GENOTYPING OF SINGLE NUCLEOTIDE POLYMORPHISMS BY HIGH RESOLUTION MELTING OF SMALL AMPLICONS

Michael Liew¹, Robert Pryor², Robert Palais³, Cindy Meadows¹, Maria Erali¹, Elaine Lyon^{1, 2} and Carl Wittwer^{1, 2}.

¹Institute for Clinical and Experimental Pathology, ARUP, Salt Lake City UT

84108-1221, USA.

²Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT

84132.

³Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090

Address for correspondence:

Carl T. Wittwer Department of Pathology University of Utah School of Medicine Salt Lake City, UT 84132

Phone: (801) 581-4737 Fax: (801) 581-4517 E-mail: carl.wittwer@path.utah.edu

Running title: SNP genotyping using LCGreen I

Keywords: SNP, Genotyping, Melting analysis, LCGreen I, HR-1, PCR, MTHFR,

Prothrombin, Factor V, Hemochromatosis, β-globin

ABSTRACT

Background. Homogeneous polymerase chain reaction (PCR) methods for genotyping single nucleotide polymorphisms (SNPs) usually require fluorescentlylabeled oligonucleotide probes or allele specific amplification. However, high-resolution melting of PCR amplicons with the DNA dye LCGreen[™] I was recently introduced as a homogeneous, closed-tube method of genotyping that does not require probes or realtime PCR. We adapted this system to genotype SNPs after rapid-cycle PCR (12 min) of small amplicons (<=50 bp).

Methods. Engineered plasmids were used to study all possible SNP base changes. In addition, clinical protocols for factor V (Leiden) G1691A, prothrombin G20210A, MTHFR A1298C, HFE C187G, and β -globin (HbS) A17T were developed. LCGreen I was included in the reaction mixture prior to PCR and high-resolution melting obtained within 2 min after amplification.

Results. In all cases, heterozygotes were easily identified because heteroduplexes altered the shape of the melting curves. About 84% of human SNPs involve a base exchange between A::T and G::C base pairs and the homozygotes are easily genotyped by melting temperatures (Tms) that differ by $0.8-1.4^{\circ}$ C. However, in about 16% of SNPs, the bases only switch strands and preserve the base pair, resulting in very small Tm differences between homozygotes (< 0.4° C). Although most of these cases can be genotyped by Tm, one-quarter (4% of total SNPs) show nearest neighbor symmetry, and, as predicted, the homozygotes cannot be resolved from each other. In these cases, adding 15% of a known homozygous genotype to unknown samples results in melting curve separation of all three genotypes.

Conclusions. SNP genotyping by high-resolution melting analysis is simple, rapid and inexpensive, requiring only PCR, a DNA dye, and melting instrumentation. The method is closed-tube, performed without probes or real-time PCR, and can be completed in less than two minutes.

INTRODUCTION

LCGreenTM I is a new fluorescent DNA dye designed to detect heteroduplexes during homogeneous melting curve analysis (<u>1</u>). Genotyping of SNPs by high-resolution melting analysis in products as large as 544 bp has been reported. Unlike SYBR[®] Green I, LCGreen I saturates the products of PCR without inhibiting amplification and does not redistribute as the amplicon melts. This allows closed-tube, homogeneous genotyping without fluorescently-labeled probes (<u>2-4</u>), allele-specific PCR (<u>5, 6</u>), or real-time PCR instruments. Heterozygotes are identified by a change in melting curve shape, and different homozygotes are distinguished by a change in melting temperature (Tm). However, it was not clear whether all SNPs can be genotyped by this method.

SNP genotyping by amplicon melting analysis requires high-resolution methods. The differences between genotypes are easier to see when the amplicons are short (7). Using small amplicons of <50 bps also allows for very rapid thermal cycling (8) – amplification is complete in less than 12 min followed by high-resolution melting requiring less than 2 min.

All possible homozygous and heterozygous genotypes with differences at one base position were studied using engineered plasmids. In addition, assays were developed to genotype the common clinical markers, prothrombin G20210A (9), factor V (Leiden) G1691A (2), methylenetetrahydrofolate reductase (MTHFR) A1298C (10), hemochromatosis (HFE) C187G (11) and β -globin (HbS) A17T (12) as examples for each class of SNP.

METHODS

DNA samples. Most of the samples used in this study were blood specimens submitted to ARUP (Salt Lake City, UT) for routine clinical genotyping of prothrombin, factor V, MTHFR, or HFE mutations. DNA was usually extracted with the MagNa Pure instrument (Roche Indianapolis, IN) according to the manufacturer's instructions. Additional samples genotyped at the β -globin locus for HbS were provided as dried bloodspots by Pediatrix Screening Inc. (Pittsburgh, PA) and extracted as previously described (<u>13</u>). All samples were genotyped at ARUP or Pediatrix Screening by melting curve analysis on the LightCycler[®] (Roche) using adjacent hybridization probe (HybProbeTM) technology, using either commercial kits (Roche), or in-house methods (<u>2</u>, <u>11, 12</u>). At least three different individuals of each genotype for prothrombin G20210A, MTHFR A1298C, HFE C187G and β -globin A17T SNPs were selected. One hundred and four samples (35 wild type, 35 heterozygote, 34 homozygous mutant) previously genotyped for factor V (Leiden) G1691A were obtained. All samples were de-identified according to a global ARUP protocol under IRB #7275.

DNA samples obtained with the MagNa Pure or from dried blood spots were not routinely quantified, but contained approximately 10-50 ng/µl. However, for HFE C187G genotyping, DNA was extracted using a QIAamp DNA Blood Kit (QIAGEN, Inc., Valencia, CA), concentrated by ethanol precipitation and quantified by A₂₆₀.

Engineered plasmids with either an A, C, G, or T at a defined position amid 50% GC content (<u>14</u>) were kindly provided by Cambrex BioScience, Inc. (Rockland, ME). Plasmid copy number was quantified by A₂₆₀.

Primer selection and synthesis. To maximize the melting temperature difference between normal and homozygous mutant genotypes, the amplicons were made as short as possible. The following process was systematized as a computer program using LabView (National Instruments, Austin, TX) and is available for remote use as "SNPWizard", at <u>DNAWizards.path.utah.edu</u>. After input of sequence information surrounding the SNP, the 3'-end of each primer is placed immediately adjacent to the SNP. The length of each primer is increased in its 5' direction until its predicted melting temperature (Tm) is as close to a user-selectable temperature (usually 55-60°C) as possible. Then, the primer pair is checked for the potential to form primer dimers or alternative amplicons. If the reaction specificity is acceptable, the primers are selected. If alternative products are likely, the 3'-end of one of the primers is shifted one base away from the SNP and the process is repeated until an acceptable pair is found.

Duplex Tms were calculated using nearest-neighbor thermodynamic models described previously (<u>15-22</u>). Best-fit values of 0.2 μ M for the amplicon concentration at the end of PCR and the Mg⁺⁺ equivalence (74-fold that of Na⁺) were obtained using a data set of 475 duplexes (<u>23</u>). The effective concentration of Mg⁺⁺ was decreased by the total dNTP concentration, assuming stoichiometric chelation. The effect of Tris⁺ was assumed equal to Na⁺ and the [Tris⁺] (20 mM) was calculated from the buffer concentration and pH. Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA), IT Biochem (Salt Lake City, UT), Qiagen Operon (Alameda, CA) and the University of Utah core facility (Salt Lake City, UT).

For the *in silico* calculation of the expected Δ Tm distribution of SNP homozygotes, the six amplicons studied (factor V, prothrombin, MTHFR, HFE, β -globin,

and pBR322) were considered. For each target, all combinations of the three bases pairs centered on the SNP were considered. A/A, C/C, G/G and T/T homozygotes were paired with each other giving six pairs of homozygotes for Δ Tm calculations. For each pair, 16 possible combinations of neighboring bases were considered. Therefore, 6 x 6 x 16, or 576 Δ Tm values were calculated and plotted as a frequency distribution, adjusted for the frequency of each SNP class as given in Table 1.

PCR. Reaction conditions for the engineered plasmids and the β -globin samples consisted of 50 mM Tris, pH 8.3, 500 µg/ml BSA, 3 mM MgCl₂, 200 µM of each dNTP, 0.4U Taq Polymerase (Roche), 1x LCGreen I (Idaho Technology, Salt Lake City, UT) and 0.5 µM each primer in 10 µl. The DNA templates were used at 10⁶ copies (plasmids) or 20 ng (genomic) and a two-temperature PCR was performed with 35 cycles of 85°C with no hold and 55°C for 1 s on either the LightCycler (Roche) or the RapidCycler II (Idaho Technology). PCR was completed within 12 min.

PCR for the prothrombin, factor V, MTHFR and HFE targets was performed in a LightCycler with reagents commonly used in clinical laboratories. Ten microliter reaction mixtures consisted of 10-50 ng of genomic DNA, 3 mM MgCl₂, 1x LightCycler FastStart DNA Master Hybridization Probes master mix, 1x LCGreen I, 0.5 µM forward and reverse primers and 0.01U/µl *Escherichia coli (E. coli)* uracil N-glycosylase (UNG, Roche). The PCR was initiated with a 10 min hold at 50°C for contamination control by UNG and a 10 min hold at 95°C for activation of the polymerase. Rapid thermal cycling was performed between 85°C and the annealing temperature at a programmed transition rate of 20°C/s. The online supplement lists primer sequences, amplicon sizes, the number of thermal cycles, and the annealing temperatures for each target.

Differentiating HFE wild type and mutant homozygotes required spiking the samples with a known genotype. The known DNA spike could be added either before or after PCR. To spike after PCR, equal volumes of known wild type and unknown PCR homozygous products were mixed. To spike before PCR, precisely 50 ng of unknown genomic DNA was used as template, along with an additional 7.5 ng of known wild-type DNA.

Melting curve acquisition and analysis. Melting analysis was performed either on the LightCycler immediately after cycling, or on a high-resolution melting instrument (HR-1, Idaho Technology). When the LightCycler was used, 20 samples were analyzed at once by first heating to 94°C, cooling to 40°C, heating again to 65°C (all at 20°C/s) followed by melting at 0.05°C/s with continuous acquisition of fluorescence until 85°C. LightCycler software was used to calculate the derivative melting curves.

When high-resolution melting was used, amplified samples were heated to 94° C in the LightCycler and rapidly cooled to 40° C. The LightCycler capillaries were then transferred to the HR-1 high-resolution instrument and heated at 0.3° C/s. Samples were analyzed between 65° C and 85° C with a turn-around time of 1-2 min. High-resolution melting data was analyzed with HR-1 software. In most cases, fluorescence vs temperature plots were normalized as previously described (<u>1</u>, <u>7</u>). For direct comparison to LightCycler data, derivative plots were used without normalization. All curves were plotted using Microsoft Excel after export of the data.

RESULTS

Melting analysis of short PCR products in the presence of the heteroduplexdetecting dye, LCGreen I, was used to genotype SNPs. Rapid-cycle PCR of short products allowed amplification and genotyping in a closed-tube system without probes or allele-specific amplification in less than 15 min. The primer locations surrounding the six polymorphic sites analyzed are shown in Fig. 1. The PCR products were 38-50 bp in length and the distance from the 3'-end of the primers to the polymorphic site varied from one to six bases.

The difference between standard and high-resolution melting techniques is shown in Fig. 2, using derivative melting curves of different factor V (Leiden) genotypes. Although the heterozygotes can be identified by the presence of a second, low temperature melting transition even with standard techniques, genotype differentiation is much easier with high-resolution methods. All subsequent studies were done at highresolution.

Engineered pBR322 constructs (<u>14</u>) were used to study all possible SNP base combinations at one position. Four plasmids (identical except for an A, C, G, or T at one position) were either used alone to simulate homozygous genotypes, or in binary combinations to construct "heterozygotes". The normalized melting curves of the four homozygotes and six heterozygotes are shown in Fig. 3. All homozygotes melt in a single transition (Fig 3A) and the order of melting is correctly predicted by nearest neighbor calculations as A/A < T/T < C/C < G/G (<u>22</u>). Heterozygotes result in more complex melting curves (Fig. 3B) arising from contributions of two homoduplexes and two heteroduplexes (<u>7</u>)). Each heterozygote traces a unique melting curve path according

to the four duplex Tms. The order of melting is again according to nearest neighbor calculations (A/T < A/C < C/T < A/G < G/T < C/G) using the average of the two homoduplex Tms. The six heterozygote curves merge at high temperatures into three traces, predicted by the highest melting homoduplex present (T/T for the A/T heterozygote, C/C for the A/C and C/T heterozygotes, and G/G for the A/G, G/T, and C/G heterozygotes). All genotypes can be distinguished from each other with highresolution melting analysis.

The genomic SNPs shown in Fig. 1 include all four classes of SNPs. Table 1 lists the four classes of SNPs that result from grouping the six different binary combinations of bases by the homoduplex and heteroduplex products that are produced when a heterozygote is amplified. Class 1 SNPs are C/T and G/A transitions that result in C::G and A::T homoduplexes and C::A and T::G heteroduplexes. In contrast, class 2 SNPs (C/A and G/T) are transversions that result in C::T and A::G heteroduplexes. Class 3 SNPs (C/G) result in C::G homoduplexes with C::C and G::G heteroduplexes. Class 4 SNPs (A/T) result in A::T homoduplexes with A::A and T::T heteroduplexes. The clinical SNPs studied were chosen to include two examples (factor V and prothrombin) in the most common SNP class and one example in each of the other three classes.

The melting curves for the five clinical SNP targets are shown in Fig. 4. For all SNP classes, heterozygotes were easily identified by a low and/or broad melting transition. For SNPs in class 1 or 2 (factor V, prothrombin, MTHFR), homozygous wild type and homozygous mutant samples were easily distinguished from each other by a shift in Tm. However, the Tm difference between homozygous genotypes for SNPs in class 3 or 4 was smaller than in class 1 or 2. Homozygous HbS (A17T, class 4) could be

distinguished from wild type with a Tm difference of about 0.2°C, but the HFE homozygous mutant (C187G, class 3) could not be distinguished from wild type.

Complete genotyping of HFE C187G by high-resolution melting analysis was possible by spiking in a known genotype into the unknown sample. Fig. 5A shows the result of mixing wild type amplicons with unknown homozygous amplicons after PCR. If the unknown sample is wild type, the melting curve does not change. However, if the unknown sample is homozygous mutant, heteroduplexes are produced and an additional low temperature transition appears. An alternative spiking option is to add a known genotype to the unknown sample before PCR. If a small amount of wild type DNA is added, wild type samples generate no heteroduplexes, homozygous mutant samples show some heteroduplexes, and heterozygous samples result in the greatest amount of heteroduplex formation (Fig 5B). Table 2 shows complete concordance between fluorescent hybridization probe (HybProbe[™]) and high-resolution amplicon melting methods for 167 samples.

Fig. 6 shows the *in silico* frequency distribution (24) of the calculated Δ Tms between the homozygous genotypes of SNPs, adjusted for the frequency of each SNP class in the human genome. All possible Δ Tm combinations (576) were calculated by varying the 3 bases centered on the natural SNPs of the 6 amplicons studied here. Class 1 and 2 SNPs form the broad cluster around 0.8 – 1.4°C and are easily distinguishable by melting analysis. Class 3 and 4 SNPs include the minor peaks around 0.00 and 0.25°C. Although most Class 3 and 4 SNPs can be fully genotyped by high-resolution melting analysis, one quarter have identical predicted Tms and the homozygotes cannot be differentiated without spiking.

DISCUSSION

There are many ways to genotype SNPs (25). Available techniques that require a separation step include restriction fragment length polymorphism analysis, single nucleotide extension, oligonucleotide ligation and sequencing. Additional methods, including pyrosequencing (26) and mass spectroscopy (27), are technically complex but can be automated for high-throughput analysis.

Homogeneous, closed-tube methods for SNP genotyping that do not require a separation step are attractive for their simplicity and containment of amplified products. Most of these methods are based on PCR and use fluorescent oligonucleotide probes. Genotyping occurs either by allele-specific fluorescence (28, 29) or by melting analysis (30). Melting analysis has the advantage that multiple alleles can be genotyped with one probe (31). Most of these techniques can be performed after amplification is complete, even though they are often associated with real-time PCR (32-36).

Some closed-tube fluorescent methods for SNP genotyping do not require probes. Allele-specific PCR can be monitored in real-time with SYBR Green I (<u>5</u>). The method requires three primers, two PCR reactions for each SNP, and a real-time PCR instrument that can monitor each cycle of PCR. An alternate method uses allele-specific amplification, SYBR Green I and melting curve analysis at the end of PCR (<u>6</u>). Monitoring each cycle is not necessary and an SNP genotype can be obtained in one reaction. However, a melting instrument and three primers are necessary with one of the primers modified with a GC-clamp. Both techniques are based on allele-specific PCR and each allele-specific primer is designed to recognize only one allele.

SNP genotyping by high-resolution melting with the dye LCGreen I does not require probes, allele-specific PCR or real-time PCR. Only two primers, one PCR reaction, and a melting instrument are required. Reagent costs for genotyping by amplicon melting are low because only PCR primers and a generic dye are needed. No probes or specialized reagents are required.

Although SNPs have been genotyped within amplicons up to 544 bp long (<u>1</u>), using a small amplicon for genotyping has numerous advantages. Assay design is simplified because primers are selected as close to the SNP as possible. The Tm differences between genotypes increases as the amplicon size decreases, allowing better differentiation. Cycling times can be minimized because the melting temperatures of the amplicons (74-81°C in Figs. 3 and 4) allow low denaturation temperatures during cycling that in addition increase specificity. Furthermore, the amplicon length is so small that no temperature holds are necessary for complete polymerase extension. Potential disadvantages of small amplicons include less flexibility in the choice of primers, less effective contamination control with UNG (<u>37</u>), and difficulty distinguishing between primer dimers and desired amplification products on gels or during real-time analysis.

Small amplicons allow rapid-cycle protocols that complete PCR in 12 min with popular real-time (LightCycler, Roche) or inexpensive instruments (RapidCycler II, Idaho Technology). Heteroduplex detection in small amplicons is favored by rapid cooling before melting, rapid heating during melting, and low Mg⁺⁺ concentrations (<u>7</u>). Although conventional real-time instruments can be used for melting (Fig. 2), their resolution is limited. Small Tm differences between homozygotes (e.g., Fig. 4E) are not distinguished on the LightCycler (data not shown).

Dedicated melting instruments have recently become available (LightTyper,

Roche; HR-1, Idaho Technology). The HR-1 provides the highest resolution and is by far the least expensive. Although only one sample is analyzed at a time, the turn-around time is so fast (1-2 min), that the throughput is reasonable. The LightTyper is an interesting platform for high-throughput melting applications. However, the temperature homogeneity across the plate needs to be improved before homozygotes can be reliably distinguished (data not shown).

Can all SNPs be genotyped by simple high-resolution melting of small amplicons? Studies with engineered plasmids of all possible base combinations at one location initially suggested that the answer was, "yes" (Fig. 3). Heterozygotes were always easily identified. Whether the different homozygotes were easy to distinguish depended on the class of SNP (Table 1). The six possible binary combinations of bases (C/T, G/A, C/A, G/T, C/G, and T/A) group naturally into 4 classes based on the homoduplex and heteroduplex base pairings produced when a heterozygote is amplified. SNP homozygotes are easy to distinguish by Tm in the first two classes because one homozygote contains an A::T pair and the other a G::C pair. These short amplicons show homozygotes Tm differences mostly between 0.8-1.4°C and these two classes make up over 84% of human SNPs (<u>38</u>).

It is more difficult to distinguish the homozygotes of SNPs in class 3 and 4 (Table 1) because the base pair (A::T or C::G) is simply inverted, that is, the bases switch strands but the base pair remains the same. Differences in amplicon Tm still result from different nearest neighbor interactions with the bases next to the SNP site, but are usually less than 0.4°C (Fig 3A, Fig. 4D and 4E). Class 3 and 4 SNPs make up about 16% of

human SNPs. Genotyping of homozygotes is still possible in most cases with highresolution analysis.

Clinical SNPs of each class were selected for concordance studies with standard genotyping methods. Factor V (Leiden) G1691A and prothrombin G20210A were class 1 SNPs, MTHFR A1298C was class 2, HFE C187G was class 3 and β -globin (HbS) A17T was class 4.

The class 3 SNP studied (Fig. 4D) was unique in that we could not discriminate the different homozygotes by Tm using simple melting analysis. Inspection of the bases neighboring the SNP site reveals why (Fig. 1D). In this case, the neighboring bases are complementary, resulting in nearest neighbor stability calculations that are identical for the two homozygotes. To the extent that nearest neighbor theory is correct, the duplex stabilities are predicted to be the same. By chance alone, this nearest neighbor "symmetry" is expected to occur 25% of the time. When this occurs in class 1 or 2 SNPs, nearest neighbor calculations indicate that the stability of the two heteroduplexes formed are identical. This is not of consequence to SNP typing because all three SNP genotypes still have unique melting curves. However, nearest neighbor symmetry in class 3 or 4 SNPs predicts that the two homoduplex Tms (homozygous genotypes) are identical. This will occur in approximately 4% of human SNPs.

When nearest neighbor symmetry of class 3 or 4 SNPs predicts that the homozygotes will not be distinguished, complete genotyping is still possible by spiking the reactions with a known genotype, either before or after PCR. If amplicon is spiked after PCR, only the homozygotes need to be tested, but potential amplicon contamination is a disadvantage. Spiking before PCR requires either that the DNA concentration of the

samples is carefully controlled or that samples are run both with and without spiked DNA.

High-resolution amplicon melting with LCGreen I can also be used to scan for sequence differences between two copies of DNA (<u>1</u>). In mutation scanning, the method is similar to other heteroduplex techniques such as denaturing high performance liquid chromatography (<u>39</u>) or temperature gradient capillary electrophoresis (<u>40</u>). However, high-resolution melting is unique in that homozygous sequence changes can often be identified without spiking. In the case of SNP genotyping with small amplicons, spiking is rarely required.

ACKNOWLEDGEMENTS

The authors would like to thank Jamie Williams for her technical assistance and Noriko Kusukawa for reviewing the manuscript. This work was supported by grants from the University of Utah Research Foundation, the State of Utah Center of Excellence program, ARUP and Idaho Technology.

REFERENCES

- Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ. High-resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem 2003;49:853-60.
- Lay MJ, Wittwer CT. Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. Clin Chem 1997;43:2262-7.
- Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl 1995;4:357-62.
- Crockett AO, Wittwer CT. Fluorescein-labeled oligonucleotides for real-time pcr: using the inherent quenching of deoxyguanosine nucleotides. Anal Biochem 2001;290:89-97.
- Germer S, Holland MJ, Higuchi R. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. Genome Res 2000;10:258-66.
- Germer S, Higuchi R. Single-tube genotyping without oligonucleotide probes.
 Genome Res 1999;9:72-8.
- Gundry CN, Vandersteen JG, Reed GH, Pryor RJ, Chen J, Wittwer CT. Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. Clin Chem 2003;49:396-406.
- Wittwer CT, Garling DJ. Rapid cycle DNA amplification: time and temperature optimization. Biotechniques 1991;10:76-83.

- 9. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. Blood 1996;88:3698-703.
- Weisberg I, Tran P, Christensen B, Sibani S, Rozen R. A second genetic polymorphism in methylenetetrahydrofolate reductase (MTHFR) associated with decreased enzyme activity. Mol Genet Metab 1998;64:169-72.
- Bernard PS, Ajioka RS, Kushner JP, Wittwer CT. Homogeneous multiplex genotyping of hemochromatosis mutations with fluorescent hybridization probes. Am J Pathol 1998;153:1055-61.
- 12. Herrmann MG, Dobrowolski SF, Wittwer CT. Rapid beta-globin genotyping by multiplexing probe melting temperature and color. Clin Chem 2000;46:425-8.
- Heath EM, O'Brien DP, Banas R, Naylor EW, Dobrowolski S. Optimization of an automated DNA purification protocol for neonatal screening. Arch Pathol Lab Med 1999;123:1154-60.
- Highsmith WE, Jr., Jin Q, Nataraj AJ, O'Connor JM, Burland VD, Baubonis WR, et al. Use of a DNA toolbox for the characterization of mutation scanning methods. I: construction of the toolbox and evaluation of heteroduplex analysis. Electrophoresis 1999;20:1186-94.
- 15. Allawi HT, SantaLucia J, Jr. Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. Biochemistry 1998;37:2170-9.
- Allawi HT, SantaLucia J, Jr. Thermodynamics of internal C.T mismatches in DNA. Nucleic Acids Res 1998;26:2694-701.

- Allawi HT, SantaLucia J, Jr. Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. Biochemistry 1998;37:9435-44.
- Allawi HT, SantaLucia J, Jr. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry 1997;36:10581-94.
- Bommarito S, Peyret N, SantaLucia J, Jr. Thermodynamic parameters for DNA sequences with dangling ends. Nucleic Acids Res 2000;28:1929-34.
- Peyret N, Seneviratne PA, Allawi HT, SantaLucia J, Jr. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. Biochemistry 1999;38:3468-77.
- SantaLucia J, Jr., Allawi HT, Seneviratne PA. Improved nearest-neighbor parameters for predicting DNA duplex stability. Biochemistry 1996;35:3555-62.
- 22. SantaLucia J, Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U S A 1998;95:1460-5.
- 23. von Ahsen N, Wittwer CT, Schutz E. Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg(2+), deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. Clin Chem 2001;47:1956-61.
- 24. von Ahsen N, Oellerich M, Armstrong VW, Schutz E. Application of a thermodynamic nearest-neighbor model to estimate nucleic acid stability and optimize probe design: prediction of melting points of multiple mutations of apolipoprotein B-3500 and factor V with a hybridization probe genotyping assay on the LightCycler. Clin Chem 1999;45:2094-101.

- Kwok PY, Chen X. Detection of single nucleotide polymorphisms. Curr Issues Mol Biol 2003;5:43-60.
- Ronaghi M. Pyrosequencing for SNP genotyping. Methods Mol Biol 2003;212:189-95.
- 27. Sauer S, Gut IG. Genotyping single-nucleotide polymorphisms by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci 2002;782:73-87.
- Lee LG, Connell CR, Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Res 1993;21:3761-6.
- 29. Mhlanga MM, Malmberg L. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. Methods 2001;25:463-71.
- Ronai Z, Sasvari-Szekely M, Guttman A. Miniaturized SNP detection: quasisolid-phase RFLP analysis. Biotechniques 2003;34:1172-3.
- Wittwer CT, Herrmann MG, Gundry CN, Elenitoba-Johnson KS. Real-time multiplex PCR assays. Methods 2001;25:430-42.
- Wittwer C, Kusukawa N. Real-time PCR. In: Persing D, Tenover F, Relman D, White T, Tang Y, Versalovic J, Unger B, eds. Diagnostic molecular microbiology; principles and applications. Washington: ASM Press, 2003:71-84.
- 33. Meuer S, Wittwer C, Nakaguwara K, eds. Rapid cycle real-time PCR: methods and applications. Berlin: Springer-Verlag, 2001:408pp.
- Reischl U, Wittwer C, Cockerill F, eds. Rapid cycle real-time PCR: methods and applications-microbiology and food analysis. Berlin: Springer-Verlag, 2002:258pp.

- Dietmaier W, Wittwer C, Sivasubramanian N, eds. Rapid cycle real-time PCR: methods and applications-genetics and oncology. Berlin: Springer-Verlag, 2002:180pp.
- 36. Wittwer C, Hahn M, Kaul K, eds. Rapid cycle real-time PCR: methods and applications-quantification. Berlin: Springer-Verlag, 2004:223pp.
- Espy MJ, Smith TF, Persing DH. Dependence of polymerase chain reaction product inactivation protocols on amplicon length and sequence composition. J Clin Microbiol 1993;31:2361-5.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001;291:1304-51.
- Wolford JK, Blunt D, Ballecer C, Prochazka M. High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC). Hum Genet 2000;107:483-7.
- Li Q, Liu Z, Monroe H, Culiat CT. Integrated platform for detection of DNA sequence variants using capillary array electrophoresis. Electrophoresis 2002;23:1499-511.

FIGURE LEGENDS

Figure 1. Details of the SNPs studied including the primer positions and the SNP class (see Table 1). Both strands of DNA are shown. The large arrows above and below the sequences indicate the 3' position and direction of the primers. The small vertical arrows indicate the SNP base change. For the pBR322 constructs, N indicates that all possible changes were studied.

Figure 2. Derivative melting curves for Factor V Leiden genotyping obtained on the LightCycler (A) and the HR-1 high-resolution instrument (B). Three individuals of each genotype were analyzed: wild type (solid black), homozygous mutant (dashed black) and heterozygous (solid grey).

Figure 3. Normalized, high-resolution melting curves of all possible SNP genotypes at one position using engineered plasmids. Three samples of each genotype were analyzed and included four homozygotes (A) and six heterozygotes (B).

Figure 4. Normalized, high-resolution melting curves from: A) factor V Leiden G1891A (Class 1), B) prothrombin G20210A (Class 1), C) MTHFR A1298C (Class 2), D) HFE C187G (Class 3), and E) β -globin A17T (Class 4) SNPs. Three individuals of each genotype were analyzed and are displayed for each SNP.

Figure 5. Genotyping at the HFE C187G locus by adding wild type DNA to each sample. In A) wild type amplicons were mixed with amplicons from three individuals of each homozygous genotype after PCR. In B) 15% wild type genomic DNA was added to the DNA of three individuals of each genotype before PCR.

Figure 6. *In silico* estimation of the Tm difference between homozygous genotypes of small amplicon SNPs. The frequency distribution is adjusted for the relative occurrence of each SNP class in the human genome (see Table 1). The larger the Δ Tm, the easier it is to differentiate the homozygous genotypes. Approximately 4% of human SNPs have a predicted Δ Tm of 0.00°C and are expected to require spiking with known homozygous DNA for genotyping of the homozygotes.

Table 1. SNP classification according to the homoduplexes and heteroduplexes produced after amplification of a heterozygote and the predicted number of distinct nearest neighbor thermodynamic duplexes $(Tms)^a$.

| Class | SNP (frequency) ^b | Homoduplex Matches (# of Tms) | Heteroduplex Mismatches (# of Tms) | Example (Figure Number) |
|-------|---------------------------------|-------------------------------------|--|----------------------------|
| 1 | C/T or G/A (0.675) | C::G and A::T (2) | C::A and T::G $(2 \text{ or } 1)^c$ | 3B, 4A, 4B |
| 2 | C/A or G/T (0.169) | C::G and A::T (2) | C::T and A::G $(2 \text{ or } 1)^c$ | 3B, 4C |
| 3 | C/G (0.086) | C::G $(2 \text{ or } 1)^c$ | C::C and G::G (2) | 3B, 4D, 5 |
| 4 | T/A (0.070) | A::T $(2 \text{ or } 1)^c$ | T::T and A::A (2) | 3B, 4E |

^{*a*}SNPs are specified with the alternative bases separated by a slash, for example C/T indicates that one DNA duplex has a C and the other a T at the same position on the equivalent strand. There is no bias for one allele over the other, that is, C/T is equivalent to T/C. Base pairing (whether matched or mismatched) is indicated by a double colon and is not directional. That is, C::G indicates a C::G base pair without specifying which base is on which strand.

^b The human SNP frequencies were taken from the Kwok data set as reported in (25).

^{*c*}The number of predicted thermodynamic duplexes depends on the nearest neighbor symmetry around the base change. One quarter of time, nearest neighbor symmetry is expected, that is, the position of the base change will be flanked on each side by complementary bases. For example, if a C/G SNP is flanked by an A and a T on the same strand (Fig. 1D), nearest neighbor symmetry occurs and only one homoduplex Tm is expected (as observed in Fig. 4D).

Table 2. Genotype concordance using adjacent hybridization probes (HybProbe[™]) and

| MARKER | GENOTYPES | HybProbe ^a | Amplicon melting ^b |
|-------------|-------------------|-----------------------|-------------------------------|
| Factor V | Wild type | 35 | 35 |
| G1891A | Heterozygous | 35 | 35 |
| | Homozygous mutant | 34 | 34 |
| Prothrombin | Wild type | 8 | 8 |
| G20210A | Heterozygous | 3 | 3 |
| | Homozygous mutant | 11 | 11 |
| MTHFR | Wild type | 6 | 6 |
| A1298C | Heterozygous | 7 | 7 |
| | Homozygous mutant | 7 | 7 |
| HFE | Wild type | 4 | 4^c |
| C187G | Heterozygous | 4 | 4 |
| | Homozygous mutant | 4 | 4^c |
| β-globin | Wild type | 3 | 3 |
| A17T | Heterozygous | 3 | 3 |
| | Homozygous mutant | 3 | 3 |

small amplicon, high resolution melting analysis (Amplicon melting).

^aAll samples were originally genotyped by ARUP (Factor V, prothrombin, MTHFR and

HFE) or Pediatrix Screening (β -globin) as clinical samples with adjacent hybridization

probes and melting curve analysis.

^bGenotyping results of the same samples using LCGreen I, the HR-1 high-resolution

melting instrument, and amplicon melting.

^cGenotyping required spiking with homozygous DNA (see text).

Table for online data supplement. Primer sequences, amplicon size and thermal

| SNP | PRIMER SEQUENCES (AMPLICON SIZE) | ANEALING TEMPERATURE (# of PCR cycles) |
|------------------------|--|--|
| Factor V G1891A | CAGATCCCTGGACAGG CAAGGACAAAATACCTGTATTC (42bp) | 55°C (32) |
| Prothrombin G20210a | GTTCCCAATAAAAGTGACTCTCAG GCACTGGGAGCATTGAGG (45bp) | 63°C (39) |
| MTHFR A1298C | GGAGGAGCTGACCAGTGAA AAGAACAAAGACTTCAAAGACACTT (46bp) | 55°C (35) |
| HFE C187G | CCAGCTGTTCGTGTTCTATGAT CACACGGCGACTCTCAT (40bp) | 63°C (35) |
| β-globin A17T | TGG TGCACCTGACTCCT AGTAACGGCAGACTTCTCC (38bp) | 55°C (35) |
| pBR322 | TCTGCTCTGCGGCTTTCT CGAAGCAGTAAAAGCTCTTGGAT (50bp) | 55°C (35) |











