Single Nucleotide Polymorphism (SNP) Discovery and Kompetitive Allele-Specific PCR (KASP) Marker Development with Korean *Japonica* Rice Varieties

Kyeong-Seong Cheon¹, Jeongho Baek¹, Young-il Cho², Young-Min Jeong², Youn-Young Lee², Jun Oh¹, Yong Jae Won³, Do-Yu Kang¹, Hyoja Oh¹, Song Lim Kim¹, Inchan Choi¹, In Sun Yoon⁵, Kyung-Hwan Kim¹, Jung-Heon Han¹, Hyeonso Ji¹*

¹Department of Agricultural Biotechnology, National Institute of Agricultural Sciences, Rural Development Administration (RDA), Jeonju 54874, Korea
³Department of Central Area Crop Science, National Institute of Crop Science, Rural Development Administration (RDA), Suwon 16429, Korea

**ABSTRACT** Genome resequencing by next-generation sequencing technology can reveal numerous single nucleotide polymorphisms (SNPs) within a closely-related cultivar group, which would enable the development of sufficient SNP markers for mapping and the identification of useful genes present in the cultivar group. We analyzed genome sequence data from 13 Korean *japonica* rice varieties and discovered 740,566 SNPs. The SNPs were distributed at 100-kbp intervals throughout the rice genome, although the SNP density was uneven among the chromosomes. Of the 740,566 SNPs, 1,014 SNP sites were selected on the basis of polymorphism information content (PIC) value higher than 0.4 per 200-kbp interval, and 506 of these SNPs were converted to Kompetitive Allele-Specific PCR (KASP) markers. The 506 KASP markers were tested for genotyping with the 13 sequenced Korean *japonica* rice varieties, and polymorphisms were detected in 400 KASP markers (79.1%) which would be suitable for genetic analysis and molecular breeding. Additionally, a genetic map comprising 205 KASP markers was successfully constructed with 188 F₂ progenies derived from a cross between the varieties, Junam and Nampyeong. In a phylogenetic analysis with 81 KASP markers, 13 Korean *japonica* varieties showed close genetic relationships and were divided into three groups. More KASP markers are being developed and these markers will be utilized in gene mapping, quantitative trait locus (QTL) analysis, marker-assisted selection and other strategies relevant to crop improvement.

**Keywords** Korean *japonica* rice, Resequencing, SNP, Kompetitive Allele-Specific PCR (KASP) marker, Genotyping

**INTRODUCTION**

Rice (*Oryza sativa* L.) is one of the important food crops in the world and has dramatically changed owing to the green revolution. Varieties of common rice are genetically differentiated into *indica* and *japonica* varieties, and *japonica* rice varieties are the main food item in East Asia (Mackill 1995; Lu et al. 2002; Garris et al. 2005). In general, a great deal of diversity has been observed between *indica* and *japonica* varieties, whereas little diversity is observed among *japonica* varieties (Glaszmann 1987; Ni et al. 2002; Gao and Innan 2008). The diversity level among Korean *japonica* rice varieties is much lower than that among whole *japonica* rice varieties. Nonetheless, Korean *japonica* rice varieties exhibit marked phenotypic differences in important agronomic traits, including disease
resistance, plant architecture, eating quality, and yield. Therefore, the development of DNA markers for discriminating among Korean *japonica* rice varieties would enable the mapping of useful genes within Korean *japonica* rice varieties and would accelerate molecular breeding of Korean *japonica* rice varieties. Recently, next-generation sequencing (NGS) has enabled the identification of numerous single nucleotide polymorphisms (SNPs) between highly homologous genomes, and it became possible to develop enough DNA markers for gene mapping with populations derived from crosses between closely-related cultivars through resequencing of the parental varieties.

Since the 1980s, molecular marker systems in rice breeding have advanced, from restriction fragment length polymorphism (RFLP) and randomly amplified polymorphic DNA (RAPD) to amplified fragment length polymorphism (AFLP), cleaved amplified polymorphic sequence (CAPS), and simple sequence repeats (SSRs) in rice genetics and breeding (Bernardo 2008; Toal et al. 2016; Shabir et al. 2017). There are limited numbers of molecular markers for the construction of genetic maps and quantitative trait locus (QTL) mapping studies in Korean *japonica* rice breeding because many of the rice markers lack sufficient polymorphism among the temperate *japonica* rice varieties (Ni et al. 2002; Garris et al. 2005; Hori et al. 2017). To resolve this problem, genome sequence variation analysis using NGS technology that can rapidly generate information regarding high-throughput SNP discovery has been used in attempts to develop effective molecular marker systems for genotyping and genetic analysis among Japanese *japonica* rice varieties. For example, through resequencing the *japonica* variety Koshihikari, total 67,051 SNPs were identified in comparison with another *japonica* rice variety Nipponbare (Yamamoto et al. 2010). Moreover, a core set of 768 SNPs were discovered by comparison of genome sequences between two Japanese rice varieties (Riku132 and Eiko), and these core SNPs could be efficiently and reliably used in the classification of the Japanese rice population (Nagasaki et al. 2010). Although NGS technologies have been used to re-sequence diverse rice genomes for genotyping, the detected variations have only been made available for genotyping on array-based platforms, with the result that the high cost of genotyping and the complexity of data processing limit their adoption for large-scale genotyping (Ertiro et al. 2015). Therefore, cost-effective and rapid methods for large-scale genotyping in *japonica* varieties and/or their progenies must be established to help researchers in terms of genomic analysis and breeding.

Of the high-throughput SNP-genotyping platforms, such as TaqMan and Kompetitive Allele-Specific PCR (KASP) assays, the TaqMan SNP assay has limited capacity to be used for SNP multiplexing and involves increased costs for the synthesis of allele-specific assays (Yuan et al. 2014). In contrast to the TaqMan assay, the KASP assay is a single-step genotyping technology that reveals, via fluorescence resonance energy transfer, pre-identified co-dominant alleles for both SNP and InDel variants between parents and progeny (Semagn et al. 2014) and has a scalability that makes it suitable for a wide range of experimental designs with widely different target loci and sample numbers (Semagn et al. 2014; Steele et al. 2018). Additionally, the KASP marker system has the major advantage of improved cost-effectiveness because it is cheaper than other SNP marker technologies (Yuan et al. 2014). The KASP marker system has already been used for genetic and genotyping analysis in pigeonpea (Saxena et al. 2012), chickpea (Hiremath et al. 2012), and *indica* rice (Pariasca-Tanaka et al. 2015; Steele et al. 2018).

In this study, data obtained from resequencing 13 Korean *japonica* rice varieties were analyzed to detect SNPs among the varieties. In addition, we designed 506 KASP markers based on the detected SNPs, and these markers were screened by genotyping the 13 Korean *japonica* varieties which were sequenced. Furthermore, a genetic map was successfully constructed from an F$_2$ population derived from a cross between two Korean *japonica* rice varieties, Junam and Nampyeong.

**MATERIALS AND METHODS**

**Plant materials and DNA extraction**

Thirteen Korean *japonica* rice varieties, Samgwang, Saenuri, Odae, Nampyeong, Junam, Ilpum, Hwayeong,
Hwacheong, Hiami, Joun, Dongjin, Giho, and Dongan, were grown for genomic DNA extraction in a greenhouse of the National Institute of Agricultural Sciences (NIAS) of the Rural Development Administration (RDA, Jeonju, Korea) (Table 1). Genomic DNA for resequencing was extracted from the leaves of one representative seedling per variety using the DNeasy Plant Maxi Kit (QIAGEN, Hilden, Germany). For genotyping with KASP markers, genomic DNA from the leaves of each of the 13 Korean japonica rice varieties and 188 F₂ progenies from a cross between Junam and Nampyeong was extracted using the Biomedic Plant gDNA Extraction Kit (Biomedic, Bucheon, Korea).

**Table 1. Summary of the sequencing data.**

<table>
<thead>
<tr>
<th>Variety</th>
<th>Raw sequencing data</th>
<th>After quality trimming (Q20)</th>
<th>After read mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no. of reads</td>
<td>Nucleotides (bp)</td>
<td>no. of reads</td>
</tr>
<tr>
<td>Samgwang</td>
<td>311,859,288</td>
<td>31,497,788,088</td>
<td>249,953,570</td>
</tr>
<tr>
<td>Saenuri</td>
<td>283,067,568</td>
<td>28,589,824,368</td>
<td>220,683,072</td>
</tr>
<tr>
<td>Odae</td>
<td>144,090,148</td>
<td>14,553,104,948</td>
<td>125,949,833</td>
</tr>
<tr>
<td>Nampyeong</td>
<td>426,572,058</td>
<td>43,083,777,858</td>
<td>370,334,608</td>
</tr>
<tr>
<td>Junam</td>
<td>164,391,782</td>
<td>16,603,569,982</td>
<td>139,207,477</td>
</tr>
<tr>
<td>Joun</td>
<td>147,885,312</td>
<td>14,936,416,512</td>
<td>129,274,464</td>
</tr>
<tr>
<td>Ilpum</td>
<td>225,038,108</td>
<td>22,728,848,908</td>
<td>195,013,534</td>
</tr>
<tr>
<td>Hwayeong</td>
<td>216,569,822</td>
<td>21,873,552,022</td>
<td>183,405,738</td>
</tr>
<tr>
<td>Hwacheong</td>
<td>251,447,722</td>
<td>25,396,219,922</td>
<td>214,183,790</td>
</tr>
<tr>
<td>Hwacheong</td>
<td>147,966,550</td>
<td>14,944,621,550</td>
<td>131,235,210</td>
</tr>
<tr>
<td>Hwacheong</td>
<td>226,594,766</td>
<td>22,886,071,366</td>
<td>191,911,724</td>
</tr>
<tr>
<td>Giho</td>
<td>204,586,220</td>
<td>20,663,208,220</td>
<td>177,626,551</td>
</tr>
<tr>
<td>Dongjin</td>
<td>71,034,132</td>
<td>7,174,447,332</td>
<td>60,068,249</td>
</tr>
</tbody>
</table>
SNP list was extracted using the clc_find_variations command, and SNPs with mapping depth greater than seven were selected to exclude mismatches due to sequencing errors. Based on the SNPs detected between each variety and Nipponbare, SNPs among the 13 varieties were analyzed using a Python program developed in-house. Additionally, polymorphism information content (PIC) values of SNPs were calculated according to a simplified version of the formula proposed by the Anderson et al. (1993):

\[ PIC_i = 1 - \sum_{j=1}^{n} P_{ij}^2 \]

where \( P_{ij} \) is the frequency of the jth allele for the ith marker and summed over \( n \) alleles (Ni et al. 2002).

SNPs were annotated as genic and intergenic based on positional information obtained from the reference genome. SNPs in genic regions were classified as coding sequence (CDS), untranslated regions, non-coding exons, and introns. SNPs in CDS were separated into synonymous SNPs and non-synonymous SNPs on the basis of amino-acid substitutions.

Development of KASP markers

For the development of potential KASP markers based on the SNPs detected among the 13 Korean japonica rice varieties, SNP sites with PIC values > 0.4 per 200-kb interval were selected on each chromosome. In total, 1,014 SNP sites were selected, and 506 SNP sites were used for developing KASP marker. The SNP sequence, 100-bp left-flanking sequence and 100-bp right-flanking sequence of each SNP site were used to design two allele-specific forward primers and a common reverse primer. The design and manufacture of the KASP marker primers were performed by LGC genomics (London, UK). The markers are listed in Supplementary Table S1 along with their chromosomal location and primer sequences.

The PCR profile and composition of KASP reactions are described at https://www.lgcgroup.com/services/genotyping/projects/#.W72qRWZReUk. KASP amplifications and allelic discriminations were performed using the Nexar system (LGC Douglas Scientific, Alexandria, USA) in the Seed Industry Promotion Center (Gimje, Korea) of Foundation of Agri. Tech. Commercialization & Transfer in Korea. An aliquot (0.8 \( \mu \)L) of 2× Master Mix (LGC Genomics, London, UK), 0.02 \( \mu \)L of 72× KASP assay mix (LGC Genomics, London, UK), and 5 ng genomic DNA template were mixed in 1.6 \( \mu \)L KASP reaction mixture in a 384-well Array Tape. The reactions were run in duplicates and with non-template controls in each run. KASP amplification was performed using the following thermal cycling profile, starting with 15 minutes at 94°C, a touchdown phase of 10 cycles at 94°C for 20 seconds, and at 61°C-55°C (dropping 0.6°C per cycle) or 68°C-62°C (dropping 0.6°C per cycle) for 60 seconds, followed by 26 cycles at 94°C for 20 seconds and 55°C or 62°C for 60 seconds (first PCR stage). Next, recycling was performed at three cycles of 94°C for 20 seconds and 57°C for 60 seconds (second PCR stage). The recycling was performed two times, and the fluorescence read was taken for KASP genotyping after PCR.

Construction of genetic map and phylogenetic tree

Total 188 F\(_2\) plants from the cross between Junam and Nampyeong were genotyped using 235 KASP markers showing polymorphism between Junam and Nampyeong. Based on the genotype data, a genetic map was constructed using the MapDisto 1.7 program (Lorieux 2012). The genetic map was processed for graphical presentation using the MapChart version 2.32 software (Voorrips 2002). Phylogenetic analysis with the 13 Korean japonica varieties was performed based on the result of genotyping with 81 KASP markers by the Unweighted Pair Group Method with Arithmetic Mean method, using POPTREE2 software (Takezaki et al. 2010).

RESULTS

Discovery of SNPs among Korean japonica rice varieties

The amount of raw genome sequence data from different varieties ranged from 7.1 Gbp (Dongan) to 43.1 Gbp (Nampyeong) (Table 1). After quality trimming, the remaining sequence data ranged from 5.8 Gbp (Dongan) to

...
36.0 Gbp (Nampyeong) with sequencing depth of 15.5-96.6×. After read mapping on the Nipponbare reference genome, the mapped nucleotides ranged from 4.7 Gbp (Dongan) to 29.5 Gbp (Nampyeong), with a mapping depth of 12.7-78.9×.

In the comparisons between each variety and Nipponbare, Hwacheong showed the lowest total SNP number of 95,132, with an SNP density of 0.25 per kbp, whereas Nampyeong showed the highest total SNP number of 253,258 and an SNP density of 0.68 per kbp (Table 2). The SNP distribution patterns on the individual rice chromosomes differed among the varieties. For example, the chromosome harboring the highest SNP number was chromosome 8 in Samgwang, Junam, and Hwayeong, chromosome 11 in Saenuri and Nampyeong, and chromosome 12 in Joun, Giho, Dongjin, and Dongan.

We identified 740,566 SNPs among the 13 Korean japonica rice varieties. Overall, chromosome 5 had the lowest SNP number (16,017 SNPs, 2.2%), and chromosome 11 had the highest SNP number (163,557 SNPs, 22.1%) (Fig. 1). The distribution of SNPs per 100-kbp interval over the 12 rice chromosomes is shown in Fig. 2. Most of the intervals had SNPs; however, the density was uneven within each chromosome. This indicates that SNP marker development among the Korean japonica rice varieties would be possible in almost all intervals. The distribution of PIC values of the detected SNPs is shown in Fig 3. The PIC values varied from 0.142 to 0.5, where the maximum possible PIC value is 0.5. SNPs with high PIC values were located across almost the entire genome, except for a few regions. The exceptionally low-PIC regions were 12-17

---

**Table 2. The number of SNPs on individual chromosomes detected between each variety and Nipponbare.**

<table>
<thead>
<tr>
<th>Variety</th>
<th>Chr1</th>
<th>Chr2</th>
<th>Chr3</th>
<th>Chr4</th>
<th>Chr5</th>
<th>Chr6</th>
<th>Chr7</th>
<th>Chr8</th>
<th>Chr9</th>
<th>Chr10</th>
<th>Chr11</th>
<th>Chr12</th>
<th>Total</th>
<th>SNP density</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samgwang</td>
<td>3,461</td>
<td>7,067</td>
<td>3,527</td>
<td>8,383</td>
<td>1,788</td>
<td>2,553</td>
<td>1,985</td>
<td>84,389</td>
<td>6,898</td>
<td>5,666</td>
<td>14,963</td>
<td>17,776</td>
<td>36.0 Gbp</td>
<td>0.48</td>
</tr>
<tr>
<td>Saenuri</td>
<td>19,209</td>
<td>6,931</td>
<td>2,854</td>
<td>12,469</td>
<td>2,016</td>
<td>6,107</td>
<td>14,752</td>
<td>5,923</td>
<td>1,934</td>
<td>26,094</td>
<td>102,846</td>
<td>8,739</td>
<td>95,132</td>
<td>0.25</td>
</tr>
<tr>
<td>Odae</td>
<td>9,418</td>
<td>5,640</td>
<td>9,119</td>
<td>10,373</td>
<td>3,190</td>
<td>5,499</td>
<td>25,649</td>
<td>17,699</td>
<td>5,656</td>
<td>17,035</td>
<td>35,628</td>
<td>8,165</td>
<td>63,557</td>
<td>0.32</td>
</tr>
<tr>
<td>Nampyeong</td>
<td>22,730</td>
<td>4,022</td>
<td>8,649</td>
<td>10,373</td>
<td>3,190</td>
<td>5,499</td>
<td>25,649</td>
<td>17,699</td>
<td>5,656</td>
<td>17,035</td>
<td>35,628</td>
<td>8,165</td>
<td>63,557</td>
<td>0.32</td>
</tr>
<tr>
<td>Junam</td>
<td>9,293</td>
<td>8,039</td>
<td>7,602</td>
<td>9,917</td>
<td>3,190</td>
<td>5,499</td>
<td>25,649</td>
<td>17,699</td>
<td>5,656</td>
<td>17,035</td>
<td>35,628</td>
<td>8,165</td>
<td>63,557</td>
<td>0.32</td>
</tr>
<tr>
<td>Ilpum</td>
<td>11,502</td>
<td>7,416</td>
<td>8,563</td>
<td>10,466</td>
<td>3,190</td>
<td>5,499</td>
<td>25,649</td>
<td>17,699</td>
<td>5,656</td>
<td>17,035</td>
<td>35,628</td>
<td>8,165</td>
<td>63,557</td>
<td>0.32</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** SNP distribution on each chromosome among 13 Korean japonica rice varieties.
Fig. 2. Distribution of SNPs among 13 Korean *japonica* rice varieties per 100 kb on the 12 rice chromosomes. The x-axis represents the physical distance along each chromosome. The y-axis indicates the common logarithm of the number of SNPs.
Fig. 3. Distribution of PIC values of SNPs on the 12 rice chromosomes among 13 Korean *japonica* rice varieties. The x-axis represents the physical distance along each chromosome. The y-axis indicates the PIC values of SNPs. Red bars indicate the regions of low PIC values.
Mbp on chromosome 2, 21-26 Mbp on chromosome 8, and 4-12 Mbp on chromosome 10 (Fig. 3). Except for these regions, it would be possible to develop SNP markers with high PIC values across the entire genome.

As a result of annotation of the 740,566 SNPs detected, 572,691 SNPs and 167,875 SNPs were located in the intergenic (77.3%) and genic (22.7%) regions, respectively (Fig. 4A). Of the genic region SNPs, 86,863, 48,044, 20,714, 10,314, and 1,940 SNPs were detected from introns (51.7%), coding sequences (28.6%), 3’ untranslated region (UTR) (12.3%), 5’UTR (6.1%), and non-coding exon (1.2%), respectively (Fig. 4B). The 48,044 SNPs located in CDS were classified into 21,118 synonymous SNPs (44%) and 26,926 non-synonymous SNPs (56%) (Fig. 4C).

**Development and validation of KASP markers**

Among the 740,566 SNP sites detected, 1,014 SNP sites were chosen for KASP marker development by selecting SNP sites with PIC values > 0.4 per 200 kbp. Subsequently, 506 KASP markers were designed, representing approximately 50% of the selected high-PIC SNP sites. The list of KASP markers is shown in Supplementary Table S1, and includes primer sequences and chromosomal locations. These 506 KASP markers were tested for genotyping with the 13 Korean *japonica* rice varieties used in the above-mentioned SNP detection studies. Of the 506 KASP markers, amplification of 17 markers (3.4%) failed in all the Korean *japonica* varieties tested. Among the remaining 489 KASP markers, 89 markers (17.5%) showed monomorphism and/or heterozygous type and were deemed unusable as markers. The remaining 400 KASP markers (79.1%) showed polymorphisms among the tested varieties and were deemed usable (Table 3, Supplementary Table S2). Among the 400 polymorphic

---

### Table 3. Summary of comparison of SNP and KASP marker genotypes among 13 Korean *japonica* rice varieties.

<table>
<thead>
<tr>
<th></th>
<th>Usable</th>
<th>Unusable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>81</td>
<td>62</td>
</tr>
<tr>
<td>Partial</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>Partial miss-matched</td>
<td>308</td>
<td>17</td>
</tr>
<tr>
<td>Mono</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*KASP genotypes corresponded with SNP genotypes completely in all varieties.*

*KASP genotypes corresponded with SNP genotypes in amplified varieties.*

*KASP genotypes corresponded with SNP genotypes partially in amplified varieties.*

*KASP genotypes showed monomorphism among all varieties.*

*KASP genotypes showed monomorphism and/or heterozygous type in all varieties.*

*No amplification in all varieties.*
KASP markers, 81 marker genotypes completely corresponded with the SNP genotypes, whereas further 11 marker genotypes largely corresponded with SNP genotypes, with a small amount of missing data, and remaining 308 marker genotypes corresponded with SNP genotypes at least partially in the amplified varieties with a few mismatches (Table 3, Supplementary Table S2).

Of the 400 polymorphic KASP markers, 303 (75.7%) were distributed in the intergenic regions, whereas the remaining 97 KASP markers (24.3%) were located in the...
Phylogenetic analysis among 13 Korean japonica rice varieties

Of the 400 usable KASP markers developed in this study, 81 KASP marker genotypes corresponded completely with the SNP genotypes in 13 Korean japonica varieties (Table 2, Supplementary Table S2), and the genotype information from these markers was used for the construction of a phylogenetic tree of the 13 Korean japonica rice varieties (Fig. 6). As a result, the varieties were divided into three groups: Hwayeong, Junam, Dongjin, and Samgwang were included in group A, Saenuri, Nampyeong, Dongan, and Hwacheong in group B, and Ilpum, Giho, Joun, Odae, and Hiami in group C. This indicates that the KASP markers developed in this study can be utilized in revealing the phylogenetic relationships among Korean japonica rice varieties.

DISCUSSION

To detect genetic variation and to use this detected variation in gene mapping and breeding within a genetically closely-related group, such as the Korean japonica rice varieties group tested here, large-scale discovery of SNPs among the varieties and the development of markers suitable for massive and cost-effective genotyping assays are required.

With respect to the SNP discovery among the Korean japonica rice varieties, the number of SNPs differed widely among the individual chromosomes. It has already been reported that rice chromosomes 6, 8, 11, and 12 have large numbers of DNA polymorphisms, whereas chromosomes 3, 5, and 9 have small numbers of polymorphisms. Relative to Nipponbare, rice chromosomes 6, 8, 11, and 12 have large numbers of DNA polymorphisms, whereas chromosomes 3, 5, and 9 have small numbers of polymorphisms. It has already been reported that rice chromosomes 6, 8, 11, and 12 have large numbers of DNA polymorphisms, whereas chromosomes 3, 5, and 9 have small numbers of polymorphisms. Relative to Nipponbare, rice chromosomes 6, 8, 11, and 12 have large numbers of DNA polymorphisms (Jeong et al. 2013; Jeong et al. 2015), which were included in the current study. When genetic variation was investigated by assessing SNPs on individual chromosomes in the Japanese japonica varieties Yamadanishiki, Gohyakumangoku, Kameji, Omachi, Koshihikari, and Norin 8, chromosomes 4 and 11 showed high densities of SNPs, whereas chromosome 9 has low densities of variants in all varieties, except Gohyakumangoku (Arai-Kichise et al. 2014). The results of SNP detection in each chromosome among the 13 Korean japonica varieties in the current study (Fig. 1) was similar to the previously reported findings. Chromosomes 6, 8, 11, and 12 had high number of SNPs, whereas chromosomes 2, 3, 5, and 9 had low number of SNPs. From this, it could be inferred that the genes on chromosome 6, 8, 11, and 12 appear to be highly diversified within the Korean japonica rice varieties.
Conversion of the detected polymorphisms to molecular markers is essential to establish the applicability of such polymorphisms to breeding. Several SNP-genotyping platforms, such as TaqMan and KASP markers, are available to convert SNPs into genotyping markers. As KASP markers have been reported to be appropriate and cost-effective for large-scale assays (Yuan et al. 2014), we designed KASP markers for the 506 SNP sites that had been detected among the 13 Korean japonica rice genomes (Supplementary Table S1). Several reports have been published on the development of KASP assays in other plants. In legumes, Saxena et al. (2012) developed 1,616 KASP markers and they were used to screen 24 pigeonpea genotypes representing the parents of 14 mapping populations. Moreover, conversion of the SNPs to cost-effective and high-throughput KASP markers generated 2,144 candidate SNPs that were selected on the basis of their genomic distribution and the origin of the species Oryza glaberrima (African rice) and O. sativa (indica rice), 2,015 SNPs were successfully converted into KASP markers and subsequently validated (Pariasca-Tanaka et al. 2015). Of these 2,015 KASP markers (Pariasca-Tanaka et al. 2015), only 1,890 KASP markers were found to be applicable to indica rice, and an additional 39 novel KASP markers were developed in nine indica rice lines (Steele et al. 2018). In the current study, although attempts were made to convert 506 SNP sites into KASP markers, only 400 SNP sites (79.1%) could be successfully converted to polymorphic KASP markers (Table 2, Supplementary Table S2). The unusable KASP markers (20.9%), showing either no amplification or monomorphism (Table 2, Supplementary Table S2), could have resulted from the presence of duplicate or paralogous loci, incorrect primer design near the SNP, identification of false SNPs initially, and/or the need to optimize PCR conditions (Saxena et al. 2012). This rate of conversion from selected SNPs to functional KASP markers could probably be increased further by optimizing primer design and amplification conditions (Saxena et al. 2012; Islam et al. 2015). Furthermore, of the 400 polymorphic markers, 81 KASP genotypes completely corresponded with SNP genotypes, but mismatched genotypes were found mostly from a limited number of varieties (Supplementary Table S2). More experiments, therefore, will be needed to resolve these mismatched genotypes between KASP markers and SNP analysis by resequencing.

When the genetic relationships among Korean japonica rice varieties were analyzed using 81 of the KASP markers developed (Table 2), short genetic distance was revealed among the 13 Korean japonica varieties; however, they could still be separated into three groups, of which four varieties, Hwayeong, Junam, Samgwang, and Dongin, were clustered into one group, A (Fig. 5). Grouping patterns similar to those from our results among Korean japonica varieties were obtained in other studies, using SSR and AFLP markers (Ji et al. 1998; Xiao et al. 2012). In Korean japonica rice breeding, Hwayeong has been reported to be the parent of the Junam and Samgwang varieties (Kwon et al. 2017), and high levels of genetic similarity between varieties sharing a parent have been reported among Korean japonica varieties (Ji et al. 1998). Based on these reports and our results, it is possible that Korean japonica rice varieties were bred from genetically similar parents and consequently, genetic diversity has been narrowed. Nevertheless, the results of our phylogenetic analysis suggested that the 81 KASP markers may be sufficient for genetic relationship analysis among even closely-related Korean japonica rice varieties.

As the genetic diversity has been narrowed among Korean japonica rice varieties, more markers are required than the KASP markers which have been established so far. Because SNPs were shown to be present at a high frequency among 13 Korean japonica rice varieties, we can readily develop additional KASP markers. Further, the remaining 508 SNP sites of the originally selected 1,014 SNP sites among the Korean japonica rice varieties will be converted to KASP markers. In addition, we are planning to focus more on KASP markers based on the SNP sites located in the genic region.

Taken all together, the results obtained in this current study showed that large-scale SNP identification through
resequencing has enabled the development of numerous molecular markers within the Korean japonica rice varieties. The KASP makers developed here will be helpful in conducting mapping studies and in identifying useful genes present in Korean japonica rice varieties and will ultimately accelerate the molecular breeding of rice in Korea.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Next-Generation BioGreen 21 Program (Plant Molecular Breeding Center No PJ01323401), Rural Development Administration, Republic of Korea.

REFERENCES


