Enhanced Primer Selection and Synthetic Amplicon Templates Optimize High-Resolution Melting Analysis of Single-Nucleotide Polymorphisms in a Large Population

Carol M. Bruzzone,¹ Patrick S. Tawadros,² Lisa A. Boardman,³ and Clifford J. Steer^{1,4}

Aims: High-resolution melting (HRM) screening and scanning for single-nucleotide polymorphisms (SNPs) afford the advantages of a quicker, less expensive, and less demanding option compared to other methods for sequence analysis. The evaluation of large populations of patients for multiple SNPs in a high-throughput manner is the next phase in individualized medicine. **Results:** We demonstrated that T_m profiles can be generated from gDNA samples that clearly differentiate homozygous ancestral, homozygous SNP, and heterozygous genotypes, while identifying samples of unique outcome without the cumbersome processes of normalization, temperature shifting, and difference plot generation. **Conclusions:** Through expanded primer selection criterion and inclusion of a cloning fragment length double-stranded DNA sequence-specific control template, we are now able to generate additional data via HRM melt domains that are greatly simplified, while considering both the peak melt temperature and profile.

Introduction

SINGLE-NUCLEOTIDE POLYMORPHISM (SNP) analysis has greatly improved the detection of transitions and transversions from an ancestral (A) sequence. In genomic transitions, a double ring to double ring or single ring to single ring nucleotide occurs with the loss or gain of a hydrogen bond (example, an A to G or a C to T). Transversions, also called strand swap rearrangements, are rare due to the chemistry of changing a two ring and one ring molecule and do not produce a change in hydrogen bond melting (example, a G to C change). Each person is heterozygous for ~3 million bases, has about one SNP every 1000 bases or ~3.2 million differences in our diploid genome. A homozygous SNP genotype (S/S) present in as few as 2% of the population can be expected to indicate an ~14% heterozygous (A/S) population. SNP analysis is becoming increasingly valuable as our understanding of the genome deepens and clinically relevant correlations are made that allow personalization of medical therapies. Accessing the secrets of genetic individuality will be part of moving medicine from descriptions of very low-frequency, but highly deleterious genetic changes to the study of less highly penetrant, but more prevalent SNPs that have a public health importance for a much larger population (Tricarico *et al.*, 2011).

SNP technology has evolved steadily through mobility shift studies, mass cleavage product evaluations, heterodimer

differences in chemical, conformational, and enzymatic differences, and mass spectroscopy and sequencing, to allele-specific hybridization probe methods. Each method presented challenges of labor intensity, unreliable efficiencies, complicated optimizations, and issues of sample quantity and quality (Liew *et al.*, 2004; Er and Chang, 2012). The introduction of LCGreen dye, nearly a decade ago (Wittwer *et al.*, 2003), began the era of double-stranded DNA dyes capable of differentiating homozygous A/A and S/S states and heterozygous (A/S) samples. The exquisite capability of the intercalation dyes in differentiation of wild-type mutations is highly regarded (Lin *et al.*, 2010). Double-stranded dyes that do not interfere with polymerase chain reaction (PCR) amplification in saturation are necessary to eliminate dye jumping. To make transversions obvious, the nearest neighbor sequence changes must influence melt temperature through locking chemistry of the dye.

Current data handling software can generate an absolute value of resolution difference (Laurie *et al.*, 2007). However, using this approach in analyzing data is blind to the shape of curve variations that represent potential mutations (Lin *et al.*, 2010). Each method and the current software have limitations to simple identification in a high-throughput setting. In this report, we describe the development of a method capable of differentiating homozygous cases and identifying heterozygote samples for transitions and transversions in a single

Departments of ¹Medicine and ²Surgery, University of Minnesota Medical School, Minneapolis, Minnesota.

³Department of Medicine, Mayo Clinic, Rochester, Minnesota.

⁴Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, Minnesota.

straightforward protocol without the need for post-amplification measures.

Materials and Methods

SNP targets

Seven transitions and one transversion associated with DNA repair, apoptosis, cell cycle, immune response, oxidative stress, or metabolism were selected for the study.

Synthetic amplicon targets

The genomic DNA sequence for the eight targets was retrieved and targets were synthesized based on NCBI consensus. Sequence-verified genomic blocks, gBlocks (Integrated DNA Technologies), were used as PCR targets to verify T_m differences in homozygous states during assay optimization and to verify the identity of homozygous peaks in the transversion. The higher melt temperature sequence-specific templates were also titrated to a second derivative maximum approximately equal to the patient population and were added as a target in spiked assays.

Primer design

Primers that flanked the SNP of interest and did not anneal to the SNP itself were designed using Primer3 (<http://fokker.wi.mit.edu/primer3>). Two or three additional candidates were hand selected for both the left and right side to generate amplicons of ~100 bp containing a single mismatch site. The primer candidates for each SNP were cross tested such that a SNP with three left and three right primers was evaluated for nine possible combinations using a multicriteria protocol. Primers were evaluated by software, against commercial gDNA, amplicon targets, and gDNA representative of our cohort samples.

First, the primers were evaluated by software for the absence of a possible secondary structure formation (<http://dinamelt.bioinfo.rpi.edu/twos-tate-fold.php>) and specificity (<http://blast.ncbi.nlm.nih.gov/ezp2.lib.umn.edu/Blast.cgi>). An additional pair of primers was designed (<http://blast.ncbi.nlm.nih.gov/ezp2.lib.umn.edu/Blast.cgi>) to generate amplicons of ~300 bp for use in sequencing validation studies. Second, all possible primer pairs were evaluated for tight peak, high fluorescence, lack of primer dimers or off target formation, and the crossing point through cross testing in a touchdown protocol against genomic DNA (Roche). Third, candidate primers against the appropriate synthetic targets were evaluated for clean separation of all three genotype amplicon melt profiles in a range of magnesium concentrations.

Finally, genomic DNA purified in the same laboratory and method as the data set samples was tested for performance and linearity. This multistep process is much more extensive than the usual selection process that often ends with the first step. Seven of our eight SNPs were assayed using primer pairs identified through the additional measures.

Patient samples

Genomic DNA from 382 patients was isolated from peripheral blood using a Qiagen kit. Patient gDNA samples at 20 ng/ μ L in the TE buffer were stored frozen at -80°C in metal foil-sealed 96-well plates until use. Seven reactions were

linear at 15 ng/15 μ L. One SNP amplification reaction was inhibited by the template DNA and was assayed at 1.5 ng/15 μ L reaction.

PCR and spiked reactions

Ninety-six well plates of 48 ng/well (or 4.8 ng/well) gDNA were mixed with 3.2 $\times$ SNP assay master mix containing 200 nM appropriate unlabeled primers, optimized magnesium, water, and Roche High Resolution Melting kit using pipette mixing. Triplicate samples were split from the 96-well plate to 15 μ L/well on a 384-well white plate and centrifuged for 5 min at 3000 g before loading in the LightCycler 480 (Roche). All assays were touchdown 62°C – 55°C at 0.5°C for 55 cycles followed by 94°C complete melting of amplicons, rapid cooling, and complete melting from 70°C to 90°C with data acquisition at 25 data points per degree melt. Three of the seven transitions were spiked by addition of ~20% higher T_m amplicon target to destabilize the colder of the homozygous profiles and enhance the homozygous profile differences (Evrard *et al.*, 2007). Based on crossing point, the higher T_m target was added to all samples as part of the master mix. The cooler T_m homozygous patient samples were weakly destabilized to a lower T_m . The higher T_m samples and heterozygous samples were unaffected, and the spiking allowed for greater separation of the homozygous states.

Validation studies

Patient gDNA samples were selected based on high-resolution melting (HRM) data to represent each of the three genotypes. Longer amplicons of each were created (Expand HiFi; Roche), cleaned (Qiagen PCR clean up), quantified by BioTek2, and submitted for Sanger classic sequencing at our core facility using one of the amplification primers.

Melt profile analysis and patient sample outcome recording

Melting profile data were reviewed using the manufacturer's instructions for two analysis options. Outcomes were evaluated by T_m calling and by using the gene scanning program, including normalization, temperature offset, and difference plotting. Patient sample outcome was generated from the T_m calling option, visualizing the three outcomes, and matching the tracings to the plate location.

Statistical analysis

Melt profile data were collected at 25 data points per degree and basic peak attributes were compiled by the HRM software. Peak widths and T_m values were exported and analyzed by the unpaired t test.

Results

Gene scanning versus T_m profiling

Our initial studies included gene scanning, but this resulted in significant difficulties in generating the expected difference plots from our synthetic targets. Thus, we included steps to evaluate the primer selection to optimize the differences in amplicons. Standard evaluations of second derivative maximums and T_m profiles were used to evaluate and select the candidate primer pairs.

FIG. 1. Twenty-four patients are shown in three analysis methods. **(A)** Using gene scanning protocol at 5° threshold offset, heterozygous (red) and homozygous (blue) patients are clearly and correctly separated with loss of differentiation between the homozygous ancestral and homozygous single-nucleotide polymorphism (SNP). **(B)** Using the gene scanning protocol at 0° threshold offset multiple groups are created, including erroneous subsets in the heterozygous and ancestral homozygous states, and two distinct homozygous SNP patients (pink). **(C)** T_m profiles clearly and correctly demonstrated three separate profiles.

Analysis of patient samples and synthetic control targets by both methods revealed the gene scanning option to be an inferior method even with enhanced primer selection and optimization. Following the manufacturer's standard protocol for gene scanning at 5° threshold offset, both homozygous states were combined and the data demonstrated the heterozygous patients as one group and both ancestral homozygous and SNP homozygous combined as a second group. Using the operator choices in normalization and threshold offset, we were able to create an interpretation of the melt profiles that placed our known SNP homozygotes in a separate difference plot outcome. However, manipulation of the operator inputs, including shifting to a 0° threshold offset, created multiple outcome groups, erroneously splitting the true ancestral homozygous category as well. A comparison of gene scanning Relative Signal Difference (RSD) plots for both standard threshold offsets and T_m melt profile plot demonstrated correct genotype separation for the T_m melt profile method (Fig. 1).

Validation with Sanger sequencing

Clear, definite, reproducible T_m differentiation of genotypes was observed for the eight SNPs of the study. One pa-

tient per genotype is shown in triplicate T_m tracings of a spiked transition SNP (Fig. 2). In our validation samples the automatic base calling software did not identify the heterozygous locations but rather, identified the location as either of the homozygous outcomes.

Transversion SNP representative replicate T_m tracings are shown in Figure 3. Intrasample variations were minimal, making a technical failure among replicates highly obvious. Direct inspection was required to identify the double tracing at all SNP heterozygous locations in the Sanger sequencing reactions. Sanger sequencing is not a viable option in high-throughput settings and the additional effort required to identify heterozygotes further calls into question its use as the gold standard for detecting heterozygosity.

Unambiguous differences of genotype were obvious for an entire triplicate plate of 96 samples even with intersample variations of melting temperature and amplification efficiency. Tracings from 288 samples for one SNP demonstrate three distinct genotype clusters (Fig. 4). The assay is clearly resilient to intersample variation, independent of amplification efficiency, and resistant to contamination. The method is resistant to false positives and failed samples are shown as baseline tracings. All samples were taken as triplicates to

FIG. 2. A/G transition. Melt peak profiles and sequence reactions for the three genotypes of an A/G transition. Triplicate tracings from three genotype representative patients are very nearly superimposed and demonstrate the reproducibility of the method. Homogeneous amplicon populations melt at a higher temperature in the genotype with three hydrogen bonds at the SNP and cooler in the genotype with two hydrogen bonds at the SNP. This reaction was spiked with 20% G target in all samples, shifting the A peak slightly to the left and facilitating clear separation of the peaks with no effect to the G or heterozygous samples.

FIG. 3. G/C transversion. Melt peak profiles and sequence reactions for the three genotypes of a G/C transversion. Taking advantage of the locking dye chemistry and the internal control, analysis of a transversion was simplified. In the absence of hydrogen bond change at the SNP, locking chemistry of the dye generates the nearest neighbor sequence change based on T_m differences. The T_m profiles of the gBlock internal control established the identity of the homozygous genotypes.

protect against incorrect interpretations such as early cycle amplification errors. Disparities within triplicates were very rare and obvious. The unpaired t test of peak width allowed significant separation of the heterozygous from the homozygous samples ($p < 0.0001$). A similar statistical analysis based on T_m demonstrated a significant separation of the homozygous samples ($p < 0.0001$).

Unexpected SNPs generated unique amplicon melting patterns (de Juan *et al.*, 2009; Er and Chang, 2012), and failed to generate amplicons. In 370 of 382 patient samples, the genotype status for all eight SNPs was generated. Of the 3056 T_m patient triplicate profiles generated, 3044 matched one of the three anticipated profiles for the indicated SNP. Three samples demonstrated unusual melt profiles inconsistent

with the expected genotype profiles. When analyzed by sequencing of longer amplicons, they showed either deletions (Fig. 5) or insertions. We speculate that the nine reactions that failed to generate amplicons represent unique sequences affecting primer binding.

Discussion

In an effort to establish SNP characterization as a widespread tool of individualized medicine, it is critical to reliably distinguish the homozygous condition (Wittwer *et al.*, 2003). Heterozygous individuals must be identified to study complete or partial inactivation of a gene, the penetrance of a trait, and fitness. An optimal assay will identify unique outcomes,

FIG. 4. Tracings of 288 reactions from 96 patients, assayed in triplicate, clearly demonstrate the three genotypes of this transversion SNP. Although intersample variations in salt content broaden the temperature peak width and differences in reaction efficiency affect the peak height, the three genotypes are obvious. Primer dimers are clearly seen in the colder temperature tracings and a single failed sample tracing takes its own path under the peaks. A/S, SNP heterozygote; S/S, SNP homozygote; A/A, ancestral homozygote melt clusters.

FIG. 5. A melt profile different from the SNP three genotypes and sequence reaction reveals a deletion heterozygous patient. Shown in triplicate are samples that demonstrate the expected profiles and a fourth profile seen only in one patient sample. A longer amplicon of the unique patient gDNA sample was sequenced revealing two deletions.

differentiate homozygous states, and identify heterozygous samples in a large sample population. An amplification internal control should be included in assays to distinguish differences in homozygous states caused by rare nucleotide changes and to confirm the identity of homozygous peaks in a transversion.

Previous methods to characterize SNPs have been associated with many technical issues. Likewise, HRM by gene scanning has been limited by cumbersome data processing, normalization, temperature shifting, and difference plot generation. Gene scanning not only requires a great deal of operator choices, but was highly unreliable in differentiation of the homozygous states and detection of the rare SNP homozygote. Probe-based assays are prone to false negative results especially in the identification of unique outcomes due to the absence of probe, and where a heterozygous individual may be identified incorrectly as homozygous. Sanger sequencing, which is unable to accommodate high-throughput settings, also does not accurately identify heterozygous locations. HRM was judged to be less time consuming or costly than conformational polymorphism, chromatography, or sequencing while able to differentiate primary mutations and identify novel mutations (Cui *et al.*, 2013).

Differentiation of homozygous states is predominantly based on melt temperature. Identification of the primer pair that generates amplicons with the largest T_m difference is essential to optimizing the assay. Using this method of optimization, we generated T_m differences between homozygous samples that were statistically different by the unpaired t test. The profile shape is critical for the detection of heterozygous samples in which, the melt profiles of both homozygous genotypes and the heterodimer are superimposed to create a wide profile. The peak width of heterozygous samples was also statistically different from a homozygous peak. HRM by T_m profile analysis considers both the temperature and shape, and thus, it is highly informative for distinguishing homozygous states and identifying the heterozygous individual. The basic characteristics of T_m and peak width provided the basis for separation of the three genotypes.

Not all SNP targets are excellent candidates for our method. Three of the SNP assays produced homozygous melt profiles that had temperature separations too close to allow inter-sample variations. These assays were optimized further by spiking with synthetic amplification targets, thus allowing the higher T_m target to increase the separation of homozygous state profiles. However, tri-allelic (Swen *et al.*, 2012) and difficult SNP targets (Liu *et al.*, 2012) that cannot meet the primer selection criteria can be interrogated by HRM through asymmetric PCR and the melting domain of an unlabeled extension blocked oligonucleotide. In these situations, similar attention to the melt profile differences must be made by optimization studies that include the expected genomic sequence synthetic targets against the oligonucleotide.

Using off the shelf products, an expanded primer selection criteria, and a double-stranded target control, we have generated information for the melt domains of seven transitions and a transversion, which are obvious and reproducible, based on both the peak melt temperature and shape. Variations in the samples, such as salt concentrations, exact DNA quantity, or template quality, will be present and affect the exact melt temperature or amplification efficiency. This method is robust in handling intersample quantity and quality variations. The assay is inexpensive, and using unlabeled primers it differentiates transition and transversion genotypes. Synthetic targets allow informative optimization and assay internal controls. Employing analysis that exploits both shape and temperature outcomes, unambiguous non-quantitative data that clearly demonstrate expected genotypes and unusual outcomes are now available as an inexpensive and straight forward protocol.

In summary, we have established an approach to HRM classification of SNPs that is more end user and template friendly, in part, because of greater attention to primer selection and the inclusion of template controls. It provides an improved method to carry out reproducible and accurate high-throughput analysis of SNP profiles. In addition, the separation of A/A and S/S homozygous samples was reliably based on T_m differences, in contrast to the separation of heterozygous A/S from homozygous samples, which was based on melt profile.

Acknowledgments

This work was supported by The Clinical Core of the Mayo Clinic Center for Cell Signaling in Gastroenterology (P30DK084567) to L.A.B., the Bernstein Colon and Rectal Surgery Library and Research Fund, and the Minnesota Medical Foundation fund to P.S.T. The authors would also like to thank Joseph Donnenhoffer from Roche Applied Science for technical support.

Author Disclosure Statement

The authors declare no competing interests.

References

- Cui G, Ding H, Xu Y, *et al.* (2013) Applications of the method of high resolution melting analysis for diagnosis of Leber's disease and the three primary mutation spectrum of LHON in the Han Chinese population. *Gene* 512:108–112.
- de Juan I, Esteban E, Palanca S, *et al.* (2009) High-resolution melting analysis of rapid screening of BRCA1 and BRCA2 Spanish mutations. *Breast Cancer Res Treat* 115:405–414.

- Er T-K, Chang JG (2012) Invited critical review: high resolution melting: applications in genetic disorders. *Clin Chim Acta* 414:197–220.
- Evrard A, Rayna C, Boyer J, *et al.* (2007) Mutation scanning of the cytidine deaminase gene by high-resolution melting curve analysis using the lightcycler^R 480 system. *Biochemica No 3*. <http://rocheappliedscience.com/techresources/techpage> (accessed December 23, 2012).
- Laurie D, Smith MP, George PM (2007) Detection of factor VIII gene mutations by high resolution melting analysis. *Clin Chem* 53:2211–2214.
- Liew M, Pryor R, Palais R, *et al.* (2004) Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem* 50:1156–1164.
- Lin CT, Er T, Tsai F, *et al.* (2010) Development of a high resolution melting method for the screening of Wilson disease-related ATP7B gene mutations. *Clin Chim Acta* 411:1223–1231.
- Liu S, Xu F, Shen F, *et al.* (2012) Rapid genotyping of APOA5-1131 T>C polymorphism using high resolution melting analysis with unlabeled probes. *Gene* 498:276–279.
- Swen J, Baak-Pablo R, Guchelaar H, *et al.* (2012) Alternative methods to a TaqMan assay to detect a tri-allelic single nucleotide polymorphism rs757210 in the HNF1 β gene. *Clin Chem Lab Med* 50:279–284.
- Tricarico R, Cruicianielli F, Alvau A, *et al.* (2011) High resolution melting analysis for a rapid identification of heterozygous and homozygous sequence changes in the MUTYH gene. *BMC Cancer* 11:305.
- Wittwer CT, Reed GH, Gundry CN, *et al.* (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin Chem* 49:853–860.

Address correspondence to:

Carol M. Bruzzone, MS

Department of Medicine

University of Minnesota Medical School

Phillips Wangenstein Building 14-212

516 Delaware Street S.E.

Minneapolis, MN 55455

E-mail: bruzz001@umn.edu

Clifford J. Steer, MD

Department of Medicine

University of Minnesota Medical School

406 Harvard Street, S.E.

Minneapolis, MN 55455

E-mail: steer001@umn.edu