparisons, the Jupiter/Trenasse and Trenasse/Bengal pairs show higher numbers of SNPs per 1Mb genome than the Jupiter/Bengal pair (Figures 2A and 2B).

Figure 2: Frequency and distribution of SNPs in individual chromosomes when compared between Jupiter and Trenasse, Jupiter and Bengal, and
Trenasse and Bengal. A) Frequency of SNPs within 1-Mb window on individual chromosomes identified from pairwise comparisons between Jupiter and Trenasse, Jupiter and Bengal, and Trenasse and Bengal. Blue-colored circle, red-colored square, and green-colored triangle in the graph represent SNP frequency between Jupiter and Trenasse, Jupiter and Bengal, and Trenasse and Bengal, respectively. B) Average frequency of SNPs per 1 Mb in each chromosome. Bars with blue-colored, red-colored, and green-colored in the graph represent the density of SNPs between Jupiter and Trenasse, Jupiter and Bengal, and Trenasse and Bengal, respectively.

In all three pairwise comparisons, the highest density of SNPs was found on chromosome 11 (9,375, 4,312, and 8,501 per Mb, respectively), while the lowest density of SNPs was found on chromosome 2 in Jupiter/Trenasse and Trenasse/Bengal (2,525 and 2,381 per Mb, respectively), and on the chromosome 12 in Jupiter/Bengal (1,056 per Mb) (Table 1). Small insertions and deletions (InDels), ranged from 1 to 18 bp, were also analyzed with each comparison. Like SNPs, larger numbers of InDels were detected from the Jupiter/Trenasse and Bengal/Trenasse comparisons compared to the Jupiter/Bengal comparison (Figure 3A and 3B, Table 2 and 3).

**Figure 3:** Average frequency of insertions (A) and deletions (B) per 1 Mb on individual chromosomes identified from pairwise comparisons between Jupiter and Trenasse, Jupiter and Bengal, and Trenasse and Bengal. Bars with blue-colored, red-colored, and green-colored in the graph represent the density of insertions or deletions between Jupiter and Trenasse, Jupiter and Bengal, and Trenasse and Bengal, respectively.
As shown in (Figure 4), number of InDels decreases exponentially in proportion to the sizes of InDels (Figure 4).

The density of variations (SNPs and InDels) between the reference genome ‘Nipponbare’ and the three US varieties in this study was much higher than that between ‘Nipponbare’ and ‘Omachi’, another temperate japonica variety [28], which is likely due to the genomic portions of tropical japonica in the US varieties (Figure 1). In other studies, comparable levels of variations have been observed between ‘Nipponbare’ and indica varieties [29, 31], as well as between ‘Nipponbare’ and other elite japonica varieties including cold temperature-tolerant Hokkaido varieties [24,27,41]. In pairwise comparisons among the three US rice varieties, variants from the Jupiter/Bengal pair were much lower compared to those from the Jupiter/Trenasse or the Bengal/Trenasse pair (Figures 2 and 3, and Tables 1-3), which is also consistent with their genomic divergence revealed in this study (Figure 1).

Nucleotide Substitutions

SNPs can be classified into transitions (C/T and G/A) and transversions (C/G, T/A, A/C, and G/T). In this study, the frequency of transitions was higher than that of transversions in all three comparisons. Among transitions, little difference was found between the C/T substitution and the G/A substitution in all comparisons (Table 3).

For transversions (A/T, C/G, A/C, and G/T), however, the A/T substitution was most frequent with more than 60% higher numbers compared to the least frequent substitution, C/G, in all the comparisons (Table 3). Among the three comparisons, Jupiter/Bengal showed lower numbers of substitutions than Jupiter/Trenasse and Trenasse/Bengal. The transitions (Ts)/Transversions (Tv) ratio was ≥2.5 in all cases (Table 3), which is higher than the previous study on rice [30,31]. The higher ratio was the result of higher transitions substitutions compared to transversions, indicating a ‘transition bias.’ It is suggested that transition bias occurs in natural selection because transition may conserve the protein structure better than transversions [42]. Transition bias has been previously reported in rice and chickpea [30,31,43]. Methylation causes higher frequency of C to T mutation, so higher C/T substitutions might have occurred compared to G/A [44]. Furthermore, A/T substitutions were abundance in transversions compared to other remaining substitutions C/G, A/C and G/T (Table 3), which is similar to the previous report on rice [30].

Positions of Variants in Different Regions of The Genome

Frequencies of SNPs and InDels in various genomic features, including intergenic region, upstream and downstream of gene models, UTR5’ and 3’, exon, and intron, were determined for all three comparisons (Table 4), using SnpEff v3.5e [36]. Those variants in the upstream of genes may have role in altering regulation of various downstream gene expression, which will ultimately alter phenotypic traits [45,46]. Variants in exons, especially those causing amino acid changes and frameshifts (i.e. non-synonymous SNPs and InDels causing frameshifts), may directly affect the functionality of the encoded protein. Regardless of the regions in the genome, highest number of variants was
detected between Jupiter and Trenasse, while lowest one was between Jupiter and Bengal (Table 4), which is congruent with other data shown in this study.

<table>
<thead>
<tr>
<th></th>
<th>Jupiter vs Trenasse</th>
<th>Jupiter vs Bengal</th>
<th>Trenasse vs Bengal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Upstream</strong></td>
<td>1,920,546</td>
<td>704,179</td>
<td>1,844,469</td>
</tr>
<tr>
<td><strong>Intergenic</strong></td>
<td>1,347,790</td>
<td>494,130</td>
<td>1,284,046</td>
</tr>
<tr>
<td><strong>Intron</strong></td>
<td>393,156</td>
<td>143,872</td>
<td>373,427</td>
</tr>
<tr>
<td><strong>UTRs</strong></td>
<td>62,961</td>
<td>21,934</td>
<td>61,441</td>
</tr>
<tr>
<td><strong>Exon</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(nsSNP)</td>
<td>231,390</td>
<td>92,272</td>
<td>221,997</td>
</tr>
<tr>
<td>(sSNP)</td>
<td>208,179</td>
<td>86,709</td>
<td>205,009</td>
</tr>
<tr>
<td>(Frameshift)</td>
<td>5,212</td>
<td>2,069</td>
<td>4,951</td>
</tr>
<tr>
<td><strong>Intragenic</strong></td>
<td>203</td>
<td>68</td>
<td>203</td>
</tr>
<tr>
<td><strong>Downstream</strong></td>
<td>1,837,883</td>
<td>679,579</td>
<td>1,770,798</td>
</tr>
</tbody>
</table>

*a* Upstream/Downstream: Variants in the region which is up to 5K base upstream/downstream of a gene. 
*Intergenic*: Variants that occur in intergenic but not in upstream/downstream regions. 
*UTRs*: Variants that hit 5′ and 3′ untranslated region. 
*Intragenic*: Variants that occur within a gene but fall outside of all transcript features.

Table 4: Annotation of variants at various genomic regions identified after pairwise comparisons among three rice cultivars.

### Synonymous and Non-Synonymous SNPs

51.6 - 52.6% of SNPs in the coding regions (CDS) were non-synonymous (nsSNPs), while 47.4 - 48.4% were synonymous SNPs (sSNPs) in all the three comparisons, resulting in the ratio of non-synonymous SNPs to synonymous SNPs (nsSNPs/sSNPs) to be around 1.06 - 1.11 (Table 5).

<table>
<thead>
<tr>
<th></th>
<th>Jupiter vs Trenasse</th>
<th>Jupiter vs Bengal</th>
<th>Trenasse vs Bengal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nsSNPs</td>
<td>231,390</td>
<td>92,272</td>
<td>221,997</td>
</tr>
<tr>
<td>%</td>
<td>52.6</td>
<td>51.6</td>
<td>52</td>
</tr>
<tr>
<td>nsSNPs/sSNPs</td>
<td>1.11</td>
<td>1.06</td>
<td>1.08</td>
</tr>
<tr>
<td>sSNPs</td>
<td>208,179</td>
<td>86,709</td>
<td>205,009</td>
</tr>
<tr>
<td>%</td>
<td>47.4</td>
<td>48.4</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 5: Synonymous and non-synonymous SNPs in coding sequences.

This value is similar to previous studies on *indica*, and *tropical* and *temperate japonica* rice [30,31,47], in which the nsSNP/sSNP ratios were around 1.2. It has been reported that the nsSNPs/sSNPs ratio tends to be lower in protein families with essential biological functions, such as cellulose synthases, but higher in the protein families with regulatory functions [38,47,48].

### Conclusion

In this study, we conducted a genome-wide comparative analysis of the three US rice varieties with different quantitative resistances to BPB and SB, and detected SNP- and InDel-based polymorphisms among them. Regarding that whole genome sequence data have been considered as an excellent source for the development of reliable molecular markers [49], the information of genomewide sequence variations gained from this study is a useful resource to develop new molecular markers for future genetic studies and marker-assisted breeding of disease resistant rice using US rice varieties. Due to the close genetic relatedness among many US rice varieties, marker development relying on random screening has often been very inefficient and costly. The whole genome sequence information from this study and other similar studies with US rice accessions will greatly improve the efficiency in the development of new molecular markers, avoiding tedious screening processes with random candidate markers. In addition, this study will provide valuable information for functional and molecular studies of quantitative rice disease resistance.
Acknowledgements

This study was supported by the USDA NIFA (Hatch Project #: LAB93918 and LAB94203), the Louisiana State University Agricultural Center, and the Louisiana Rice Research Board. D.-H. O. and M. D. were supported by the National Science Foundation (MCB-1616827) and the Next Generation BioGreen21 Program (PJ0011379) of the Rural Development Administration, Republic of Korea.

References


