Research Paper

Selection criteria for SNP loci to maximize robustness of high-resolution melting analysis for plant breeding

Yoshiyuki Yamagata[1], Atsushi Yoshimura[1], Toyoaki Anai[2] and Satoshi Watanabe*[2]

[1] *Faculty of Agriculture, Kyushu University*, 744 Motooka, Nishi, Fukuoka 819-0395, Japan
[2] *Faculty of Agriculture, Saga University*, 1 Honjo-machi, Saga, Saga 840-8502, Japan

DNA markers are useful for identifying genes and developing new genetic materials for breeding and genetic research. High-resolution melting (HRM) analysis can detect a single nucleotide polymorphism (SNP) in two polymerase chain reaction (PCR) fragments as a melting temperature (Tm) difference without additional experimental steps, such as gel electrophoresis. To design a method for developing reliable HRM markers that discriminate between homozygous alleles containing SNPs, we tested new evaluation indexes related to the thermodynamics of double-stranded DNA to find one that maximizes the difference in Tm values between PCR fragments. We found that differences in the change in Gibbs free energy ($\Delta G°$) correlated with actual differences in Tm values. Optimization of the nearest neighboring nucleotide (NNN) of a SNP by nucleotide substitution in the primer and reducing the size of the PCR fragment both enlarged the actual differences in Tm. The genetic DNA markers we developed by NNN substitution, termed NNNs-HRM markers, could be precisely mapped within soybean chromosomes by linkage analysis. We developed a Perl script pipeline to enable the automatic design of a massive number of NNNs-HRM markers; these scripts are freely available and would be useful for practical breeding programs for other plant species.

Key Words: high-resolution melting, DNA marker, thermodynamics of DNA, plant breeding, marker-assisted selection.

Introduction

DNA markers are an indispensable tool in plant and animal breeding. Polymorphisms based on restriction fragment length (RFLP), simple sequence length (SSLP), amplified fragment length (AFLP), and random amplified polymorphic DNA (RAPD) have been widely used as DNA markers in genetic analysis and marker-assisted selection (MAS) in practical breeding programs (Collard and Mackill 2008). More recently, with the advent of high-throughput genotyping platforms based on third-generation sequencing technology, genome-wide single nucleotide polymorphisms (SNPs) have become easily attainable for use as markers (Nielsen *et al.* 2011, Pabinger *et al.* 2014). Utilization of SNPs plays a critical role in cost-effective and rapid MAS programs, especially in breeding using modern elite varieties that are highly genetically related to each other, e.g., *Glycine max* (soybean), and show genetic bottlenecks during domestica-

tion (Hyten *et al.* 2006). SNPs are grouped into four classes: class I SNPs are non-complementary transitions, e.g., cytosine (C) to thymine (T) and guanine (G) to adenine (A); class II SNPs are non-complementary transversions, e.g., C to A and G to T; class III SNPs are polymorphisms of complementary nucleotides with three hydrogen bonds, e.g., G to C; and class IV SNPs are polymorphisms of complementary nucleotides with two hydrogen bonds, e.g., A to T. In many organisms, including *G. max* and *Oryza sativa* (rice), class I and II SNPs are quite frequent, whereas class III and IV SNPs are relatively rare.

Highly multiplexed genotyping platforms for SNPs, such as Infinium BeadChip assay (Illumina, Inc., San Diego, CA, USA), MASS-array systems (Sequenom, Inc., San Diego, CA, USA), and high-throughput sequencers (e.g., HiSeq, Illumina; Ion Torrent, Thermo Fisher Scientific Inc., Waltham, MA, USA; and RSII, Pacific Biosciences, Menlo Park, CA, USA), are particularly useful for genotyping a huge number of loci in a few samples; however, due to their excessive power and cost, they are not ideal for genotyping a few loci in a huge number of samples. A well-designed experimental plan that considers cost-effectiveness is therefore required prior to such large-scale genotyping. In

contrast, in the average laboratory, for practical reasons, clinical diagnosis and agricultural breeding studies require rapid and cheap uniplex SNP genotyping methods at a small number of loci. Uniplex SNP genotyping platforms such as TaqMan (Thermo Fisher Scientific) and Molecular Beacon assay (Premier Biosoft International, Inc., Palo Alto, CA, USA; Szemes and Schoen 2003) require fluorescence-labeled probes to be prepared for each target locus. Another uniplex SNP genotyping method, allele-specific polymerase chain reaction (ASPCR; Ugozzoli and Wallace 1991), and its derivatives have been applied in several procedures: the Mismatch Amplification Mutation Assay (MAMA; Cha *et al.* 1992), Amplification-Refractory Mutation System-PCR (ARMS-PCR; Ye *et al.* 2001), Kompetitive Allele Specific PCR (KASP; Semagn *et al.* 2014), and Tm-shift for discrimination of SNPs or wildtype/mutant using two kinds of allele-specific oligos (ASO1 and ASO2; Wang *et al.* 2005). Since ASPCR is based on the difference in amplification efficiency of a perfectly matched and mismatched primer to template DNA, amplification from mismatched priming is not completely avoided (Kwok *et al.* 1990, Liu *et al.* 2012) even with optimization of annealing temperature by use of gradient PCR (Tortajada-Genaro *et al.* 2017); this frequently causes pseudo-positive detection of mismatched alleles. Introduction of an artificial mismatch within two to five bases closest to the 3′ end of the primers can produce more

accurate genotyping (Cha *et al.* 1992, Hayashi *et al.* 2004, Kwok *et al.* 1990, Liu *et al.* 2012); however, the best patterns of introduced mismatch bases near the 3′ end of primers have not been quantitatively predicted based on thermodynamics so far.

Another genotyping method, high-resolution melting (HRM), has been used to detect SNPs or mutations within target PCR fragments (**Fig. 1A**; Vossen *et al.* 2009). HRM uses a double-stranded DNA (dsDNA)–specific fluorescent dye to monitor decreases in fluorescence signals due to thermal dissociation of dsDNA to single-stranded DNA (ssDNA) during highly regulated temperature increases. Large differences between thermal dissociation of two dsDNA fragments containing single nucleotide substitutions, insertions, or deletions allows more accurate discrimination of genotypes in HRM analysis. The most important parameter to predict thermodynamics of dsDNA-to-ssDNA conversion is Tm value (melting temperature calculated from the above decreases in fluorescence signals). Tm values are determined mainly by the number of hydrogen bonds at each Watson–Crick base pairing and the base-stacking interaction between two base pairs of nearest-neighboring (NN) nucleotides (Crothers and Zimm 1964, SantaLucia 1998). The NN model predicts Tm based on two independent parameters: enthalpy ($\Delta H°$) and entropy ($\Delta S°$) at two base pairs of neighbor nucleotides as follows:
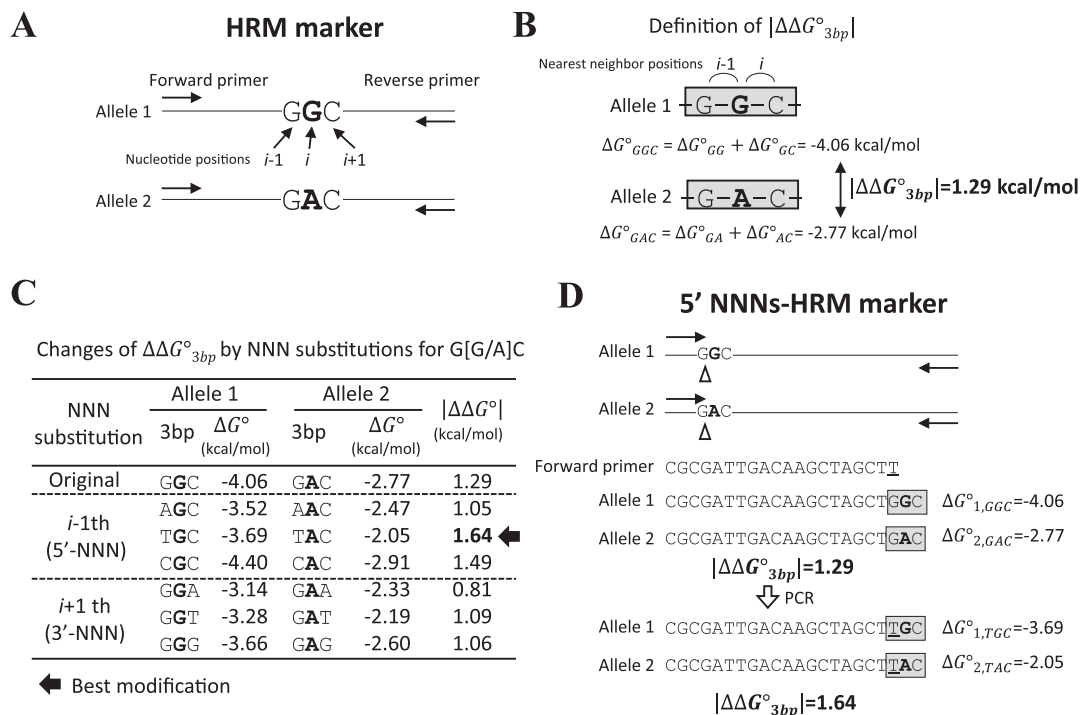
Fig. 1. Example calculation of the $|\Delta\Delta G°_{3bp}|$ index and an approach to maximize the $|\Delta\Delta G°_{3bp}|$ for the nearest neighbor nucleotide substitution in the forward primer. (A) Primer locations of an HRM marker for amplification of dsDNA harboring a class I SNP A/G at position *i*. (B) Definition of $|\Delta\Delta G°_{3bp}|$. (C) Combination of the nearest neighbor nucleotide and SNP giving highest $|\Delta\Delta G°_{3bp}|$ for class I SNPs A/G. Introduction of T nucleotide at the 5′-nearest neighbor nucleotide flanking the SNP site is the best modification. $|\Delta\Delta G°_{3bp}|$ scores of all other combinations of 3-bp nucleotides are listed in **Supplemental Table 3**. (D) Site-directed mutagenesis by PCR for maximization of $|\Delta\Delta G°_{3bp}|$.

$$Tm = \frac{\Delta H^\circ_{Total}}{R\ln(C/4) + \Delta S^\circ_{Total}} \qquad (1)$$

where R is the gas constant (1.987 cal/K mol), C is total oligonucleotide concentration ($C/4$ is used because it is assumed that the two single-stranded DNAs resulting from denaturation of each dsDNA fragment are of equal concentration; SantaLucia 1998), T is Kelvin temperature, and ΔH°_{Total} and ΔS°_{Total} of a single dsDNA with n base pairs are calculated by summing at the i-th neighboring nucleotide:

$$\Delta H^\circ_{Total} = \sum_{i=1}^{n-1} \Delta H^\circ_i \text{ and } \Delta S^\circ_{Total} = \sum_{i=1}^{n-1} \Delta S^\circ_i.$$

Tm is also predicted based on Gibbs free energy (ΔG°) as follows:

$$Tm = \frac{\Delta G^\circ_{Total}}{R\ln(C/4)} \qquad (2)$$

where $\Delta G^\circ_{Total} = \sum_{i=1}^{n-1} \Delta G^\circ_i$.

Equation (2) can be derived from Equation (1) through the following relationship:

$$\Delta G^\circ_{Total} = \Delta H^\circ_{Total} - T\Delta S^\circ_{Total} = -RT\ln(K) \qquad (3)$$

where K is the equilibrium constant for DNA fragments (ratio of single-stranded to double-stranded DNA in solution).

Although there have been some reports of the use of HRM markers for genetic fingerprinting, gene mapping, and MAS in plants (Simko 2016), application of HRM analysis to plant research has been limited. Therefore, we expect that the rational design of HRM markers will promote the use of SNPs in a wide range of applications. The aim of the current study was to use the NN model to design reliable HRM markers to discriminate homozygous alleles, especially those containing a SNP within a PCR amplicon.

We proposed some evaluation indexes calculated based on ΔG°, ΔH°, and ΔS° to easily select SNPs with a large difference in predicted Tm values between PCR fragments harboring a SNP, and we evaluated the correlations between these evaluation indexes and the observed difference in Tm values (ΔTm_{obs}) between two PCR fragments (experiment [Exp.] 1). We then designed primers that had a nucleotide substitution compared to the original genome sequence to enlarge the difference in predicted ΔG°, and we evaluated the effects of the nucleotide substitution (Exp. 2) on the difference in Tm values between the resultant two PCR fragments. We also evaluated the correlation between PCR product size and ΔTm_{obs} (Exp. 3). Finally, we evaluated the usefulness of the obtained HRM markers for linkage analysis and for more genetically diversified samples (Exp. 4). Our Perl scripts for the design of primers for HRM markers are freely available for plant breeding and research studies.

Materials and Methods

Evaluation indexes for predicting the difference between Tm values of two PCR fragments

We assume two dsDNA fragments of n bp, dsDNA 1 and 2, derived from alleles 1 and 2, respectively, which share the same sequence except for one nucleotide (1 SNP) at position i (**Fig. 1A**). The dsDNA fragments are at concentration C in the reaction mixture and have Tm values of Tm_1 and Tm_2, respectively. According to equation (1), the absolute value of $\Delta Tm = |Tm_1 - Tm_2|$ is defined as follows:

$$\Delta Tm = \left| \frac{\Delta H^\circ_{1,Total}}{R\ln(C/4) + \Delta S^\circ_{1,Total}} - \frac{\Delta H^\circ_{2,Total}}{R\ln(C/4) + \Delta S^\circ_{2,Total}} \right|$$

$$(4)$$

where total enthalpies and entropies of the dsDNA 1 and dsDNA 2 are $\Delta H^\circ_{1,Total}$ and $\Delta H^\circ_{2,Total}$, and $\Delta S^\circ_{1,Total}$ and $\Delta S^\circ_{2,Total}$, respectively. The entropy correction for salt concentration such as [Na$^+$] in reaction mixtures is given in SantaLucia (1998) by the equation $\Delta S^\circ([\text{Na}^+]) = \Delta S^\circ(1 \text{ M NaCl}) + 0.368 \times N \times \ln[\text{Na}^+]$. Since accounting for the salt concentration did not improve the ΔTm prediction, we used the unified oligonucleotide ΔH° and ΔS° NN parameters in 1 M NaCl from SantaLucia (1998).

According to equation (2), we can also estimate ΔTm as follows:

$$\Delta Tm = \left| \frac{\Delta G^\circ_{1,Total}}{R\ln(C/4)} - \frac{\Delta G^\circ_{2,Total}}{R\ln(C/4)} \right|$$

$$= \left| \frac{(\Delta H^\circ_{1,Total} - Tm_1\Delta S_{1,Total}) - (\Delta H^\circ_{2,Total} - Tm_2\Delta S_{2,Total})}{R\ln(C/4)} \right|$$

$$(5)$$

When we assume $Tm_1 = Tm_2 + \Delta Tm$,

$$\Delta Tm =$$
$$\left| \frac{(\Delta H^\circ_{1,Total} - Tm_2\Delta S^\circ_{1,Total}) - (\Delta H^\circ_{2,Total} - Tm_2\Delta S^\circ_{2,Total})}{R\ln(C/4) + \Delta S^\circ_{1,Total}} \right|$$

$$\Delta Tm =$$
$$\left| \frac{(\Delta H^\circ_{1,Total} - \Delta S^\circ_{2,Total}) - Tm_2(\Delta H^\circ_{1,Total} - \Delta S^\circ_{2,Total})}{R\ln(C/4) + \Delta S^\circ_{1,Total}} \right|$$

$$(6)$$

We defined the ΔH° and ΔS° of NN between the $(i-1)$-th and i-th nucleotides as $\Delta H^\circ_{NN=i-1}$ and $\Delta S^\circ_{NN=i-1}$, and that between the i-th and $(i+1)$-th nucleotides as $\Delta H^\circ_{NN=i}$ and $\Delta S^\circ_{NN=i}$ (**Fig. 1B**). We defined ΔH°, ΔS°, and ΔG° of NN at 3-bp ($i-1$, i, and $i+1$) for dsDNA s ($s = \{1, 2\}$) (**Fig. 1B**) as follows:

$$\Delta H^{\circ}_{s,3bp} = \Delta H^{\circ}_{s,NN=i-1} + \Delta H^{\circ}_{s,NN=i}$$
$$\Delta S^{\circ}_{s,3bp} = \Delta S^{\circ}_{s,NN=i-1} + \Delta S^{\circ}_{s,NN=i}$$
$$\Delta G^{\circ}_{s,3bp} = \Delta H^{\circ}_{s,3bp} - T\Delta S^{\circ}_{s,3bp}$$

Since the dsDNA1 and dsDNA2 are the same except for one SNP at position i,

$$\Delta H^{\circ}_{1,3bp} - \Delta H^{\circ}_{2,3bp} = \Delta H^{\circ}_{1,Total} - \Delta H^{\circ}_{2,Total}$$
$$\Delta S^{\circ}_{1,3bp} - \Delta S^{\circ}_{2,3bp} = \Delta S^{\circ}_{1,Total} - \Delta S^{\circ}_{2,Total}$$

Therefore, equation (6) can be rewritten as follows:

$$\Delta Tm = \left| \frac{(\Delta H^{\circ}_{1,3bp} - \Delta H^{\circ}_{2,3bp}) - Tm_2(\Delta S^{\circ}_{1,3bp} - \Delta S^{\circ}_{2,3bp})}{R\ln(C/4) + \Delta S^{\circ}_{1,Total}} \right|$$
$$\propto |\Delta G^{\circ}_{1,3bp} - \Delta G^{\circ}_{2,3bp}| \equiv |\Delta\Delta G^{\circ}_{3bp}|$$

(7)

In equation (7), ΔTm is proportional to $|\Delta\Delta G^{\circ}_{3bp}|$ at $T = Tm_2$ since $R\ln(C/4) + \Delta S^{\circ}_{1,Total}$ is a constant when the sequence of dsDNA 1 is known. Since $|\Delta\Delta G^{\circ}_{3bp}|$ at $T = 37°C$ is highly correlated with that at 73°C ($r = 0.99$), we used the list of NN parameters calculated with $T = 37°C$ (SantaLucia 1998) for our prediction.

For ΔTm estimation, equation (4) uses all the sequence information of the PCR fragments, whereas equation (7) uses only ΔH°_{3bp} and ΔS°_{3bp}. The $|\Delta\Delta G^{\circ}_{3bp}|$ index uses the three nucleotides around the SNP without any estimation of the ΔTm of two DNA fragments containing a single SNP. In the experiments described below, we used two indexes: (1) ΔTm_{pred}, the difference between the Tm values of PCR fragments obtained from dsDNA 1 ($Tm1_{pred}$) and 2 ($Tm2_{pred}$), was predicted using equation (4) as follows: $\Delta Tm_{pred} = |Tm1_{pred} - Tm2_{pred}|$; and (2) $|\Delta\Delta G^{\circ}_{3bp}|$ was used as prediction index for the observed differences of Tm between dsDNA 1 and dsDNA 2 (ΔTm_{obs}) in HRM analysis. Tm_{obs} (see below) was defined as the maximum peak of $-\Delta f/\Delta t$ during dissociation, and the absolute difference in the Tm_{obs} values of the two PCR fragments was used for comparison.

Plant materials

Glycine max cultivars, 'Fukuyutaka' (accession no., JP29668 in Genebank of the National Agriculture and Food Research Organization [NARO]), 'Toyoshirome' (not available in Genebank; bred at the Kyushu Okinawa Agricultural Research Center, NARO), and Chinese *G. max* landrace 'Aokimame' (accession no., JP28298) were used as parental lines. The F$_7$ population containing 96 plants derived from a cross between 'Fukuyutaka' and 'Aokimame' (hereafter, AFF$_7$) was used as the mapping population. *O. sativa* Japonica rice cultivar 'Nipponbare' and a wild rice AA-genome species *Oryza barthii* (accession no., W0652) were used. W0652 was kindly provided by the National Institute of Genetics, Mishima, Japan.

SNP data

To develop the HRM DNA markers, we used two sets of soybean SNP data obtained with restriction-site-associated DNA sequencing (RAD-seq) analysis: the first set was between 'Aokimame' and 'Fukuyutaka' with *Hae*III digestion (hereafter, AF SNPs); and the second set was between 'Toyoshirome' and 'Fukuyutaka' with *Pst*I–*Msp*I–*Cvi*QI digestion (TF SNPs). Our previous study (Watanabe *et al.* 2017) shows details of the RAD-seq analyses; lists of SNPs identified with high confidence (vcf format) containing information about 8372 loci for AF SNPs and 6142 loci for TF SNPs were used as input files. SNP data between 'Nipponbare' and W0652 (hereafter, SB SNPs) were obtained from variant analysis of whole genome sequences of W0652 and the reference sequence of 'Nipponbare' (Os-Nipponbare-Reference-IRGSP-1.0) by using Burrows–Wheeler Aligner (Li and Durbin 2009) and Samtools mpile-up version 1.1 (Li *et al.* 2009) with a filtering quality score of ≥100. Variant whole genome sequences of W0652 were downloaded from DDBJ sequence archives (accession no., DRP003801).

Primer design

In Exp. 1, we selected 48 HRM primer pairs for AF SNPs by 1) retrieving 201-bp genomic sequences around SNPs (SNP located at the center of sequence) from *G. max* reference genome sequence (cultivar 'Williams82', Glyma1.0 version); 2) designing primers with the stand-alone Primer 3 program (Untergasser *et al.* 2012) with maximum length of PCR product set to 80 bp; and 3) conducting BLASTN searches to count the number of loci showing identical sequence with the designed primer sequence in the *G. max* genome (Camacho *et al.* 2009). We chose primer pairs in which one primer matched a single identical sequence in the soybean genome database and the other matched two or less identical sequences in the database. After designing the primers, we calculated two evaluation indexes, ΔTm_{pred} and $|\Delta\Delta G^{\circ}_{3bp}|$, for each predicted PCR fragment. These steps for designing primers for each SNP were conducted automatically with Perl scripts (we provide a flow chart and all scripts used in this study as **Supplemental Fig. 1** and **Supplemental Text 1**). We then calculated the Pearson's correlation coefficient between each of these proposed indexes and the ΔTm_{obs}.

In Exp. 2, we designed HRM primers with or without consideration of nearest neighboring nucleotide (NNN) substitution; the resultant primers were termed 'NNNs-HRM markers' and 'non-NNNs-HRM markers', respectively. Taking into account the nucleotides at positions $i-1$ and $i+1$, where i is the position of the SNP, we identified the two-nucleotide combination giving the highest $|\Delta\Delta G^{\circ}_{3bp}|$ index (**Fig. 1C**); using this information, we designed a forward primer for site-directed mutagenesis to convert the 5′ or 3′-NNN of the SNP site to the optimal nucleotide (**Fig. 1D**). The reverse primer was designed by using the Primer 3 program. We designed two sets of 16 combinations

of forward and reverse primers; in the first set, forward primers contained a nucleotide substitution just upstream of the SNP (M65 to M80 in **Supplemental Table 1**), compared with the original sequence; in the other set, all primers (M49 to M64) contained sequence identical to the reference genome sequence.

In Exp. 3, we designed 24 pairs of NNNs-HRM primers producing 50–100 bp PCR products with $|\Delta\Delta G°_{3bp}|$ index thresholds of >1.43 kcal/mol from TF SNPs data sets on *G. max* chromosome 10. All primers used in this study are listed in **Supplemental Table 1**. For more genetically divergent cases, we calculated $|\Delta\Delta G°|$ by subtraction of two $\Delta G°$ scores calculated from the sequence of dsDNA 1 and dsDNA2 (**Supplemental Text 2**) according to equation (3).

DNA marker analysis

Total genomic DNAs of young trifoliate leaflets from parental lines and the AFF_7 population, and adult leaves from 'Nipponbare' and W0652 plants were isolated by using the cetyl trimethylammonium bromide method (Murray and Thompson 1980). HRM analysis was conducted using quantitative PCR equipment (LightCycler 96 System; Roche Diagnostics K.K., Minato, Tokyo, Japan). PCR mixtures contained 1× PCR buffer (400 mM Tricine-KOH [pH 9.2], 150 mM Sodium acetate [pH 6.0], 35 mM Magnesium acetate, 37.5 μg/mL bovine serum albumin, 0.5% (v/v) Nonidet P-40, 0.5% (v/v) Tween 20), 0.2 mM each dNTP, 20–30 ng genomic DNA, 0.25 units of home-made recombinant *Thermus aquaticus* (*Taq*) DNA polymerase (activity roughly equivalent to 0.25 units of *Takara Ex Taq* Hot Start Version from Takara, Kusatsu, Shiga, Japan), 0.5× fluorescent dye EvaGreen (Biotium, Inc., Fremont, CA, USA), and 0.1 pmol each HRM primer. To prepare the home-made recombinant *Taq* DNA polymerase, we used the pET-vector protein expression system (Merck Japan, Meguro, Tokyo, Japan) as described briefly below. The gene encoding DNA polymerase I (accession number J04639) of *Taq* strain JCM 10724 (provided by National Bio Resource Project) was cloned into pET-28a (+); the resulting vector was used to transform *Escherichia coli* strain XL21-AI (Thermo Fisher Scientific). Production of recombinant *Taq* polymerase was induced according to the manufacturer's protocols. The pellet of *E. coli* cells from a 1 L culture with 0.8 OD was suspended in 30 mL buffer B (20 mM Tris-HCl [pH 7.2], 500 mM NaCl, 5 mM imidazole, 0.1% Triton X-100) and homogenized with an Ultrasonic Generator 201M (Kubota, Hongo, Tokyo, Japan). The cell lysate was incubated at 75°C for 30 min and then centrifuged at 16,000 ×g for 20 min to eliminate insoluble proteins. Polyhistidine-tag fusion recombinant *Taq* polymerase was purified with 4 mL Ni Sepharose 6 Fast Flow (GE Healthcare Japan, Shinjuku, Tokyo, Japan) according to the manufacturer's instructions. After elution of the recombinant *Taq* polymerase with buffer B containing 100 mM imidazole (instead of 5 mM imidazole), the buffer was substituted with the storage buffer (20 mM Tris-HCl [pH 8.0], 100 mM KCl, 0.1 mM EDTA,

1 mM dithiothreitol, 0.5% (v/v) Nonidet-P40, and 0.5% (v/v) Tween 20) by dialysis. Enzyme activity was roughly estimated by comparing the PCR results of commercial *Taq* DNA polymerase (*Takara Ex Taq* Hot Start Version) with that of a dilution series of purified recombinant *Taq* DNA polymerase. PCR conditions were pre-incubation at 95°C for 5 min, and 40 cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 20 s; melting analysis was then performed at 95°C for 10 s, 65°C for 1 min, and then increasing temperatures to 97°C with a slope of 0.1°C/s with continuous measurement of the fluorescence of the PCR product. The reproducibility of genotypes obtained by HRM analysis was evaluated with 7 NNNs-HRM markers (AF_M65, AF_M67, AF_M69, AF_M73, AF_M75, AF_M76, and AF_M77) with the same conditions mentioned above except that the DNA polymerase and PCR buffer were provided by Takara. The post-run analysis of HRM data was conducted with software provided by the LightCycler 96 System, and the Tm_{obs} of each genotype was calculated as the peak value of $-\Delta f/\Delta t$, where Δf is change in fluorescence and Δt is change in temperature. Single band amplification for all PCR amplicons was confirmed with 12% polyacrylamide gel electrophoresis.

Data analysis

Linkage map analysis of the AFF_7 population was performed with Antmap software with default parameters for recombinant inbred lines (Iwata and Ninomiya 2006). Correlation analysis and paired *t*-tests (alternative hypotheses, Tm_{obs} of non-NNNs HRM markers < that of NNNs-HRM markers) were conducted with R (ver. 3.3.1) software (R Development Core Team 2008).

Results

Correlation between $|\Delta\Delta G°_{3bp}|$ and observed Tm

We investigated SNP data sets obtained from intraspecific comparison of *G. max* data (AF SNPs and TF SNPs) and from interspecific comparison of genus *Oryza* data (SB SNPs) (**Table 1**). The class I and II SNPs displayed the first and second highest frequencies in all data sets, whereas class III and IV SNPs showed the lowest frequencies; only class I and II SNPs were sufficiently abundant to be used as HRM markers.

From the 48 primer pairs designed from the AF SNP dataset (see Materials and Methods), we prepared 38 (79.2%) primer pairs that each produced single-band amplification of a different fragment that was polymorphic

Table 1. Frequencies of SNP classes in soybean and rice SNP datasets

Dataset	Genus	SNP class				Total No. of SNPs
		I	II	III	IV	
AF	*Glycine*	64.2%	17.7%	6.2%	16.4%	8372
TF	*Glycine*	56.9%	19.8%	6.9%	11.3%	6142
SB	*Oryza*	71.3%	15.1%	5.8%	7.8%	4138822
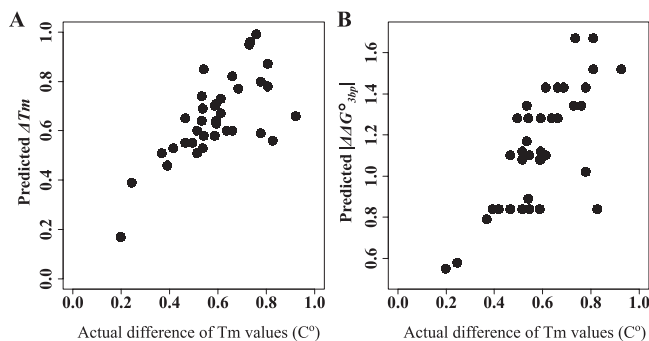
Fig. 2. Correlation between ΔTm_{obs} and evaluation indexes. (A) Correlation between ΔTm_{obs} and ΔTm_{pred} ($r = 0.648$). (B) Correlation between ΔTm_{obs} and $|\Delta\Delta G°_{3bp}|$ ($r = 0.647$).

between the parents, 'Aokimame' and 'Fukuyutaka' (Exp. 1). For each primer pair, absolute differences in the Tm_{obs} values of the PCR products obtained from two varieties (i.e., $|Tm1_{obs} - Tm2_{obs}| = \Delta Tm_{obs}$) were calculated (**Supplemental Table 2**). The ΔTm_{obs} and ΔTm_{pred} values (predicted using whole sequences of dsDNA fragments) were positively correlated ($r = 0.648$, $P < 0.01$; **Fig. 2A**), as were the ΔTm_{obs} and $|\Delta\Delta G°_{3bp}|$ scores ($r = 0.647$, $P < 0.01$; **Fig. 2B**). Because the correlation coefficients for the above relationships were almost identical, these results demonstrate that $|\Delta\Delta G°_{3bp}|$ score has an equivalent prediction power to ΔTm_{pred}. Selection for primer pairs with the highest $|\Delta\Delta G°_{3bp}|$ scores provides us a selection criterion for those that would detect the largest differences in Tm_{obs} values between two variants at an HRM marker.

Optimization of HRM markers

Since larger ΔTm_{obs} values in HRM analysis reduce genotyping errors due to factors such as variance in reaction volumes and amplification efficiency across wells, we attempted to optimize the design of the HRM markers for more reliable genotyping by conducting site-directed mutagenesis of the 5′ or 3′ NNN of the SNP with forward primer (**Fig. 1C**, **1D**).

To determine which, if any, 5′ or 3′ NNNs maximize $|\Delta\Delta G°_{3bp}|$, we calculated $|\Delta\Delta G°_{3bp}|$ for various types of SNPs (**Figs. 1C**, **1D**, **3**). For C/T (Y) SNPs (an example of class I SNPs), $|\Delta\Delta G°_{3bp}|$ was maximized by the presence of G or A nucleotide at the 5′ or 3′ NNN (upstream or downstream of SNP, respectively), respectively; for G/A (R) SNPs, $|\Delta\Delta G°_{3bp}|$ was maximized by the presence of T or C nucleotide at the 5′ or 3′ NNN, respectively (**Fig. 3**). Among class II SNPs, G nucleotides at 5′ or 3′ NNN for C/A (M) SNPs and C nucleotides at 5′ or 3′ NNN for G/T (K) SNPs provided maximum $|\Delta\Delta G°_{3bp}|$. These data suggest that nucleotide substitutions immediately upstream or downstream of the SNP are critical determinants of $|\Delta\Delta G°_{3bp}|$ at class I and class II SNP sites. However, neither 5′ nor 3′ NNNs at class III and IV SNP sites showed a clear trend for maximizing $|\Delta\Delta G°_{3bp}|$. In **Supplemental Table 3**, we provide a quick reference matrix for practical guidelines to select suitable NNNs for any SNP.

We then developed a pipeline to automatically design primers to maximize $|\Delta\Delta G°_{3bp}|$ scores in HRM analysis. Of the 32 primer pairs, which comprised NNNs-HRM markers (M65 to M80, **Supplemental Table 1**) and their original non-NNNs-HRM markers (M49 to M64) at 16 SNP loci from the AF SNP dataset, 24 (75.0%) showed single amplification and polymorphism between parental lines. For example, the non-NNNs-HRM marker, M60, and the NNNs-HRM marker, M76, amplified genomic sequences containing the class II SNP (5′-AKT-3′) in DNA strands identical to reference sequence of *G. max* at 42.7 Mbp on chromosome 1 (Gm01). The results in **Fig. 3** suggest that, to maximize $|\Delta\Delta G°_{3bp}|$, the A nucleotide at i-1th position should be converted to C by PCR-mediated site-directed mutagenesis. After this substitution, the $|\Delta\Delta G°_{3bp}|$ scores improved from 0.84 to 1.34 kcal/mol and the ΔTm_{obs} values improved from 0.53 to 0.64°C (**Supplemental Table 4**, **Supplemental Fig. 2**). Even for the non-NNNs-HRM marker, M64, an example of a class IV SNP (A/T), which had an extremely low $|\Delta\Delta G°_{3bp}|$ of 0.05 kcal/mol, the NNN substitution improved ΔTm_{obs} values and provided successful genotyping of the AFF$_7$ population (marker M80_AF, **Supplemental Table 4**, **Supplemental Fig. 2**; see section below). Consequently, the average ΔTm_{obs} values of non-NNNs-HRM and NNNs-HRM markers were 0.43°C and 0.63°C, respectively (**Supplemental Table 4**), and the effect of NNN substitution in primer to increase the ΔTm_{obs} values was significant (11 combinations of non-NNNs and NNNs-HRM markers were used in **Supplemental Table 4**; df = 10, t value = 7.6, $P < 0.01$, paired t-test), demonstrating that the concept of $|\Delta\Delta G°_{3bp}|$ allowed rational design of HRM markers via NNN substitution.

Next, we investigated the correlation between PCR product size and ΔTm_{obs}. We designed 24 NNNs-HRM markers that displayed $|\Delta\Delta G°_{3bp}|$ of >1.43 kcal/mol (**Supplemental Table 5**) with PCR product sizes of 54–100 bp, and then used them for HRM analysis. Out of 24 primers, 3 showed no polymorphism between parents and another 3 showed a dominant segregation pattern; these primers were not analyzed further. The remaining 18 primers (75%) showed co-dominant segregation among three genotypes. The ΔTm_{obs} values between parental lines (mean, 0.74°C; **Supplemental Table 5**) were significantly negatively correlated with the length of PCR product ($r = -0.75$, $P < 0.01$; **Fig. 4**). This result indicates that not only maximization of $|\Delta\Delta G°_{3bp}|$ but also minimization of PCR product length increased ΔTm_{obs}. We also evaluated reproducibility of genotyping by using different commercial *Taq* polymerases with 7 NNNs-HRM markers: 653 out of 672 genotypes (97.2%) showed identical genotype results regardless of the source of *Taq* polymerase (**Supplemental Fig. 3**, **Supplemental Table 6**).

A drawback of thresholding by $|\Delta\Delta G°_{3bp}|$ and limiting PCR product length is the decreased flexibility in PCR primer design. Without consideration of NNN substitution, when the ΔTm_{pred} of PCR products was thresholded at
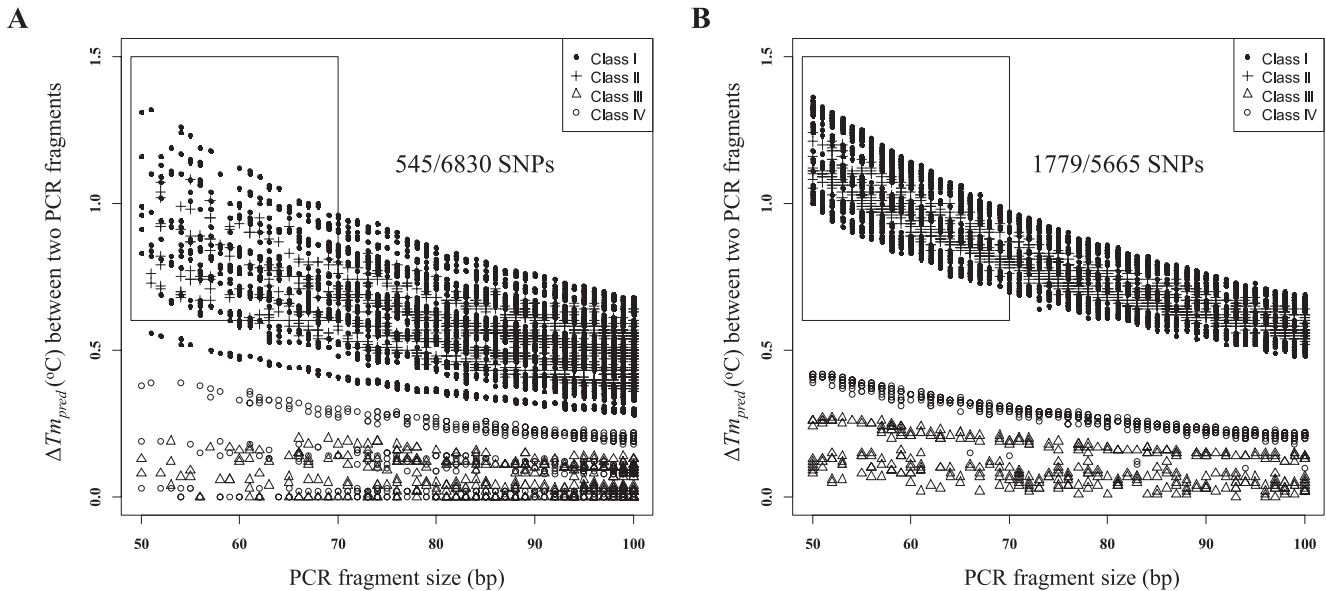
Nucleotide substitution of 5′ NNN   Nucleotide substitution of 3′ NNN
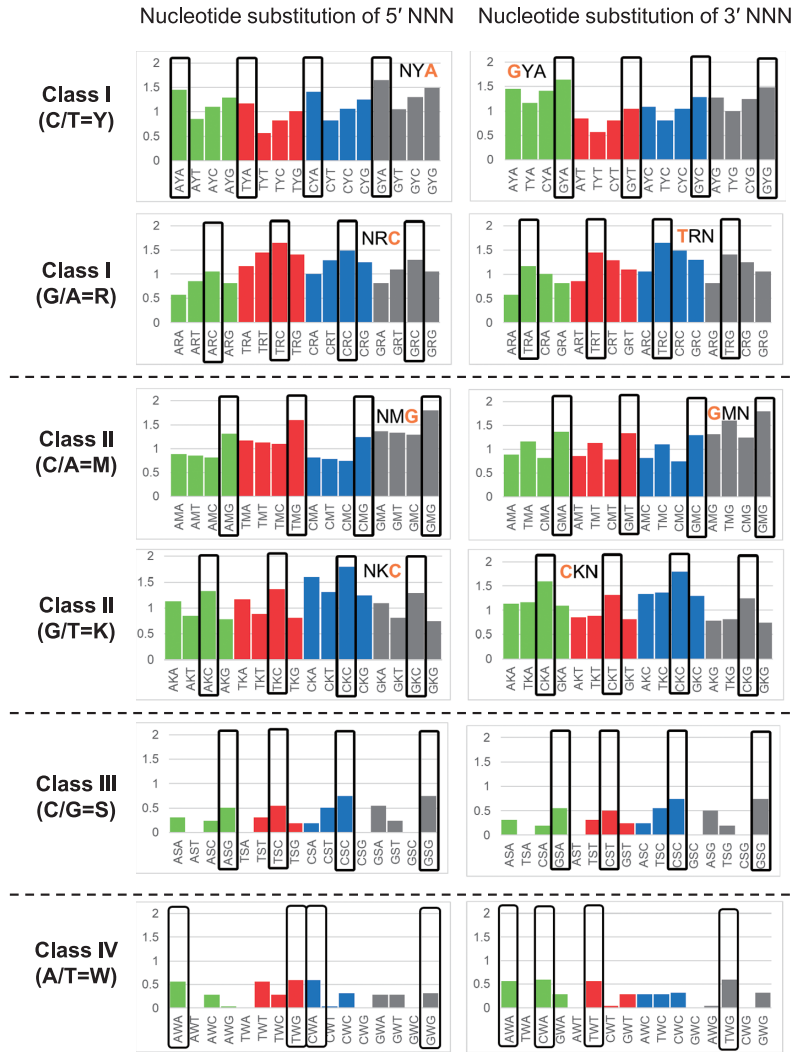
Fig. 3. |ΔΔG°$_{3bp}$| for class I, II, III, and IV SNPs sorted by 5′ or 3′ nearest neighbor nucleotides (NNN) relative to the SNP. In the groups of three nucleotides, first the 5′ NNNs are arranged A, T, C, G, and then the 3′ NNNs are arranged A, T, C, G. In each panel, SNPs with the highest |ΔΔG°$_{3bp}$| are marked by rectangles.
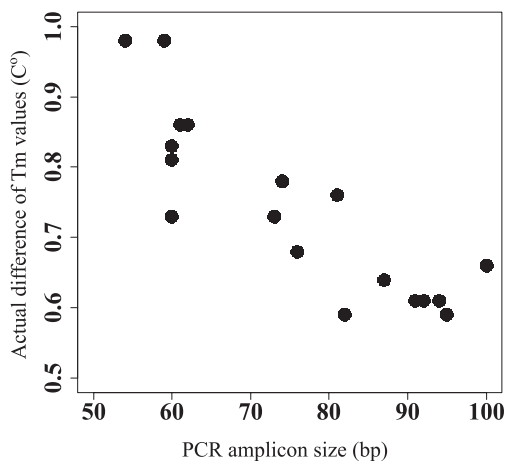
Fig. 4. Correlation between the ΔTm_{obs} values for NNNs-HRM markers and size of PCR fragment.

>0.6°C and the PCR product length was limited to <70 bp in the AF SNP dataset, we succeeded in designing primers at only 545 (8.0%) of 6830 loci (**Fig. 5A**). However, under the same conditions, consideration of NNN substitution greatly improved the percentage of primer design such that it increased to 1779 (31.4%) of 5665 loci (**Fig. 5B**). Primer3_ core was used to calculate the melting temperatures of oligos and the tendencies of oligos to form hairpins or dimers (Untergasser *et al.* 2012). We fixed the position of primer including nucleotide substitution, and optimization of the substituted nucleotide decreased the variations of nucleotides and increased the percentage of NNN markers having hairpin structures (e.g., GC and CG) at the 3′ end of the sequence: 3.5% in non-NNNs and 13.2% in NNNs for AF SNP (data not shown). Optimization of NNNs is the likely cause of the low number of successfully designed primer pairs (6860 to 5665 loci). However, 5269 SNP loci were commonly used to design non-NNNs and NNNs-HRM

A

545/6830 SNPs

B

1779/5665 SNPs

Fig. 5. Effect of optimization for NNN substitution flanking SNPs. Scatter plot of ΔTm_{pred} values of two PCR fragments and PCR product size without nucleotide substitution (A) and with nucleotide substitution (B). All AF SNPs were used to design primer pairs and simulate ΔTm_{pred} of two PCR amplicons containing a SNP. The square boxes include the NNNs-HRM markers with the ΔTm_{pred} of >0.6 and PCR product size of <70 bp. The numbers of markers (SNPs) in the square boxes are indicated in the figure.

markers, and even if we use a threshold for primer inclusion of maximum number of loci for a single primer of ≤2 according to BLASTN search, close to 1000 NNNs-HRM markers would be available for analysis with AF and TF SNPs (**Supplemental Tables 7**, **8**).

Validation of genotypes obtained with NNNs-HRM markers

We further validated that the NNNs-HRM markers mapped at precise loci on the soybean genetic linkage map. We analyzed the genotype of AFF₇ plants by using ten combinations of primer pairs and found that non-NNNs-HRM and NNNs-HRM markers targeting the same SNP mapped on the same genetic locus (**Fig. 6A**). Moreover, the relationship between genetic and physical positions showed a typical sigmoid-like pattern, with recombination suppressed around the pericentromeric region (**Fig. 6B**). These results indicate that use of NNNs-HRM markers did not affect the determination of genetic location or genotype.

Availability of HRM markers for more genetically divergent cases

We tried to extend application of $|\Delta\Delta G°|$ for estimation of ΔTm_{obs} to PCR fragments carrying more than one SNP derived from more diverged species: the cultivated rice species *O. sativa* L. and wild species *O. barthii*. In a plot of $|\Delta\Delta G°|$ against ΔTm_{obs} (**Fig. 7**, **Supplemental Text 2**), the values of the PCR products containing single SNPs were close to each other on both the $|\Delta\Delta G°|$ and ΔTm_{obs} axes, whereas the values of ones containing more than one SNP were scattered widely. Although this observation is qualitative, $|\Delta\Delta G°|$ prediction accuracy for ΔTm_{obs} may be lower when the two PCR products contain more than one SNP.

A **B**
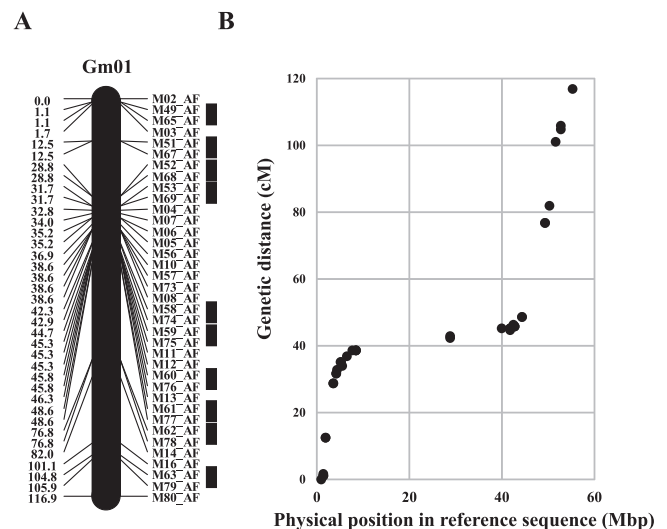
Fig. 6. Genetic linkage map of soybean chromosome 1 (Gm01) constructed by using the AFF₇ population with non-NNNs-HRM markers and NNNs-HRM markers. Non-NNNs-HRM and NNNs-HRM markers detecting the same SNP are arranged next to each other and are indicated with vertical black rectangles, e.g., M49_AF and M65_AF; numbers at the left of the map indicate genetic distance (cM) (A). A scatter plot of the physical positions versus map distances of these HRM markers is shown (B).

Discussion

Current protocols for choosing which SNPs will be more successful for genotyping in HRM analysis are based on whether the SNPs are class I, II, III, or IV. Here, we suggest a novel criterion that considers the 5′- and 3′- nucleotides
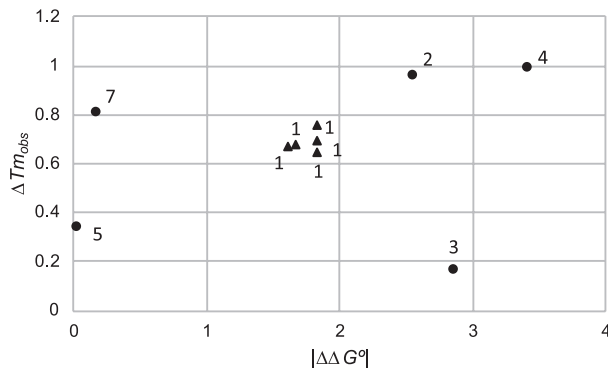
Fig. 7. Scatter plots of $|\Delta\Delta G°|$ versus ΔTm_{obs} for PCR fragments from *Oryza sativa* and *O. barthii*. Circles and triangles represent PCR fragments containing single SNPs and more than one SNP, respectively. Digits represent numbers of SNPs in the PCR fragments.

flanking the SNP sites, and we make available a Perl script pipeline for the automatic design of a massive number of suitable HRM markers from SNP datasets. For electrophoresis-based DNA markers such as simple sequence repeats (SSR) and sequence-tagged site (STS) markers, large differences of dsDNA length derived from two alleles enable more accurate genotyping. Similarly, as large as possible Tm differences between dsDNAs derived from two alleles are required in HRM analysis. Because two dsDNA fragments containing SNPs optimally have a large ΔTm_{obs}, an easy and practical technique to predict ΔTm_{obs} and thereby select reliable HRM markers would have wide applications in genotyping. Here, we compared two indexes for prediction of Tm_{obs}, i.e., $|\Delta\Delta G°_{3bp}|$ and ΔTm_{pred}, and we found that both indexes had equivalent power to predict ΔTm_{pred}. Because calculation of $\Delta\Delta G°_{3bp}$ (which requires addition of thermodynamic parameters for three nucleotides) is easier than calculation of ΔTm_{pred} (which requires analysis of the entire nucleotide sequence of the PCR amplicon), we chose to use $|\Delta\Delta G°_{3bp}|$ to select SNP sites available for genotyping or to predict the best modification of the 5′- or 3′-NNN at a given SNP site in further experiments. Our empirical observation revealed that a ΔTm value >0.5°C guaranteed accurate genotyping using the Roche LightCycler 96 system (data not shown). All of the Class I and II NNNs-HRM markers with an amplicon size of <90 bp achieved a ΔTm_{pred} of >0.5°C (**Fig. 5B**), whereas 61.4% (3466/5647 for Class I and II) of non-NNNs-HRM markers did not (**Fig. 5A**). Furthermore, we noticed that some Class III and IV NNNs-HRM markers with ΔTm_{obs} of 0.2–0.3°C produced accurate genotypes. Using markers that satisfy the criteria of predicted $|\Delta\Delta G°_{3bp}|$ of >1.0 kcal/mol and amplicon size of <80 bp would generally guarantee ΔTm *of* >0.5°C (**Fig. 2B**). In HRM analysis, NNNs-HRM markers designed using this pipeline showed higher ΔTm_{obs} values, thereby improving the reliability of genotyping regardless of the kind of *Taq* polymerase. Moreover, considering the size of amplicon for HRM analysis also increased ΔTm_{obs} (**Fig. 4**). Many manufacturers of equipment for HRM analysis, such as Applied Biosystems

(Foster City, CA, USA), Roche, Bio-Rad Laboratories (Hercules, CA, USA), Thermo Fisher Scientific, and QIAGEN (Hilden, Germany), highly recommend using short amplicons (70–150 bp) in their application manuals. Furthermore, Vossen *et al.* (2009) targeted fragment sizes of 80–100 bp for SNP typing. Our experimental results (**Fig. 4**) and simulation data (**Fig. 5**) support their recommendation for amplicon sizes, with <80 bp being preferable for reliable genotyping. In a previous study, we proposed a method for developing dCAPS (derived cleaved amplified polymorphic sequences) markers from TF SNPs (one of the two SNP data sets used here); we found that 2000 out of 6000 SNPs could be used for dCAPS analysis with 17 inexpensive restriction enzymes (Watanabe *et al.* 2017). By using the current pipeline, we could design over 3700 candidate primers for HRM analysis with the same SNP data set (**Supplemental Table 7**), thereby increasing the number of loci available for genetic research and breeding programs. The ability to conduct HRM analysis in a single tube (i.e., without any analysis of PCR product or other equipment) would improve the effectiveness (time and labor cost) of genotyping individuals. Using DNA markers tightly linked to a target gene can enhance the accuracy of MAS, and increasing the number of DNA markers linked to a target locus can reduce the size of the donor chromosome segment and thereby reduce linkage drag (Ribaut and Hoisington 1998). In general breeding programs using bulk population breeding methods for self-pollinated crops, such as rice, wheat, and soybean, single plants are selected at the F_4 or F_5 stage. If we apply MAS to these generations, discrimination of homozygous alleles is necessary. Therefore, increasing the reliability of HRM markers by enlarging the difference of Tm values is relevant.

Designing HRM markers throughout the whole genome would facilitate gene mapping and development of genetic materials such as chromosome segment substitution lines (CSSLs). Tsuda *et al.* (2015) developed a mutagenized population with the genetic background of Japanese soybean cultivar 'Enrei', and screened many mutant lines based on phenotypes such as maturity, yield, and oil and protein contents. We will confirm the effects of each mutated gene by using segregated populations derived from crosses with genetically close cultivars, such as 'Toyoshirome' or 'Fukuyutaka'. These crosses between Japanese elite soybean lines should reduce the introduction of unfavorable traits and allow easy evaluation of the genetic effects of the mutated gene. Bulk segregant analysis (Quarrie *et al.* 1999) with HRM markers could rapidly identify a chromosome region containing a particular mutated gene. Furthermore, HRM-mapping of segregating populations can provide us with genetic markers linked to the target gene and breeding materials simultaneously. Accumulating melting curve and other data for short DNA fragments containing HRM markers under the same experimental condition could potentially be used to improve the NN parameter table of SantaLucia (1998) for more accurate estimation of DNA thermodynamics. For developing CSSLs, sets of HRM markers detecting

SNPs specific to fundamental elite lines could be used to more easily discriminate the donor chromosome, which contains novel genes originating from a genetically diverse germplasm. In rice, many CSSLs with combinations of chromosomes of Japanese rice cultivars, such as 'Taichung 65' (T65), and exotic *Oryza* species have been developed (Yasui *et al.* 2010) with SSR markers. SNPs specific to T65 could be easily used to develop genome-wide HRM markers that could be used across different donor types to monitor chromosomal segments.

Here, we used two SNP data sets that showed similar trends of naturally occurring nucleotide substitutions to each other. As in humans (Collins and Jukes 1994), in both sets, most SNPs were class I SNPs. Transition-type substitutions occur frequently because of the similarity between the chemical structures of the bases (Collins and Jukes 1994). This means that class I SNPs show redundancy in most species, which probably increases the opportunities to use reliable NNNs-HRM markers. Maintaining the reproducibility of the shape of the melting curve (Tm values) is important (Simko 2016); therefore, future work will consider the need for uniform amplification without unspecific amplicons, the type of dye, and the preparation of DNA samples of uniform concentration. HRM markers could also be extended to include polymorphisms related to insertions and deletions.

While current genotyping techniques, such as genotyping by sequencing and DNA typing arrays, produce a large amount of information on genotypes of individuals, the quick system for developing reliable DNA markers described in this study would accelerate the use of DNA markers for large populations and could also be adapted easily to other crop species to find genes or develop new genetic materials for breeding and genetic research.

Acknowledgements

Literature Cited

Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer and T.L. Madden (2009) BLAST+: architecture and applications. BMC Bioinformatics 10: 421.

Cha, R.S., H. Zarbl, P. Keohavong and W.G. Thilly (1992) Mismatch amplification mutation assay (MAMA): application to the c-H-*ras* gene. PCR Methods Appl. 2: 14–20.

Collard, B.C.Y. and D.J. Mackill (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 363: 557–572.

Collins, D.W. and T.H. Jukes (1994) Rates of transition and transversion in coding sequences since the human-rodent divergence. Genomics 20: 386–396.

Crothers, D.M. and B.H. Zimm (1964) Theory of the melting transition of synthetic polynucleotides: Evaluation of the stacking free energy. J. Mol. Biol. 9: 1–9.

Hayashi, K., N. Hashimoto, M. Daigen and I. Ashikawa (2004) Development of PCR-based SNP markers for rice blast resistance genes at the *Piz* locus. Theor. Appl. Genet. 108: 1212–1220.

Hyten, D.L., Q. Song, Y. Zhu, I.Y. Choi, R.L. Nelson, J.M. Costa, J.E. Specht, R.C. Shoemaker and P.B. Cregan (2006) Impacts of genetic bottlenecks on soybean genome diversity. Proc. Natl. Acad. Sci. USA 103: 16666–16671.

Iwata, H. and S. Ninomiya (2006) AntMap: constructing genetic linkage maps using an ant colony optimization algorithm. Breed. Sci. 56: 371–377.

Kwok, S., D.E. Kellogg, N. McKinney, D. Spasic, L. Goda, C. Levenson and J.J. Sninsky (1990) Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. 18: 999–1005.

Li, H. and R. Durbin (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Liu, J., S. Huang, M. Sun, S. Liu, Y. Liu, W. Wang, X. Zhang, H. Wang and W. Hua (2012) An improved allele-specific PCR primer design method for SNP marker analysis and its application. Plant Methods 8: 34.

Murray, M.G. and W.F. Thompson (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8: 4321–4325.

Nielsen, R., J.S. Paul, A. Albrechtsen and Y.S. Song (2011) Genotype and SNP calling from next-generation sequencing data. Nat. Rev. Genet. 12: 443–451.

Pabinger, S., A. Dander, M. Fischer, R. Snajder, M. Sperk, M. Efremova, B. Krabichler, M.R. Speicher, J. Zschocke and Z. Trajanoski (2014) A survey of tools for variant analysis of next-generation genome sequencing data. Brief. Bioinformatics 15: 256–278.

Quarrie, S.A., V. Lazic-Jancic, D. Kovacevic, A. Steed and S. Pekic (1999) Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. J. Exp. Bot. 50: 1299–1306.

R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ribaut, J.-M. and D. Hoisington (1998) Marker-assisted selection: new tools and strategies. Trends Plant Sci. 3: 236–239.

SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc. Natl. Acad. Sci. USA 95: 1460–1465.

Semagn, K., R. Babu, S. Hearne and M. Olsen (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol. Breed. 33: 1–14.

Simko, I. (2016) High-resolution DNA melting analysis in plant research. Trends Plant Sci. 21: 528–537.

Szemes, M. and C.D. Schoen (2003) Design of molecular beacons for AmpliDet RNA assay—Characterization of binding stability and

probe specificity. Anal. Biochem. 315: 189–201.

Tortajada-Genaro, L.A., R. Puchades and Á. Maquieira (2017) Primer design for SNP genotyping based on allele-specific amplification—Application to organ transplantation pharmacogenomics. J. Pharm. Biomed. Anal. 136: 14–21.

Tsuda, M., A. Kaga, T. Anai, T. Shimizu, T. Sayama, K. Takagi, K. Machita, S. Watanabe, M. Nishimura, N. Yamada *et al.* (2015) Construction of a high-density mutant library in soybean and development of a mutant retrieval method using amplicon sequencing. BMC Genomics 16: 1014.

Ugozzoli, L. and R.B. Wallace (1991) Allele-specific polymerase chain reaction. Methods 2: 42–48.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm and S.G. Rozen (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res. 40: e115.

Vossen, R.H.A.M., E. Aten, A. Roos and J.T. Den Dunnen (2009) High-resolution melting analysis (HRMA)—More than just sequence variant screening. Hum. Mutat. 30: 860–866.

Wang, J., K. Chuang, M. Ahluwalia, S. Patel, N. Umblas, D. Mirel, R. Higuchi and S. Germer (2005) High-throughput SNP genotyping by single-tube PCR with Tm-shift primers. Biotechniques 39: 885–893.

Watanabe, S., C. Tsukamoto, T. Oshita, T. Yamada, T. Anai and A. Kaga (2017) Identification of quantitative trait loci for flowering time by a combination of restriction site-associated DNA sequencing and bulked segregant analysis in soybean. Breed. Sci. 67: 277–285.

Yasui, H., Y. Yamagata and A. Yoshimura (2010) Development of chromosome segment substitution lines derived from *indica* rice donor cultivars DV85 and ARC10313 in the genetic background of *japonica* cultivar Taichung 65. Breed. Sci. 60: 620–628.

Ye, S., S. Dhillon, X. Ke, A.R. Collins and I.N.M. Day (2001) An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res. 29: e88.