Developing PCR-Based SNP Markers for Distinguishing Korean Waxy Corn F1 Hybrids

Sang Gon Kim*, Jin-Seok Lee*, Seonghyu Shin, Hwan Hee Bae, Jung-Tae Kim, Beom-Young Son, Seong-Bum Baek

Department of Central Area Crop Science, NICS, RDA, Suwon 16429, Korea

ABSTRACT  Single nucleotide polymorphisms (SNPs) are abundantly and evenly distributed throughout the genomes of most plant species. These markers have become popular for use in genetic research in many crops. SNP markers can be used to screen maize cultivars rapidly during the early growth stages. In this study, to develop additional SNP markers for maize, we chose 20 SNP sites per chromosome from the maizeGDB website (www.maizegdb.org) and designed primers with two base pair mismatches using Primer Designer 4 based on putative SNP sites of the B73 genome sequence. The polymerase chain reaction (PCR) products ranged from 200 to 500 bp in size, whereas no PCR product was detected when the SNP site was present in Korean waxy corn. Using nine Korean commercial F1 hybrids of waxy corn, including Chalok 1, Chalok 4, Ilmichal, Eolrukchal 1, Heukjinjuchal, Hayanchal 95, Mibaekchal, Mibaek 2, and Miheukchal, we selected 16 primer sets showing clear bands or no bands. Based on cluster analysis, we confirmed that the nine Korean waxy corn hybrids could clearly be distinguished. The SNP marker sets are easy to utilize through simple PCR and agarose gel electrophoresis. These results suggest that analysis using the SNP marker set designed in this study would be faster, cheaper, and more reproducible than that using other genotyping tools, such as cleaved amplified polymorphic sequence markers, which require the use of restriction enzymes.

Keywords  Maize, Marker, PCR, Primer, SNP, Waxy corn

INTRODUCTION

Maize (Zea mays ssp. mays) is the third most highly cultivated crop (behind wheat and rice) and has the second highest production worldwide. Since its domestication approximately 9,000 years ago, maize has been widely grown around the world in both temperate and tropical regions. Cultivated maize was domesticated from the grass teosinte (Zea mays ssp. parviglumis). The morphology of maize mainly differs from that of teosinte because a few major genes for specific alleles were selected during maize domestication (Doebly et al. 2006). Homozygous inbred maize lines have been developed, and the high level of heterosis in current hybrid varieties has been achieved through the combination of different heterotic lines.

Maize can be classified into three types, including normal, waxy, and sweet, depending on the starch composition of the endosperm in the seed. Normal corn and waxy corn differ in terms of texture or starch content (amylose and amylopectin) (Nelson and Rines 1962), whereas sweet corn has high sugar contents in the seed (Carey et al. 1984). Normal corn is widely cultivated for use as a food, forage, industrial, and bioenergy crop. However, waxy corn and sweet corn are only used in food production. Waxy corn is a popular food that is widely grown in East Asian countries, especially Korea.

With the completion of the maize inbred line B73 genome sequence (Schnable et al. 2009), various tools such as quantitative trait locus analysis, positional cloning, and physical mapping have been widely used to decode the structure and function of the maize genome (Azanza et al. 1996; Ku et al. 2010; Wang et al. 2010; Choe and Rocheford 2012). In addition, numerous molecular markers for agronomically important qualitative and quantitative
traits have been developed (Prasanna et al. 2010). Among DNA-based molecular markers, simple sequence repeats have been widely used to construct linkage maps in maize (Sharopova et al. 2002; Sa et al. 2012). Single nucleotide polymorphisms (SNPs) are the most abundant genetic variations in plants (Rafalski 2002). SNP genotyping techniques have been improved based on allele-specific hybridization, primer extension, oligonucleotide ligation, endonuclease cleavage, Invader, Taqman, and allele-specific polymerase chain reaction (PCR) (Gupta et al. 2001). These methods are suitable for multiplexing for both genetic mapping and diversity studies using a large number of markers.

Recently, the single-nucleotide amplified polymorphisms (SNAP) assay, a simple, allele-specific method, was developed for SNP analysis in mapping populations of Arabidopsis (Drenkard et al. 2000). SNAP markers are primers containing a single base mismatch with the 3’ end of one allele (the specific allele). Kim et al. (2005) developed a SNAP marker for a GmNARK gene controlling supernodulation in soybean (Kim et al. 2005). In the current study, we used a modified two base pair mismatch PCR procedure to identify SNPs that are sufficiently robust to easily discriminate among various commercial Korean maize F1 hybrids. This modified procedure can facilitate the rapid, reliable creation and analysis of a number of molecular markers using simple methods common to most molecular biology laboratories.

MATERIALS AND METHODS

Plant materials and DNA extraction

Genomic DNA was isolated from the leaves of nine waxy corn varieties (Chalok 1, Chalok 4, Ilmichal, Eolrukhchal 1, Heukjinjuchal, Hayanchal 95, and Mibaekchal [developed at the National Institute of Crop Science] and Mibaek 2 and Miheukchal [developed at the Maize Experiment Station, Kangwon Agricultural Research and Extension Services]). Genomic DNA for use with the SNP markers was extracted from lyophilized leaf tissue of all F1 individuals according to a previous method (Kim et al. 2011). In brief, the tissue was combined with extraction buffer (7 M Urea, 350 mM NaCl, 10 mM Tris-Cl pH 7.6, 20 mM ethylenediaminetetraacetic acid [EDTA], and 2% sodium dodecyl sulfate) and incubated at 65°C for 10 minutes. The extracted genomic DNA was treated with phenol/chloroform/isoamyl alcohol (25:24:1 volume) and precipitated with isopropanol. The resulting pellet was washed with 70% ethanol, briefly incubated at room temperature to remove any residual ethanol, and dissolved in Tris-EDTA buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.4).

Primer design and PCR analysis

The SNP sites for primer development were identified through a search of the MaizeGDB database website (www.maizegdb.org). Optimization of melting temperature, oligonucleotide length, and the length of amplified products was achieved using Primer Designer 4 (Sci-Ed Software, Denver, CO, USA). Primer sequences were screened against the B73 maize genome sequence to minimize mis-priming. Primers for the SNP sites contained one base pair mismatch at two bases closest to the 3’ end of the forward primer.

SNP amplifications were performed in a total volume of 25 µl containing 20 ng of genomic DNA, 12.5 µl of GoTaq® Green Master Mix 2X (Promega, Madison, WI, USA), 10 µM forward and reverse primers, and nuclease-free water. The PCR profile consisted of an initial denaturation period at 95°C for 2 minutes followed by 30 cycles of 95°C for 30 seconds, 46°C to 60°C for 1 minute, and 72°C for 5 minutes. PCR products were electrophoretically separated on a 1.5% agarose gel.

Data analysis

SNP PCR products were scored across all maize varieties investigated. Cluster analysis was performed based on SNP marker scoring to classify the maize varieties. The varieties were grouped by cluster analysis based on the average linkage method using the PROC CLUSTER procedure (SAS ver. 9.2; SAS Institute, Cary, NC, USA). A clustering tree was generated using PROC CLUSTER (SAS 9.2; SAS Institute) with the average linkage method based on SNP marker scoring.
RESULTS

Design of SNP primers

Nine Korean waxy corn varieties (Chalok 1, Chalok 4, Ilmichal, Eolrukchal 1, Heukjinjuchal, Hayanchal 95, Mibaekchal, Mibaek 2, and Miheukchal) were used in this study, and 200 candidate SNP primer sets (20 SNP sites per chromosome) were designed based on a SNP site generated from the B73 maize reference genome sequence using Primer Designer 4 (Sci-Ed Software). The method used to design the SNP PCR primers is illustrated in Fig. 1. The technique utilizes primers with one base pair mismatch before the 3’ end of a SNP site to allow no PCR amplification product to be produced. As shown in Fig. 1, the forward primer had one base pair mismatch close to the SNP site at the 3’-terminal nucleotide, and the reverse primer perfectly matched the B73 genome sequence. Consequently, this primer combination led to the production of PCR products (Fig. 1A). When a maize variety exhibited a DNA sequence variation at a SNP site compared to the B73 genome sequence, a PCR product was not created due to the presence of a two base pair mismatch close to the 3’-terminal nucleotide of the forward primer (Fig. 1B).

Screening for optimized SNP primer sets

At the beginning of SNP marker development, we had to optimize PCR conditions. The 200 SNP forward primers contained 20 to 25 bases complementary to the genome of interest and were selected to give a Tm of 46°C to 64°C, as calculated by the %GC method using Primer Designer 4 software. These primers contained the second nucleotide mismatch close to the SNP site of the 3’-terminus. The sequences for the reverse primers were chosen to avoid inverted repeats and overlaps. To optimize the SNP PCR results, we screened 200 SNP primer sets and removed those with different sized or nonspecific PCR products and retained those producing a single-band PCR product of the expected size (Fig. 2).

The 200 SNP primer pairs were mainly targeted to the nine Korean waxy corn varieties. We analyzed the PCR-amplified products by 1.5% agarose gel electrophoresis and evaluated the presence or absence of bands in sets of reactions for each primer pair. Fig. 3 shows the results obtained using 16 primer pairs (IBM440, IBM471, IBM516, IBM535, IBM549, IBM580, IBM652, IBM795, IBM911, IBM960, IBM1118, IBM1269, IBM1492, IBM1601, IBM1619, and IBM1798) using 30 cycles of

![Fig. 1](image_url). Schematic representation of the single nucleotide polymorphism (SNP) polymerase chain reaction (PCR) strategy using two mismatches. Oligonucleotide SNP primers with one mismatch site close to a SNP site are used to preferentially not amplify specific SNP site alleles. (A) Primer P1 forms a perfect match complementary to a DNA sequence without the second mismatched nucleotide close to the SNP site. (B) Primer P1 forms a mismatch with both the second mismatched nucleotide close to the SNP site and the SNP site, but other complementary sequences form a perfect match.
Fig. 2. Analysis of the specificity of single nucleotide polymorphism (SNP) primer pairs among nine Korean waxy corn varieties. Polymerase chain reactions (PCRs) for each SNP primer were performed using DNA from nine Korean waxy corn varieties for 30 PCR cycles. SNP primers producing unexpected or nonspecific PCR products were not used for SNP marker development.

Fig. 3. Photographs of gels of polymerase chain reaction products using single nucleotide polymorphism (SNP) primer sets for nine Korean waxy corn varieties (Chalok 1, Chalok 4, Ilmichal, Eolrukel 1, Heukjinjuchal, Hayanchal 95, Mibaekchal, Mibaek 2, and Miheukchal). Sixteen SNP markers (IBM440, IBM471, IBM516, IBM535, IBM549, IBM580, IBM652, IBM795, IBM911, IBM960, IBM1118, IBM1269, IBM1492, IBM1601, IBM1619, and IBM1798) are shown on 1.5% agarose gels. Varieties including SNP sites were not amplified. M: marker.
PCR amplification. Five SNPs were discovered on chromosome 2, including those amplified by primer sets IBM440, IBM471, IBM516, IBM535, and IBM549. Chromosome 4 and 8 each contained three SNPs (primer sets IBM795, IBM911, and IBM960 and IBM1492, IBM1601, and IBM1619, respectively), and chromosome 3 contained two SNPs (primer sets IBM580 and IBM652). One SNP per primer set was detected on chromosome 5, 6, and 9 (Table 1).

To confirm the reliability of the 16 SNP primer sets, we carried out SNP PCR using the parental lines of Chalok 4, Ilmichal, Eolrukchal 1, and Heukjinjuchal (Fig. 4): the primer sets for each hybrid were analyzed to detect their SNP sites in the respective parental lines. For Chalok 4, primer sets IBM911 and IBM960 detected a SNP site in KW35 and KW33, respectively, indicating that the SNP sites are heritable and the respective primer sets are reliable after hybridization. Other primer sets also revealed SNP sites between inbred parental lines and F1 hybrid lines. Interestingly, primer set IBM960 did not reveal a SNP site in KW35, the inbred parental line of Chalok 4 and Ilmichal (Fig. 4), and IBM911 did not reveal a SNP site in KW33, the inbred parental line of Chalok 4 and Eolrukchal 1 (Fig. 4).

Minimum SNP primer combination and cluster analysis for distinguishing Korean waxy corn F1 hybrids

We determined the smallest SNP combination needed to discriminate among the nine Korean waxy corn varieties, finding that only the IBM 440 SNP primer set could separate Chalok 4 from the nine Korean waxy corn varieties and that Ilmichal, Eolrukchal 1, and Heukjinjuchal could be distinguished using two SNP primer sets (Ilmichal: IBM 580, IBM 1619; Eolrukchal 1: IBM 911, IBM 1118; Heukjinjuchal: IBM 535, IBM 795) (Fig. 5). We examined the Korean waxy corn F1 hybrids using cluster analysis, a statistical method used to determine the distance among the nine varieties based on SNP scoring depending on the presence of PCR products (Fig. 6). The results show that the nine Korean waxy corn varieties could be distinguished using 16 SNP primer sets (Fig. 6).

Table 1. Single nucleotide polymorphism primer information about the specific markers developed for nine Korean waxy corn varieties.

| Primer name | Forward sequence | Reverse sequence | Temperature (°C) | Product size (bp) | Genetic coordinate (cM) | B73 genome coordinate (bp) | Chromosome number |
|-------------|-----------------|-----------------|------------------|-------------------|-------------------------|---------------------------|------------------|
| IBM440     | GCAGCTCTTAAAGATGGAGTGGGATTGACCGCCGATCAGCTTCAT | 60 | 288 | 90.3 | 169024048 | 2 |
| IBM471     | TTCTGCTGTGGTAAATGCAATTGCTGTCCAACACCTTGCTTA | 53 | 390 | 109.8 | 185379678 | 2 |
| IBM516     | AACGAGCCTTTCATTCTTGCAGATTACAAGCCACCGCTCTTGGTA | 58 | 319 | 143.1 | 211527046 | 2 |
| IBM535     | TGCTGGATTCGCTCATCCGGTGCGTCGGCAGCAAGGAACAT | 64 | 251 | 160.9 | 223290506 | 2 |
| IBM549     | TCGCAATCGATCACAGCTGTATAAGCAGCATCTTCTCCATCAAGG | 56 | 316 | 178.2 | 232392706 | 2 |
| IBM580     | TGAAGATATACCGGCTTGTGCCGTGCTCCATGCTGCACGAAT | 58 | 254 | 42.6 | 73153892 | 3 |
| IBM652     | GACGCTATTTTTTTTACTCCAGTGGACGC ACTTGTACCGAGCGTTCAT | 57 | 339 | 100.3 | 166400256 | 3 |
| IBM795     | GCTGTATGTAATGTGACCTACACGTGTTGTCCACGTCTGTCAAC | 51 | 250 | 36.6 | 16344879 | 4 |
| IBM911     | TTGGTACTTGGTATCACAATGGCATGTGTCCAGTGTTCTC | 50 | 227 | 131.2 | 191767453 | 4 |
| IBM960     | GCAAATGTGTGAACAATAACCCAGGAGGCAAGAGCATCACTGTC | 52 | 352 | 167.1 | 237321034 | 4 |
| IBM1118    | AGCGGATCCCACGTTTCTTGCGTGTTAGCTGTGCATCCATGA | 58 | 224 | 102.6 | 73153892 | 5 |
| IBM1269    | CACCTGTAGCTGTAGCTACCACCGAGGTGGTAAGTCTTG | 46 | 228 | 79.7 | 151442112 | 6 |
| IBM1492    | GGCTGACGAGTGACAGTATTTTACTGTTGAATGTGGCGGCACTT | 54 | 271 | 17.9 | 162758085 | 8 |
| IBM1601    | AGGCAAGGATCCTGTAATCCTTTOGAGCTTCTTCTGCAACAC | 51 | 337 | 112.5 | 16706645 | 8 |
| IBM1619    | TGCACTTTTGCACTGTTCATTTCCTGACGACGACGACGACGAC | 58 | 257 | 128.7 | 15478080 | 9 |
**Fig. 4.** Examples of validation results of selected single nucleotide polymorphism (SNP) primer sets (IBM471, IBM549, IBM911, IBM960, IBM1601, and IBM1118). SNP polymerase chain reaction using parental lines of Chalok 4, Ilmichal, Eolrukchal 1, and Heukjinjuchal were carried out to confirm the reliability of the 16 SNP primer sets. KW33, KW35: inbred parental lines of Chalok 4; KW51, KW35: inbred parental lines of Ilmichal; KW23, KW33: inbred parental lines of Eolrukchal 1; KW24, KBW2: inbred parental lines of Heukjinjuchal.

**Fig. 5.** Minimum single nucleotide polymorphism primer combination required for breed discrimination. IBM440 is the only marker for Chalok 4, as it does not produce bands. The combination of primers IBM580 and IBM1619 and primers IBM535 and IBM795 can discriminate between Ilmichal and Heukjinjuchal, respectively, as they do not produce bands. The combination of IBM911 and IBM1118 can discriminate Eolrukchal 1, as they produce bands. M: marker.
DISCUSSION

After the completion of the genome sequence of maize inbred line B73 (Schnable et al. 2009), the SNAP assay was utilized for SNP analysis of mapping populations (Drenkard et al. 2000). In this study, we used a modification of the original allele-specific PCR methodology in which we utilized an additional mismatch within the last four bases of a primer to increase primer specificity and to reliably discriminate between alleles (Newton et al. 1989; Kwok et al. 1994). The use of this modification allowed us to generate primers with sufficient selectivity for use as molecular markers, as mismatch extension can vary significantly depending on the sequences surrounding the mismatch (Mendelman et al. 1989; Huang et al. 1992). Indeed, we speculate that the inability to obtain specific primers for some SNP sites could be due to particular characteristics of the sequence surrounding the SNP.

Compared to other methods used for high-throughput SNP detection performed in genome centers (Ahmadian et al. 2000; Alderborn et al. 2000), the SNP methodology described in this report does not require special equipment or sophisticated methodologies to detect SNPs. Since PCR-based SNP marker detection requires only a simple gel-based assay, this method is accessible to any molecular biology laboratory. Moreover, the generation and analysis of markers used to discriminate among varieties using our

![Clustering tree showing the results of principal component analysis-based nine Korean waxy corn varieties using the PROC CLUSTER procedure (SAS 9.2).](image)

The 16 SNP primer sets identified in this study could distinguish both Korean waxy corn varieties and their inbred parental lines, and the respective SNP sites are assumed to be heritable. The failure to generate specific primers in some cases may be due to mismatch extension depending on the sequences surrounding the mismatch (Mendelman et al. 1989; Huang et al. 1992).
method is relatively inexpensive. Since the SNP method is based on the generation or lack of generation of an amplified DNA sequence, designing high-throughput methods for detecting SNP markers using this method should be relatively straightforward.

ACKNOWLEDGEMENTS

This work was carried out with the support of the “Cooperative Research Program for Agriculture Science & Technology Development (Project title: Breeding of edible corn inbred lines, Project No. PJ00874702)” funded by the Rural Development Administration (RDA), Republic of Korea. Sang Gon Kim was supported by a 2016 Postdoctoral Fellowship from the National Institute of Crop Science, RDA, Republic of Korea.

REFERENCES

Ahmadian A, Gharizadeh B, Gustafsson AC, Sterky F, Nyrén P, Uhlén M, et al. 2000. Single-nucleotide polymorphism analysis by pyrosequencing. Anal. Biochem. 280: 103-110.

Alderborn A, Kristofferson A, Hammerling U. 2000. Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. Genome Res. 10: 1249-1258.

Azanza F, Tadmor Y, Klein BP, Rocheford TR, Juvik JA. 1996. Quantitative trait loci influencing chemical and sensory characteristics of eating quality in sweet corn. Genome 39: 40-50.

Carey EE, Dickinson DB, Rhodes AM. 1984. Sugar characteristics of sweet corn populations from a sugary enhancer breeding program. Euphytica 33: 609-622.

Choe ES, Rocheford TR. 2012. Genetic and QTL analysis of pericarp thickness and ear architecture traits of Korean waxy corn germplasm. Euphytica 183: 243-260.

Doebly JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. Cell 127: 1309-1321.

Drenkard E, Richter BG, Rozen S, Stutius LM, Angell NA, Mindrinos M, et al. 2000. A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in Arabidopsis. Plant Physiol. 124: 1483-1492.

Gupta PK, Roy JK, Prasad M. 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr. Sci. 80: 524-535.

Huang MM, Arnheim N, Goodman MF. 1992. Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. Nucleic Acids Res. 20: 4567-4573.

Kim MY, Van K, Lestari P, Moon JK, Lee SH. 2005. SNP identification and SNAP marker development for a GmNARK gene controlling supernodulation in soybean. Theor. Appl. Genet. 110: 1003-1010.

Kim SG, Kim ST, Wang Y, Yu S, Choi IS, Kim YC, et al. 2011. The RNase activity of rice probenazole-induced protein1 (PBZ1) plays a key role in cell death in plants. Mol. Cells 31: 25-31.

Ku LX, Zhao WM, Zhang J, Wu LC, Wang CL, Wang PA, et al. 2010. Quantitative trait loci mapping of leaf angle and leaf orientation value in maize (Zea mays L.). Theor. Appl. Genet. 121: 951-959.

Kwok S, Chang SY, Sninsky JJ, Wang A. 1994. A guide to the design and use of mismatched and degenerate primers. PCR Methods Appl. 3: S39-S47.

Mendelman LV, Boosalis MS, Petruska J, Goodman MF. 1989. Nearest neighbor influences on DNA polymerase insertion fidelity. J. Biol. Chem. 264: 14415-14423.

Nelson OE, Rines HW. 1962. The enzymatic deficiency in the waxy mutant of maize. Biochem. Biophys. Res. Commun. 9: 297-300.

Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, et al. 1989. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res. 17: 2503-2516.

Prasanna BM, Pixley K, Warburton ML, Xie CX. 2010. Molecular marker-assisted breeding options for maize improvement in Asia. Mol. Breed. 26: 339-356.

Rafalski A. 2002. Novel genetic mapping tools in plant: SNPs and LD-based approaches. Plant Sci. 162: 329-333.

Sa KJ, Park JY, Park KC, Lee JK. 2012. Analysis of genetic mapping in a waxy/dent maize RIL population using SSR and SNP markers. Genes Genom. 34: 157-164.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. 2009. The B73 maize genome: complexity,
diversity, and dynamics. Science 326: 1112-1115.
Sharopova N, McMullen MD, Schultz L, Schroeder S, Sanchez-Villeda H, Gardiner J, et al. 2002. Development and mapping of SSR markers for maize. Plant Mol. Biol. 48: 463-481.

Wang YZ, Li JZ, Li YL, Wei MG, Li XH, Fu JF. 2010. QTL detection for grain oil and starch content and their associations in two connected F$_{2,3}$ populations in high-oil maize. Euphytica 174: 239-252.