Development of Cleaved Amplified Polymorphic Sequence Markers for Classifying Ginger (Zingiber officinale) Cultivars Using Reference Sequencing

Ji-Nam Kang¹, Gyeong-Hui Lee², Jin Yu³, Mi-Hwa Choi¹, Simyung Lee¹*
¹Genomics Division, National Institute of Agricultural Sciences, Jeonju 54874, Korea
²Wanju-gun Agricultural Technology Center, Wanju 55310, Korea
³Ginseng Division, Department of Herbal Crop Research, Eumseong 27709, Korea

ABSTRACT  Ginger (Zingiber officinale) is grown worldwide in subtropical and tropical regions and primarily used as a spice and medicinal plant. Despite the economic importance of ginger, research on its molecular aspects is limited. Moreover, although ginger is mainly cultivated through vegetative propagation owing to poor flowering and infertility, few molecular markers have been identified to distinguish cultivars. In this research, we developed five Cleaved Amplified Polymorphic Sequence (CAPS) markers that can distinguish between the “Bongdong” ginger (Bg) cultivar, indigenous to South Korea, and the Chinese imported ginger (Cg) cultivar through reference sequencing based on the recently reported complete genome information of ginger. Furthermore, the integrated application of the five CAPS markers allow us to distinguish between Bg, Cg, and Indonesian ginger. Among them, the ClaI-based CAPS marker was identified as specific to Bg cultivars. Therefore, TaqMan real-time PCR based on ClaI-based CAPS can be widely used to distinguish between Bg and Cg cultivars. This study is the first to report the development of genome-based single-nucleotide polymorphism markers in ginger and therefore provides important information for the breeding and conservation of ginger.

Keywords  Ginger, CAPS marker, TaqMan real-time PCR

INTRODUCTION  Ginger (Zingiber officinale) is a perennial monocotyledonous herb that belongs to the Zingiberaceae family. Ginger is rich in secondary metabolites such as oleoresin and shogaol, which are the main substances responsible for the spiciness and flavor of ginger (Shivakumar 2019). In addition, ginger exhibits pharmacological activity for the treatment of diabetes, obesity, diarrhea, allergy, pain, and inflammation, and has traditional applications for human health (Dhanik et al. 2017). This crop is now grown worldwide in subtropical and tropical regions, including South Korea, Japan, China, India, Brazil, Australia, West Africa, and Southeast Asia (Mahdi et al. 2013). Ginger typically propagates through the formation of new plants from rhizomes owing to its poor flowering ability and the infertility of flowers; as a result, ginger breeding occurs predominantly through vegetative growth (Shivakumar 2019). Nevertheless, ginger exhibits diverse morphological phenotypes, with more than 25 species of ginger cultivated worldwide (Nayak et al. 2005; Bhatt et al. 2013). Several promising cultivars have been characterized based on morphological, biochemical, major secondary metabolite content, and yield data; however, these characteristics may differ in their habitat and growing conditions (Nayak et al. 2005; Mahdi et al. 2013). Therefore, it is
important to clearly characterize the genetic diversity of indigenous populations for the development and conservation of various ginger cultivars.

Molecular markers enable characterization within plant species through direct access to plant genotypes (Ismail et al. 2016). To date, researchers have developed many molecular markers based on restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), inter-simple sequence repeat (ISSR), simple sequence repeats (SSR), single-nucleotide polymorphism (SNP), diversity array technology, and retrotransposons, and applied these molecular markers in plants (Nadeem et al. 2018). Among these, RAPD (Jatoi et al. 2008; Kizhakkayil and Sasikumar 2010; Sajeev et al. 2011; Ghosh and Mandi 2011; Ashraf et al. 2014; Bua-in and Paisooksantivatana 2010), AFLP (Ghosh et al. 2011), SSR (Mahdi et al. 2013; Pandotra et al. 2013), and ISSR (Kizhakkayil and Sasikumar 2010; Pandotra et al. 2013) markers have been used to identify genetic diversity and interspecies relationships in ginger, suggesting that ginger retains a wide genetic diversity despite vegetative growth. However, studies based on the molecular markers of ginger are rare, with most studies limited to polymorphism identification techniques using random genome excision and polymerase chain reaction (PCR).

Molecular markers are powerful tools for characterizing plant genotypes; however, the developed marker systems are not necessarily applicable to different plant populations. The application of molecular markers differs according to the specific purpose, which suggests that various molecular markers should be identified for ginger. Cleaved amplified polymorphic sequence (CAPS) markers have been successfully applied in plants as molecular markers by combining PCR and RFLP methods (Shavrukov 2016). These markers identify variants by amplifying small DNA fragments (amplicons) from the whole genome then applying the appropriate endonuclease (restriction enzyme) to the SNP region. For example, the application of a restriction enzyme that recognizes the variant site in an amplicon amplified with the same primer set will generate two fragments; otherwise, it will retain the original amplicon. Therefore, the identification of variant sites applicable to restriction enzymes is the most important factor for CAPS marker development (Shavrukov 2016).

Recent advances in genome sequencing technology have revolutionized the development of molecular markers, along with our knowledge of whole genomes. Technologies such as next-generation sequencing (NGS) and genotyping-by-sequencing can help identify numerous SNPs that cause high polymorphism between or within species by employing known genomic sequences. SNPs based on single-nucleotide substitutions are the simplest genetic units and provide the largest number of markers (Nadeem et al. 2018). In this regard, NGS technology may provide greater opportunities to identify CAPS in organisms. The combination of NGS and CAPS has enabled the development of effective CAPS markers for various plants (Shavrukov 2016). For example, whole-genome shotgun sequencing enabled the development of CAPS markers for powdery mildew resistance in soybean (Glycine max) (Jun et al. 2012). Illumina golden gate technology was effectively used to develop CAPS markers for micromapping the Barley Stripe Mosaic Virus Resistance Gene in Brachypodium distachyon (Cui et al. 2012). High-throughput reference sequencing (Re-Seq) using Illumina HiSeq enabled the identification of 2,458 CAPS markers in watermelon (Citrullus lanatus), which significantly increased the rate of molecular marker development for this crop (Shavrukov 2016; Liu et al. 2015).

SNP molecular markers are the most unique molecular markers in ginger (Ismail et al. 2016); however, CAPS markers have not previously been applied to ginger, which may be attributed to limited genetic information on ginger. A recent study reported a ginger genome with a total length of 3.1 Gb consisting of 11 chromosomes (haploid genome size of approximately 1.55 Gb) (Cheng et al. 2021). Using this ginger reference genome, we develop a marker for distinguishing between indigenous ginger with a long cultivation history in the “Bongdong” region of South Korea and an imported Chinese ginger cultivar through Re-Seq. Compared to the reference genome sequence, we identify a total of 3,040,656 variants from the two ginger varieties. Five CAPS markers selected by rigorous filtering clearly differentiate the two cultivars. Furthermore, the application of these CAPS markers to ginger cultivars collected from various regions shows that the ClaI-based
CAPS marker can be specifically applied to the “Bongdong” cultivar. This representative study is the first to report the development of genome-based CAPS markers in ginger and therefore provides important information for the preparation of institutional devices to protect the intellectual property rights of domestic ginger breeders and ginger breeding.

MATERIALS AND METHODS

Plant materials

A total of 63 ginger samples were collected in this study. Five rhizomes of the “Bongdong” ginger (Bg) cultivar were collected from five separate farms located in Bongdong-eup, Wanju-gun, Jeollabuk-do, Republic of Korea (Bg_1-, 2-, 3-, 4-, and 5-). Five rhizomes of Chinese imported ginger (Cg) cultivars were collected from five separate farms located in the same region (Cg_6-, 7-, 8-, 9-, and 10-). Additionally, three ginger plants grown at the Wanju-gun Agricultural Technology Center were collected as tissue culture samples for maintenance and preservation of the local indigenous cultivar Bg (Bg_0-). Two ginger rhizomes collected from Indonesia (G-1 and G-2), four ginger rhizomes collected from different regions in South Korea (G-3, G-4, G-7, and G-10), two rhizomes from Cg (G-5 and G-6), and two ginger rhizomes collected from China (G-8 and G-9) were obtained from the National Institute of Horticultural and Herbal Science (NIHHS) (Supplementary Table S1).

DNA isolation and quality control

Genomic DNA from 63 ginger samples was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer’s instructions. To extract high-quality DNA from the rhizomes, the homogenized ginger samples were incubated in a water bath at 65°C for 3 hours. The total DNA quantity and quality were measured using an Agilent 2200 TapeStation (Agilen Technologies, California, USA). Four DNA samples (Bg_0-1, Bg_1-1, Bg_2-1, and Bg_3-1) from the Bg cultivar and four DNA samples (Cg_6-1, Cg_7-1, Cg_8-1, and Cg_9-1) from the Cg cultivar were selected for Re-Seq. These DNA samples were confirmed to have a DNA integrity number of seven or higher, which means that these DNA samples were extracted with high quality. Finally, the absence of RNA contamination was confirmed using 0.8% agarose gel electrophoresis.

Whole genome sequencing and mapping

Eight genomic DNA samples were used to construct a paired-end sequencing library with an insert size of 550 bp, using the TruSeq DNA PCR-free kit (Illumina, USA). The prepared libraries were sequenced on the Illumina NovaSeq platform (Macrogen, Seoul, Korea). Low-quality and redundant reads were removed from the total reads generated using the Trimmomatic program (ver. 0.39) (Bolger et al. 2014). High-quality reads were mapped onto the ginger reference genome (ver. Zo_v1.1) (Cheng et al. 2021) using the mem option of the Burrows-Wheeler Aligner program (ver. 0.7.17) (Li and Durbin 2009).

Variant filtering and annotation

Unmapped and secondary-aligned reads were removed using SAMtools software (ver 1.11) with sort, fixmate, and markdup parameters (Li et al. 2009). The mapping result was then converted to BAM format. Variation calling was performed using the HaplotypeCaller module of the Genome Analysis Toolkit (GATK, ver. 4.2) (McKenna et al. 2010), and VCF files were generated. The GATK hard filter (SNP: QD < 2.0 | FS > 60.0 | MQ < 40.0 | MQRankSum < −12.5 | ReadPosRankSum < −8.0, In/Del: QD < 2.0 | FS > 200.0 | ReadPosRankSum < −20.0) was adjusted to remove false-positive SNPs. The annotation of identified SNPs and InDels was performed using SnpEff software (ver. 5.2e) (Cingolani et al. 2012). Variations present in the gene regions (from 5’ UTR to 3’ UTR, including introns and exons) were annotated as genic, whereas those in other genomic regions were annotated as intergenic. Variations in the coding sequences were further divided into synonymous and non-synonymous sequences.

CAPS primer design and PCR

For CAPS primer design, a 500-bp flanking sequence from the variant site was extracted from the reference genome sequence. PCR primer sets were designed using
CAPS-finder (https://github.com/mfcovington/CAPS-finder/blob/master/CAPS-finder.pl) and included the following parameters: primer size, 19-22 mer; Tm of 55-62°C; and amplicon size of 300-700 bp. The specificity of the designed primers was confirmed by BLASTN searches (cutoff E-value 1e-5) against the reference genome sequences.

The PCR reaction mixture (total 50 µL) included 1 µL (20-30 ng) of genomic DNA, 1 µL (10 pmol) of primer set, 4 µL of dNTP mixture (2.5 mM each), 10 µL of 5 × PrimeSTAR GXL buffer, and 1 µL (1.25 unit) of PrimeSTAR GXL DNA Polymerase (Takara, Kusatsu, Japan). PCR was performed using an ABI GeneAmp 9700 PCR Thermal Cycler (Applied Biosystems, Waltham, USA) under the following conditions: 10 minutes at 98°C for initial denaturation, followed by 35 cycles of 98°C for 10 seconds, 58°C for 15 seconds, and 68°C for 1 minute.

Restriction enzyme treatment and electrophoresis

All PCR products were verified by 2.5% agarose gel electrophoresis. The target fragment was purified using an Expin™ Gel SV kit (GeneAll, Seoul, Korea) according to the manufacturer’s instructions. BglII, DraI, ClaI, NaeI, SspI, and DdeI restriction enzymes were purchased from Enzymomics (Daejeon, Korea). A total of 50 µL of reaction solution containing 30 µL of PCR purified product, 20 units of each restriction enzyme, and 5 µL of 10 × activation buffer was incubated at 37°C for 16 hours. Then, 10 µL of the reaction solution was loaded onto a 2.5% agarose gel with 10 µL of sterile water, followed by electrophoresis at 100 V for 25 minutes.

TaqMan probe design and real-time PCR

TaqMan real-time PCR primer sets and probes were specifically designed for SNP validation of Clal-based CAPS using Primer Express (Thermo Fisher Scientific). The forward and reverse primer sequences were 5'-AA TTTGCCACACTATATTAAAAACAGTAATTTTT-3' and 5'-TACTATGCCGATGAATAACATCCCG-3', respectively. The probe set 5'-TTTAATCGATTGATACAA TGT-3' (Probe-1) or 5'-TTTAATCGATTGATACAA TGT-3' (Probe-2) was designed based on the SNP region G in the ClalI recognition sequence ATCGAT. The 5'-ends of probe-1 and probe-2 were labeled with asymmetric xanthene (VIC) and carboxyfluorescein (FAM), respectively, and a nonfluorescent quencher-minor groove binder was attached to the 3’-end as a fluorescent quencher.

The real-time PCR reaction mixture (total 10 µL) included 1 µL (10-20 ng) of genomic DNA, 0.5 µL of 20 × TaqMan™ SNP Genotyping Assay solution mixed with primer and prober, and 5.0 µL of TaqPath™ ProAmp™ master Mix (Thermo Fisher Scientific, Waltham, USA). Real-time PCR was performed using QuantoStudio 1 (Thermo Fisher Scientific, Waltham, USA), and the reaction conditions were as follows: 30 seconds at 60°C and 5 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Allelic discrimination plots were generated based on Delta-Rn values using QuantoStudio 1 analysis software.

RESULTS

Re-Seq data production and reference genome mapping

The production of high-quality sequencing data is essential for the development of molecular markers to distinguish between ginger cultivars. Here, we generated high-quality sequence data from eight samples of ginger DNA (four Bg cultivars and four Cg cultivars) using Illumina sequencing. Read bases with an average of 39,469,764,821 bp (ranging from 37,868,502,686 to 41,244,609,006 bp) and 261,389,171 reads (ranging from 250,784,786 to 273,143,106) were generated from the eight DNA samples. After quality trimming, the average of these read bases was 35,439,687,045 bp (ranging from 33,924,428,140 to 36,919,318,239 bp), which is an average of 11.47 times larger than the reference genome sequence (3,090,429,014 bp) (Cheng et al. 2021) (Supplementary Table S2).

Mapping of the reference genome was performed using trimmed reads. An average of 241,335,907 trimmed reads (ranging from 231,030,074 to 251,534,965) were mapped to the reference genome as input reads and an average of 96.2% (ranging from 95.6% to 97.3%) of these reads hit the reference genome. The final coverage of the reference genome was an average of 3,005,911,374 bp (ranging from
2,996,879,881 to 3,009,021,425 bp), which covered an average of 97.3% (ranging from 97.0% to 97.4%) of the entire reference genome at a depth of 1× (Supplementary Table S3). These results showed that the Re-Seq data were properly mapped to the ginger reference genome.

Calling and filtering variants from mapping data

Compared to the reference genome sequence, a total of 3,040,656 variants were called with annotation information from the eight ginger mapping data samples. The variant calling information for each sample is presented in Supplementary Table S4. To select distinct SNP and InDel sites from these vast variants, we filtered using the following parameters: from a total of 3,040,656 variants, the use of multi-allele variant removal, GATK hard filtering (see Materials and Methods), variants genotyped in all samples (genotype call rate = 1.0), and removal of variants with the same genotype in all samples reduced the number of variants to 1,154,942. Furthermore, as a result of screening for genotype variants that were identical within the Bg group or within the Cg group, but different between the two groups, the remaining number of variants was 3,467 (Fig. 1a). These variants were classified into 2,600 SNPs, 867 InDels, 2,315 homozygous (Homo) variants, 1,152 heterozygous (Hetero) variants, 323 genic variants, and 3,144 intergenic variants (Table 1). Detailed information on the 3,467 selected variants is presented in Supplementary Table S5.

Because 3,467 selected variants are still a large number, additional filtering was performed based on 2,315 Homo-variants. Among a total of 2,315 homo-variants, 609 homo-variants with read depths of 5 or higher were identified.

Table 1. Summary of variant information between Bg and Cg groups.

<table>
<thead>
<tr>
<th>Total variants</th>
<th>SNPs</th>
<th>InDels</th>
<th>Homo variants</th>
<th>Hetero variants</th>
<th>Genic variants</th>
<th>Intergenic variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>3467</td>
<td>2600</td>
<td>867</td>
<td>2315</td>
<td>1152</td>
<td>323</td>
<td>3144</td>
</tr>
</tbody>
</table>

Fig. 1. Variants filtering process. (a) Total number of variants (3,040,656) was reduced to 3,467 by five-step filtering. (b) Homo-variants with a reading depth ≥ 5× were screened.
including 160 InDel and 227 CAPS types (Fig. 1b). We extracted approximately 500 bp of flanking sequences from these 387 variant regions and attempted to design applicable primer sets. BLASTN analysis showed that the primer sets designed for the 160 InDel markers were not specific. However, a specific primer set applicable to 15 CAPS sites was designed; these CAPS sites have been identified as recognition sites for restriction enzymes including BglII, BstKI, DraI, CviQI, ClaI, TaqI, HpaII, NaeI, CviAI, AseI, SspI, DdeI, Hinfl, HpyCh4V, BsrUI, StyI, BstBI, and MvaI (Supplementary Table S6).

**Confirmation of polymorphism by CAPS digestion from amplicon**

Among the 15 CAPS marker candidates screened, we finally selected six CAPS to which BglII, DraI, ClaI, NaeI, SspI, and DdeI restriction enzymes were selected. The selected restriction enzymes were six or five cutting restriction enzymes, which were judged to increase the accuracy of CAPS markers. Table 2 shows the polymorphic pattern predictions for each amplicon using these restriction enzymes.

PCR for polymorphism verification was performed using five DNA samples (Bg_0-1, Bg_1-1, Bg_2-1, Bg_3-1, and Bg_4-1) from the Bg group and five DNA samples (Cg_6-1, Cg_7-1, Cg_8-1, Cg_9-1, and Cg_10-1) from the Cg group. Amplicons of the expected size were identified in all samples (Table 2 and Fig. 2a). The restriction enzyme treatment results for each amplicon showed that BglII-, DraI-, ClaI-, and SspI-based CAPS markers generated two distinct fragments at their predicted positions. Only one fragment was identified for the DraI-based CAPS marker near the predicted cleavage site, which is likely because the two predicted fragment sizes were very similar and appeared to merge into one fragment. The NaeI-based CAPS marker was cleaved to the expected size only in the Bg group, but an amplicon of 521 bp that was not cleaved in all Bg groups was identified. NaeI-based CAPS in the Bg group is likely to exist in a heterozygous form (Table 2 and Fig. 2b). Among the six CAPS marker candidates, BglII, DraI, ClaI, NaeI, and SspI-based CAPS markers were identified as molecular markers capable of distinguishing between Bg and Cg cultivars.

**Verification of polymorphisms in various ginger resources using developed CAPS markers**

To test the coverage of the five developed CAPS markers, we obtained ginger resources collected from Indonesia, South Korea, and China and obtained from NIHHS. The labels of these ginger resources are based only on the collection area; thus, their origin may be obscure. The results of applying five CAPS markers to ten ginger resources suggested the following results. First, various ginger cultivars are cultivated in each region: G-3 (lane 5) and G-7 (lane 9) are resources collected from “Iksan (IKS)” and “Gangjin (GAN)” regions in South Korea, respectively, but the marker application results are predicted for Indonesian (IDN) ginger (G-1 and G-2; lanes 1 and 2). G-4 (Lane 6) is a resource collected from the “Taean (TAE)” region in South Korea, but showed the same polymorphic pattern as the Cg cultivar (PC-2; lane 2). G-8 (lane 10) and

---

**Table 2. Polymorphism prediction of amplicons by CAPS digestion.**

<table>
<thead>
<tr>
<th>Reference genomeposition</th>
<th>Bg-CAPS</th>
<th>Cg-CAPS</th>
<th>Restriction enzyme</th>
<th>Predicted PCR product size (bp)</th>
<th>Amplicon cutting size</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC_055986.1_153787420</td>
<td>C</td>
<td>T</td>
<td>BglII</td>
<td>470</td>
<td>137 333</td>
</tr>
<tr>
<td>NC_055987.1_85483445</td>
<td>AC</td>
<td>A</td>
<td>DraI</td>
<td>570</td>
<td>277 293</td>
</tr>
<tr>
<td>NC_055989.1_143739477</td>
<td>G</td>
<td>A</td>
<td>ClaI</td>
<td>579</td>
<td>418 161</td>
</tr>
<tr>
<td>NC_055989.1_144594043</td>
<td>C</td>
<td>T</td>
<td>NaeI</td>
<td>521</td>
<td>390 131</td>
</tr>
<tr>
<td>NC_055995.1_7767767</td>
<td>C</td>
<td>A</td>
<td>SspI</td>
<td>572</td>
<td>222 350</td>
</tr>
<tr>
<td>NC_055997.1_6190205</td>
<td>G</td>
<td>A</td>
<td>DdeI</td>
<td>540</td>
<td>121 419</td>
</tr>
</tbody>
</table>
G-9 (lane 11) are resources collected from China (CHN) but have the same polymorphic pattern as IDN ginger. Second, the Bg cultivar can be genetically specific: Clal-based CAPS markers produced cleavage fragments specifically in only two Bg cultivars (PC-1 and G-10; lanes 1 and 12), indicating that Bg cultivars are genetically different from CHN and IDN ginger (Fig. 3). In summary, the integrated application of the five CAPS markers was able to differentiate Bg (Cluster A), Cg (Cluster B), and IDN (Cluster C) gingers.

**TaqMan real-time PCR application of Clal-based CAPS marker**

Previous results have shown that the Clal-based CAPS marker could be used as a clear molecular marker to differentiate Bg cultivars from other ginger cultivars. Here, we applied the Clal-based CAPS marker to all ginger DNA samples collected from the “Bongdong” area, and confirmed that only the Bg group produced two distinct cleavage fragments (Fig. 4a). The electrophoresis method has the disadvantage of being cumbersome and time-consuming for polymorphism identification using CAPS markers. TaqMan real-time PCR technology has successfully replaced electrophoresis-based CAPS markers in plants (Endo et al. 2020). Therefore, we employed a real-time PCR technique using TaqMan probes to improve the availability of Clal-based CAPS markers. Allele-G within the Bg groups was detected by the FAM fluorescently labeled probe, whereas allele-A within the Cg group was detected by the VIC fluorescently labeled probe. The different emission wavelengths of the two fluorescent dyes were distinguishable within the same test tube. The allelic discrimination plot based on Delta-Rn values showed a clear distinction between the Bg group (blue) and Cg group (red) groups. Additionally, the hetero-type group (green) prepared by mixing Bg group DNA and Cg group DNA also showed a clear difference from the above two groups (Fig. 4b).
Fig. 3. Prediction of ginger cultivars by CAPS polymorphism patterns. The ten ginger resources are clustered into three groups by the polymorphic patterns of the five CAPS markers. PC-1 and PC-2 are the positive controls for Bg and Cg cultivars, respectively (lanes 1 and 2). Ginger resources from G-1 to G-10 were all obtained from NIHHS. G-1 and G-2 are ginger resources collected from Indonesia (IDN) (lanes 3 and 4). G-3, G-4, and G-7 are ginger resources collected from “Iksan (IKS), “Taean (TAE)”, and “Gangjin (GAN)” regions of South Korea, respectively (lanes 5, 6, and 9). G-5 and G-6 are Cg cultivars obtained from NIHHS (lanes 7 and 8). G-8 and G-9 are ginger resources collected from China (CHN) (lanes 10 and 11). G-10 is a Bg cultivar obtained from NIHHS (lane 12). The presence or absence of cleavage by each restriction enzyme is indicated by O or X. Electrophoresis was performed at 100 V on a 2.5% agarose gel. M, 100 bp Plus DNA Ladder (Bioneer, Daejeon, Korea).

**DISCUSSION**

Although ginger is an economically important crop, research on the molecular aspects of this plant and the molecular markers distinguishing different cultivars remains insufficient (Ismail et al. 2016; Cheng et al. 2021). Limited genetic information and low genetic variation because of flower infertility are the biggest obstacles to developing molecular markers for ginger. Improved phenotypes of the progeny are selected based on important phenotypic traits (productivity and environmental tolerance) during long-term breeding, which are likely caused by natural genetic variation at specific loci (Cheng et al. 2021). Local ginger cultivars often exhibit uniform phenotypes for specific traits, which can be used for the morphological differentiation of ginger (Mahdi et al. 2013). However, it is very difficult to differentiate ginger cultivars when morphological phenotypic traits are not sufficiently unique to differentiate between ginger populations, or when used as secondary processed products in powder form, which suggests the importance of developing molecular markers for ginger. This study was conducted to detect the genetic diversity between the local indigenous ginger cultivated in the “Bongdong” region of South Korea.
and ginger imported from China, and develop a molecular marker for clearly distinguishing between these two cultivars.

Based on the recently reported complete genomic information of ginger (Cheng et al. 2021), we performed Re-Seq on Bg and Cg cultivars. As a result, we identified five CAPS markers that clearly distinguish between the two varieties. These markers were applied very stably within the Bg and Cg groups. This result indicates that the developed CAPS markers are based on homozygosity, and that the same population is likely to have fixed genetic characteristics. A genetic stability study of indigenous ginger cultivars has also been conducted for India. RAPD and ISSR results of ten indigenous gingers collected from nine sites in the “Manipur” region of northeastern India clearly revealed the genetic stability of these gingers, which is not surprising given that ginger is cultivated and propagated via vegetative growth (Singh et al. 2013). However, several studies have reported the differentiation of ginger populations according to geographic location using RAPD molecular markers. The application of RAPD markers to 12 ginger populations collected from the Indian subcontinent revealed a high degree of polymorphism between populations (Ashraf et al. 2014). Moreover, ginger populations collected from six locations in northeastern India were differentiated using RAPD markers (Sajeev et al. 2011). These studies show that RAPD markers can clearly differentiate between ginger populations according to the collection area. The five CAPS markers developed in this study also clearly revealed the genetic diversity between populations. These results suggest that, although most gingers are vegetatively propagated, they are indigenous to

Fig. 4. TaqMan real-time PCR application of ClaI-based CAPS markers. (a) Electrophoresis results of ClaI-based CAPS markers for Bg and Cg cultivars collected in the “Bongdong” region. Lanes 1 to 20 are Bg cultivars and lanes 21 to 40 are Cg cultivars. Electrophoresis was performed at 100 V on a 2.5% agarose gel. M, 100 bp Plus DNA Ladder (Bioneer, Daejeon, Korea). (b) Allelic discrimination plot. A total of 44 ginger DNA samples were used for TaqMan real-time PCR. The allele-G of 20 Bg cultivars was detected by FAM fluorescent dye linked-probes (grouped in blue). The allele-A of 20 Cg cultivars was detected by VIC fluorescent dye linked-probes (grouped as red). Green group represents DNA samples mixed with Bg and Cg cultivars, which were artificially prepared for heterozygous type detection. Black square indicates the negative control. Allelic discrimination plots were generated based on Delta-Rn values.
each region and undergo genetic mutations.

The integrated application of the five CAPS markers suggested the origin of the Bg cultivar. Specifically, the four CAPS markers revealed the same polymorphic pattern for the IDN ginger and Bg cultivars. As Southeast Asia, including Indonesia, has been reported as the origin of ginger (Ghosh and Mandi 2011; Dhanik et al. 2017; Chavan et al. 2008), it is likely that the Bg cultivar originated from Southeast Asia. Interestingly, the Clal-based CAPS marker was the only one of the five CAPS markers to be applied to Bg cultivars, which suggests that ginger from other Southeast Asian regions may be the origin of the Bg cultivar or a specific genetic variant of the Bg cultivar. In summary, the five CAPS markers identified in this study are stable and reproducible molecular markers for classifying ginger cultivars and providing insights into the geographic origin of the ginger resources used in the study.

Opportunities for crop improvement by genomic recombination are limited in ginger, as it is mainly propagated through vegetative growth (Shivakumar 2019). Therefore, the development of molecular markers is essential for ginger conservation and breeding. This study is the first to report the development of CAPS markers based on SNP variations in ginger. Integrative analysis of the five CAPS markers enabled us to distinguish between IDN ginger, the Bg cultivar, and the Cg cultivar. As the Clal-based CAPS marker was applied specifically to Bg cultivars, we expect that TaqMan real-time PCR application of the Clal-based CAPS marker will become a widely used method of easily distinguishing between Bg cultivars and Cg cultivars. Furthermore, the application of these markers to ginger cultivars collected from more diverse countries and regions could provide information regarding the diversity and geographic origin of ginger species.

ACKNOWLEDGEMENTS

We would like to thank the Wanju-gun Agricultural Technology Center and the National Institute of Horticultural and Herbal Science for supporting ginger resources.

DATA AVAILABILITY STATEMENT

All analyzed data are included in the article and its supporting information. The Re-Seq data used in the present study were submitted to the National Agricultural Biotechnology Information Center (NABIC, https://nabic.rda.go.kr) and are publicly available (Accession No.: NN-8299, NN-8300, NN-8301, NN-8302, NN-8303, NN-8304, NN-8305, and NN-8306).

FUNDING

This work was supported by the "National Institute of Agricultural Sciences" (project no. PJ015035), Rural Development Administration, Republic of Korea.

REFERENCES


