Identification of Korean Ginseng (\textit{Panax ginseng}) Cultivars Using Simple Sequence Repeat Markers

Yurry Um\textsuperscript{1}, Mei-Lan Jin\textsuperscript{1}, Ok-Tae Kim\textsuperscript{1}, Young-Chang Kim\textsuperscript{1}, Seong-Cheol Kim\textsuperscript{1}, Seon-Woo Cha\textsuperscript{1}, Ki-Wha Chung\textsuperscript{2}, Serim Kim\textsuperscript{3}, Chan-Moon Chung\textsuperscript{3}, Yi Lee\textsuperscript{3}\textsuperscript{*}

\textsuperscript{1}Department of Herbal Crop Research, National Institute of Horticultural and Herbal Science, Rural Development Administration, Eumseong 27709, Korea
\textsuperscript{2}Department of Biological Science, Kongju National University, Gongju 32588, Korea
\textsuperscript{3}Department of Industrial Plant Science and Technology, Chungbuk National University, Cheongju 28644, Korea

\textbf{ABSTRACT} \textit{Panax ginseng} has been one of the most important herbal medicines used in Eastern Asia. Recently, various molecular markers have been developed to authenticate \textit{Panax} species, but these markers cannot differentiate the exact varieties or variants of Korean ginseng cultivars. In this study, six cultivars of Korean ginseng (Chunpoong, Yunpoong, Gopoong, Gumpoong, Jakyung, and Hwangsook), \textit{P. quinquefolius}, and \textit{P. notoginseng} were differentiated by simple sequence repeat (SSR) marker development. Specific primer sets were designed for the 54 candidate sequences containing SSRs that were predicted. Finally, eight polymorphic SSR loci were developed. DNA fragment analysis was performed using fluorescence-labelled primers for the amplicons. Reproducibility tests were carried out using multiple samples of Korean ginseng cultivars and \textit{Panax} species. Eight primer sets (PgSSR07, PgSSR08, PgSSR09, PgSSR17, PgSSR37, PgSSR40, PgSSR51, and PgSSR53) showing polymorphism were used for phylogenetic relationship analysis. Consequently, six Korean ginseng cultivars (Chunpoong, Yunpoong, Gopoong, Gumpoong, Jakyung, and Hwangsook), \textit{P. quinquefolius}, and \textit{P. notoginseng} could be identified using the combination of SSR markers discovered.

\textbf{Keywords} \textit{Panax}, Korean ginseng cultivars, SSR markers, DNA fragment analysis, Phylogenetic relationship

\textbf{INTRODUCTION}

Korean ginseng (\textit{Panax ginseng}, C.A. Meyer) is an herbaceous perennial plant of the genus \textit{Panax} (Araliaceae). Korean ginseng is a plant that grows only in the Far East areas, as in Asia, China, Manchuria, and Russia (Woo \textit{et al.} 2004). The main physiologically active components of ginseng are saponins, phenols, polyacetylenes, alkaloids, and polysaccharides (Lee \textit{et al.} 2008). These components act on the central nervous, cardiovascular, endocrine, and immune systems, and are recently reported to have antioxidant and anti-stress effects (Nam \textit{et al.} 2005).

In general, the \textit{Panax} genus is divided into nine species: \textit{P. ginseng}, \textit{P. notoginseng}, \textit{P. quinquefolius}, \textit{P. japonicus}, \textit{P. pseudoginseng}, \textit{P. assamicus}, \textit{P. sokpayensis}, \textit{P. sikkimensis}, and \textit{P. bipinnatifidus} (Sharma and Pandit 2009). However, the number of \textit{Panax} species is unclear because several scholars define the same species differently. Many studies identify plants species and varieties using morphological, cytological, anatomical, physiological, and ecological methods (Woo \textit{et al.} 2004; Yu \textit{et al.} 2009; Jee \textit{et al.} 2014; Bai \textit{et al.} 2015; Kim \textit{et al.} 2015). Using such morphological characteristics to distinguish among varieties can lead to confusion because expression may be the result of a combination of genetics and environment (Jo \textit{et al.} 2013). Furthermore, molecular techniques have added another level of classification through study of DNA (Lim and Choi 1990; Lim \textit{et al.} 1993; Hon \textit{et al.} 2003). In order to identify the plant species, restriction fragment length polymorphism (RFLP), random amplified polymorphic DNAs (RAPD), amplified fragment length polymorphism (AFLP), and simple sequence repeat (SSR) were used...
Table 1. Source of the Panax species and Panax ginseng cultivars used in this study.

<table>
<thead>
<tr>
<th>Species</th>
<th>Cultivar</th>
<th>Accession</th>
<th>No. of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. ginseng</td>
<td>Chunpoong</td>
<td>KT&amp;Gz</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Yunpoong</td>
<td>KT&amp;G</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Gopoong</td>
<td>Local</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Gumpoong</td>
<td>Local</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Jakyungjong</td>
<td>KT&amp;G</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Hwangsookjong</td>
<td>Local</td>
<td>20</td>
</tr>
<tr>
<td>P. quinquefolius</td>
<td>NIHHS</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>P. notoginseng</td>
<td>NIHHS</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

aKT&G: name of a Korean ginseng corporation, NIHHS: National Institute of Horticultural and Herbal Science.

Besides, many researchers have tried to identify ginseng species and varieties using SSR markers (Kim et al. 2007; Ma et al. 2007; Nguyen et al. 2010; Bang et al. 2011a, 2011b). SSR markers are important for several types of research, including the assessment of genetic diversity, development of genetic maps, comparative genomics, marker-assisted selection, and other plant breeding fields. However, there are few studies on genetic polymorphism analyses of domestic cultivars using SSR markers. These studies were inconclusive because the ginseng genome is very large (3.12 billion base pairs) and there are many recently developed varieties. Therefore, the development of more markers is necessary for molecular breeding and variety identification.

The aim of this research was to develop SSR markers from Korean ginseng cultivars and differentiate the Panax species and Panax ginseng cultivars. The SSR markers were identified from the database which was constructed using an enriched SSR library and the polymorphism of the markers was investigated using DNA fragment analysis.

MATERIALS AND METHODS

Plant materials and DNA extraction

This study was carried out with 115 Panax accessions grown in Korea. Panax samples used in this study were Korean ginseng cultivars (17 Chunpoong, 22 Yunpoong, 15 Gopoong, 10 Gumpoong, 23 Jakyung, and 20 Hwangsook) collected from various local fields in Goesan and Suwon, Korea. Moreover, 5 P. quinquefolius and 3 P. notoginseng were obtained from KT&G cooperation company (Daejeon, Korea) and National Institute of Horticultural and Herbal Science (Eumseong, Korea), respectively (Table 1). One hundred mg of fresh leaves were used for genomic DNA extraction using TissueLyser (Qiagen, Düsseldorf, Germany) and the DNeasy plant DNA isolation kit (Qiagen GmbH, Hilden, Germany).

SSR detection and specific primer design

A total of 3,773 DNA sequences of P. ginseng were collected through downloading of 2,549 genome survey sequences from the National Center for Biotechnology Information (NCBI) database and 1,224 microsatellite-enriched library (MEL) sequences from the data collected by Kim et al. (2007). Putative Panax MEL sequences were analyzed by the SSR identification tool (http://www.gramene.org/db/markers/ssrtool) program. Primer pairs flanking the SSRs were designed using the Primer3 program (Koressaar and Remm, 2007; Untergrasser et al. 2012) with a length of 18 to 25 bp, amplification product size of 100 to 500 bp, annealing temperature (Tm) ranging from 54°C to 60°C, and GC content between 40% and 60%.

Polymerase chain reaction (PCR) amplification and data analysis

Twenty μl of PCR mixture contained 10 pmole primers (fluorescence-labeled primer), 10 ng of genomic DNA, 0.1 mM dNTP, 0.2 units Taq DNA polymerase (RBC
Bioscience, New Taipei City, Taiwan) and 10× Taq polymerase buffer using a T-gradient Thermoblock PCR system (Biometra, Göttingen, Germany) for 35 cycles. The initial cycle was pre-denaturation for 5 minutes at 94°C, denaturation for 30 seconds at 94°C, annealing for 1 minute at 59°C, and extension for 1 minute at 72°C (for the final cycle, 5 minutes added). Electrophoresis was performed using a 10% polyacrylamide gel. Polymorphic bands were visualized by gel staining using Gelstar (Lonza Inc., Rockland, ME, USA) and photographed under UV Transilluminator (CoreBio System Co., Seoul, Korea). Polymorphic amplicons visualized on polyacrylamide gels were separated in agarose gel electrophoresis. To verify the accuracy of ginseng SSR polymorphisms, the forward primer was labeled and tested on 6 Korean ginseng cultivars, *P. quinquefolius* and *P. notoginseng*. DNA fragment profile data were demonstrated by the SSR markers PgSSR08 and PgSSR51 using NED-labelled primers; PgSSR09 and PgSSR37 using FAM-labelled primers; PgSSR07 and PgSSR40 using VIC-labelled primers; and PgSSR17 and PgSSR53 using PET-labelled primers. Then, we preformed capillary electrophoresis using an ABI3730xl DNA analyzer with GeneScan-500LIZ size standard (Applied Biosystems, Foster City, CA, USA). Based on the resulting data, allele sizes were measured in terms of the numbers of alleles, gene diversity, and polymorphic information content (PIC), using PowerMarker software ver. 3.23 (Liu and Muse 2005). Coefficients of genetic similarity for the eight species used in this study were calculated using the SIMQUAL program (NTSYS-pc, ver. 2.10; Mogea 1999). A neighbor-joining dendrogram was constructed based on a genetic similarity matrix created in the SHAN clustering program from NTSYSpc using the unweighted pair group method with arithmetic mean algorithm.

**RESULTS**

**Abundance and distribution of SSRs**

A total of 3,773 sequences were collected and used for the finding of SSRs, defined as di- to octa-nucleotide motifs. Thus, 1,283 sequences (34% of the total sequences) contained SSRs and a total of 13,769 *in silico* SSR markers were identified. The observed frequencies of di-, tri-, tetra-, penta-, hexa-, hepta-, and octa-nucleotide repeats were 60.1% (n=8,279), 27.8% (n=3,832), 7.8% (n=1,070), 2.7% (n=371), 1.1% (n=153), 0.3% (n=48), and 0.1% (n=16), respectively (Table 2).

Among the dinucleotide repeats, (AT/AT)n and (TA/TA)n were the most abundant repeat motifs and represented 20.0% and 17.2% of dinucleotides, respectively. Among the trinucleotide repeats, (ATA/TAT)n and (TAA/TTA)n were most abundant, with each accounting

<table>
<thead>
<tr>
<th>Repeat type</th>
<th>Simple sequence repeat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Di-nucleotide</td>
<td>8,279 (60.1)</td>
</tr>
<tr>
<td>Tri-nucleotide</td>
<td>3,832 (27.8)</td>
</tr>
<tr>
<td>Tetra-nucleotide</td>
<td>1,070 (7.8)</td>
</tr>
<tr>
<td>Penta-nucleotide</td>
<td>371 (2.7)</td>
</tr>
<tr>
<td>Hexa-nucleotide</td>
<td>153 (1.1)</td>
</tr>
<tr>
<td>Hepta-nucleotide</td>
<td>48 (0.3)</td>
</tr>
<tr>
<td>Octa-nucleotide</td>
<td>16 (0.1)</td>
</tr>
<tr>
<td>Total</td>
<td>13,769 (100.0)</td>
</tr>
</tbody>
</table>

Values are presented as number (%).

**Table 3.** Abundant SSR motives in the *Panax ginseng* database. Motifs accounting for 4.9% or more of each repeat type are included*.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SSR motif</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC/GT</td>
<td>990</td>
<td>(12.0)</td>
</tr>
<tr>
<td>AG/CT</td>
<td>1,240</td>
<td>(15.0)</td>
</tr>
<tr>
<td>AT/AT</td>
<td>1,655</td>
<td>(20.0)</td>
</tr>
<tr>
<td>CA/TG</td>
<td>1,403</td>
<td>(16.9)</td>
</tr>
<tr>
<td>GA/TC</td>
<td>1,340</td>
<td>(16.2)</td>
</tr>
<tr>
<td>TA/TA</td>
<td>1,428</td>
<td>(17.2)</td>
</tr>
<tr>
<td>AGA/TCT</td>
<td>192</td>
<td>(5.0)</td>
</tr>
<tr>
<td>ATA/TAT</td>
<td>419</td>
<td>(10.9)</td>
</tr>
<tr>
<td>TTG/CAA</td>
<td>215</td>
<td>(5.6)</td>
</tr>
<tr>
<td>GAA/TTC</td>
<td>220</td>
<td>(5.7)</td>
</tr>
<tr>
<td>TAA/TTA</td>
<td>377</td>
<td>(9.8)</td>
</tr>
<tr>
<td>TCA/TGA</td>
<td>189</td>
<td>(4.9)</td>
</tr>
<tr>
<td>AAAT/ATT</td>
<td>65</td>
<td>(6.0)</td>
</tr>
<tr>
<td>TAAA/TTTA</td>
<td>62</td>
<td>(5.7)</td>
</tr>
</tbody>
</table>

Values are presented as number (%).

*SSR: simple sequence repeat.
for approximately 10% of the trinucleotides. Among the
tetranucleotide repeats, (AAAT/ATTT)n and (TAAA/
TTTA)n were most abundant, with each accounting for
approximately 6% of the tetranucleotides (Table 3).

**Gel electrophoresis of SSR motifs and DNA fragment analysis**

Fifty-four putative SSR markers were validated for
polymorphic markers. The allelic structures of the markers
were determined from DNA band patterns from Gelstar
stained 10% polyacrylamide gels (data not shown). Finally,
eight markers—PgSSR07, PgSSR08, PgSSR09, PgSSR17,
PgSSR37, PgSSR40, PgSSR51, and PgSSR53—amplified
polymorphic PCR product bands among Korean ginseng
cultivars. The products were analyzed using the DNA
fragment analysis method.

A total of eight SSR loci showed polymorphism, pro-
ducing a total of 33 alleles among the eight Panax species.
The observed number of alleles ranged from three
(PgSSR07, PgSSR09, PgSSR17, PgSSR40, and PgSSR53)
to eight (PgSSR51), with an average of 4.13 alleles per
locus. The genetic diversity of each locus ranged from 0.13
(PgSSR53) to 0.78 (PgSSR51), with an average of 0.45. In
addition, the heterozygosity values ranged from 0.07
(PgSSR53) to 0.59 (PgSSR08 and PgSSR09), with an
average of 0.32. PIC values were calculated for each
polymorphic marker using the method of Roldán-Ruiz et
al. (2000). For all accessions, PIC values ranged from 0.13
(PgSSR53) to 0.76 (PgSSR51), with an average of 0.42
(Table 4).

The NTSYSpc software calculated the genetic distances
among ginseng cultivars. The eight cultivars were divided
into three groups: I, II, and III (Fig. 1). The six Korean
ginseng cultivars were divided into two groups. Group I
contained Chunpoong, Yunpoong, Gopoong, and Jakyung.
Group II contained two cultivars, Gumpoong and
Hwangsook. Group III contained *P. quinquefolius* and *P.
notoginseng*.

**DISCUSSION**

*P. ginseng* is one of the most important medicinal plants in
Korea. Several cultivars that have different characteristics
have been developed. For instance, Yunpoong is a cultivar
that produces a higher yield than other cultivars and
Chunpoong has good root shape and is suitable for red
ginseng production (Lee 2002). Thus, molecular markers

---

**Table 4.** Characteristics of 8 SSR markers including GenBank accession, repeat motif, primer sequence, allele size range, number of alleles, genetic diversity, heterozygosity, and PIC among 8 cultivars of *Panax* species*.

<table>
<thead>
<tr>
<th>Marker ID</th>
<th>GenBank accession</th>
<th>Repeat motif</th>
<th>Forward primer (5'→3')</th>
<th>Allele size range (bp)</th>
<th>No. of allele</th>
<th>Genetic diversity (h)</th>
<th>Heterozygosity</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PgSSR07</td>
<td>GU565701</td>
<td>(GAA)$_3$</td>
<td>VIC-ATGGAAGTGGTTTGTTGG</td>
<td>AGGAGACCATGAAGGATTCG</td>
<td>204-276</td>
<td>3</td>
<td>0.34</td>
<td>0.26</td>
</tr>
<tr>
<td>PgSSR08</td>
<td>EF531909</td>
<td>(GAAA)$_2$</td>
<td>NED-CCTGCTGGAGATTGAAGTCAT</td>
<td>GTGGAATGCTTCAGAGAT</td>
<td>180-207</td>
<td>5</td>
<td>0.61</td>
<td>0.59</td>
</tr>
<tr>
<td>PgSSR09</td>
<td>GU565702</td>
<td>(AG)$_3$</td>
<td>FAM-TGGAATTTGACATTTCTTG</td>
<td>CCCTCACTAACCCTAAAAC</td>
<td>164-204</td>
<td>3</td>
<td>0.51</td>
<td>0.59</td>
</tr>
<tr>
<td>PgSSR17</td>
<td>BZ957342</td>
<td>(GGACC)$_3$</td>
<td>PET-ATCGAAGCGGAGCTGGAC</td>
<td>CTGAATCTCCTCTCTCTCTCC</td>
<td>103-118</td>
<td>3</td>
<td>0.21</td>
<td>0.08</td>
</tr>
<tr>
<td>PgSSR37</td>
<td>EF140892</td>
<td>(ATG)$_3$(CTGATG)$_2$</td>
<td>FAM-AATCGAAAAACAAGAAGCCTAAAC</td>
<td>CTTCTCTCTCTCTCTCTCTCTCCT</td>
<td>100-122</td>
<td>5</td>
<td>0.73</td>
<td>0.36</td>
</tr>
<tr>
<td>PgSSR40</td>
<td>EF140893</td>
<td>(ATAG)$_3$</td>
<td>VIC-GTGAAGTGGTTGAACCTTCTTTGCTAAG</td>
<td>ATTTTAACTCCTCTCTCTCTCTCCTAC</td>
<td>125-162</td>
<td>3</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>PgSSR51</td>
<td>EF140899</td>
<td>(AAG)$_3$</td>
<td>NED-GGGAGTGGATGTGAGGAACTTCC</td>
<td>GGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTTCCC</td>
<td>299-330</td>
<td>8</td>
<td>0.78</td>
<td>0.34</td>
</tr>
<tr>
<td>PgSSR53</td>
<td>EF140900</td>
<td>(CTCCTTT)$_4$</td>
<td>PET-CTACACGCTTCTTTCATGCTTCAA</td>
<td>TGCTGCTA AAAAGAGTGGTGCAGGC</td>
<td>185-194</td>
<td>3</td>
<td>0.13</td>
<td>0.07</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.13</td>
<td>0.45</td>
<td>0.32</td>
<td>0.42</td>
</tr>
</tbody>
</table>

*SSR: simple sequence repeat, PIC: polymorphic information content.
Fig. 1. Dendrograms showing genetic similarities among eight cultivars of ginseng, constructed using the *Panax ginseng* cultivars and *Panax* species similarity coefficient based on eight SSR markers (unweighted pair group method with arithmetic mean algorithm). Cultivars in Chunpoong (C), Yunpoong (Y), Gopoong (GO), Gumpoong (GU), Jakyung (J), Hwangsook (H), *P. quinquefolius* (PQ), *P. notoginseng* (PN). Group I contained C, Y, GO, and J. Group II contained GU and H. Group III contained *P. quinquefolius* and *P. notoginseng*. 
are needed to identify the cultivars for optimum utilization and further breeding. Previous researchers have developed various molecular markers—RAPD (Um et al. 2001; Cui et al. 2003; Shim et al. 2003; Lim et al. 2007), RFLP (Yang and Kim 2003), AFLP (Ha et al. 2002; Choi et al. 2008), STS (Bang et al. 2010), and SSR (Bang et al. 2011a, 2011b)—for identification of ginseng cultivars. But, it was not sufficient to identify all the Korean ginseng cultivars.

In this study, we collected 3,773 *P. ginseng* sequences from GSS and MEL databases. The selected primer sets were labeled with fluorescence at the 5’ end of the forward primer for DNA fragment analysis. The DNA fragment analysis method with different fluorescent colors made it easy to analyze the genetic diversity. Especially for the similar-sized PCR products, labeling using different colors made it possible to analyze several markers at the same time. This method is cost- and time-effective for large-sized samples. We used four colors of fluorescent dyes, FAM, NED, PET, and VIC. The product size of PgSSR09 and PgSSR17 was very similar, but it was possible to identify the markers because we used different colored dyes for each marker. Korean ginseng cultivars have very similar genotypes, therefore many markers are needed to differentiate the cultivars. Thus, we performed genotyping using eight SSR markers and DNA fragment analysis.

A dendrogram was constructed using the DNA fragment analysis data. The correlation coefficients between the cultivars were calculated from the genetic distance matrices based on PgSSR07, PgSSR08, PgSSR09, PgSSR17, PgSSR37, PgSSR40, PgSSR51, and PgSSR53 alleles. The correlation coefficients within the I, II, and III groups were 0.77, 0.87, and 0.70, respectively. The correlation coefficients between groups I+II and III was estimated as 0.56 and the coefficient between groups I and II was 0.67. The low correlation coefficient of the third combination indicates that alleles could come from all of the three genomes instead of only one or two genomes, so that genetic relationships between varieties can be accurately evaluated (Zhang et al. 2002).

Korean ginseng cultivars grouped together. Jakyung and Hwangsook have been cultivated as domestic varieties in Korea for a long time. Chunpoong, Yunpoong, and Gopoong were developed from Jakyung by selective breeding and Gumpoong was developed from Hwangsook by selective breeding. From the results, Chunpoong, Yunpoong, and Gopoong were grouped with Jakyung; and Gumpoong was grouped with Hwangsook. Therefore, dendrogram results are consistent with the history of Korean ginseng cultivars and *Panax* species. The traditional ginseng classification was based on morphology, so we generated molecular data to establish the *Panax* species lineage. Some researchers have published molecular markers for the identification of *Panax* species (Choi et al. 2011; Bang et al. 2012). But there is no report on the DNA fragment analysis method. For this reason, this study provides a very efficient method for classification study and shows the results of marker application using molecular markers in *Panax* species.

Consequently, our result is a case study demonstrating that SSR markers developed from GSS and MEL sequences of *P. ginseng* could be used for genetic studies in *Panax* species and Korean ginseng cultivars. These markers could be used for further research on cultivar development with molecular breeding techniques and conservation of the genetic diversity of the species in the genus. Moreover, the markers could be useful for genetic mapping and eventual marker-assisted selection in the crop. The use of next-generation sequencing will enable more rapid and efficient development of SSR markers.

**ACKNOWLEDGEMENTS**

This work was carried out with the support of “Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01102202)”, Rural Development Administration, Republic of Korea.

**REFERENCES**


Mogea JP. 1999. Relationships and phylogeny of the species


Yang DC, Kim MS. 2003. DNA analysis of ginseng using PCR-aided RFLP technology. J. Ginseng Res. 27: 146-150.
