DNA Melting Analysis for Detection of Single Nucleotide Polymorphisms

Robert H. Lipsky,[*†] Chiara M. Mazzanti,[†] Joseph G. Rudolph,[†] Ke Xu, Gopal Vyas,
David Bozak, Marta Q. Radel, and David Goldman

Background: Several methods for detection of single nucleotide polymorphisms (SNPs; e.g., denaturing gradient gel electrophoresis and denaturing HPLC) are indirectly based on the principle of differential melting of heteroduplex DNA. We present a method for detecting SNPs that is directly based on this principle.

Methods: We used a double-stranded DNA-specific fluorescent dye, SYBR Green I (SYBR) in an efficient system (PE 7700 Sequence Detector) in which DNA melting was controlled and monitored in a 96-well plate format. We measured the decrease in fluorescence intensity that accompanied DNA duplex denaturation, evaluating the effects of fragment length, dye concentration, DNA concentration, and sequence context using four naturally occurring polymorphisms (three SNPs and a single-base deletion/insertion).

Results: DNA melting analysis (DM) was used successfully for variant detection, and we also discovered two previously unknown SNPs by this approach. Concentrations of DNA amplicons were readily monitored by SYBR fluorescence, and DNA amplicon concentrations were highly reproducible, with a CV of 2.6%. We readily detected differences in the melting temperature between homoduplex and heteroduplex fragments 15–167 bp in length and differing by only a single nucleotide substitution.

Conclusions: The efficiency and sensitivity of DMA make it highly suitable for the large-scale detection of sequence variants.

© 2001 American Association for Clinical Chemistry

Detection of variations in the canonical sequence of the human genome is a cornerstone for the analysis of genetic diseases, genetically influenced traits, and the origin and structure of populations. Single nucleotide polymorphisms (SNPs)[1] comprise the most abundant category of DNA sequence variation, occurring at a rate of ~1 per 500 nucleotides in coding sequences and at a higher rate in noncoding sequences (1). Genome wide there are several million common SNPs. SNPs are highly amenable for high-throughput genotyping using efficient methodologies, including DNA arrays (2–4), mass spectrometry (5), and PCR end-plate read methods such as TaqMan (PE 7700 Sequence Detector) (6). For this reason, detection of SNPs of moderate or high abundance (i.e., with rare allele frequencies of >10%) has been made a high priority and may be successfully accomplished by analyzing a relatively small complement of chromosomes. Sequencing of 10 chromosomes (five persons) detects approximately two-thirds of all sequence variants with a frequency of 10%. However, SNPs that alter gene expression or affect structure of the gene product often are rare. Sequencing of 10 chromosomes detects <10% of sequence variants with a frequency of 1%, a frequently used definition for polymorphism.

The continuing requirement for detection of rare SNPs will maintain the need for high-throughput, inexpensive methods for sequence variant detection. Sequence variant detection methods access physical properties of DNA. Several of the most powerful methods for SNP detection [e.g., denaturing gradient gel electrophoresis (DGGE) and denaturing HPLC (dHPLC)] are ultimately based on the thermodynamic properties of DNA duplexes or single-stranded DNA [i.e., single strand conformational polymorphism (SSCP) analysis].

For many years it has been understood that the relative thermostability of a DNA duplex increases with increas-

---

Laboratory of Neurogenetics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Rockville, MD 20852.

*Address correspondence to this author at: 12420 Parklawn Dr., Suite 451, MSC 8110, Rockville, MD 20852. Fax 301-443-8579; e-mail rlipsky@mail. nih.gov.

†These authors contributed equally to the project.

---

[1] Nonstandard abbreviations: SNP, single nucleotide polymorphism; DGGE, denaturing gradient gel electrophoresis; dHPLC, denaturing HPLC; SSCP, single-strand conformation polymorphism; dsDNA, double-stranded DNA; T_m, melting temperature; DMA, DNA melting analysis; and AChR, acetylcholine receptor.

ing GC content. Subsequently, it has become apparent that DNA duplex stability is also dependent on additional factors. Mathematical models have predicted and empirical results verified that the largest contribution by far to DNA helix stability actually originates from the vertical stacking of base pairs *(7)* and that the stability of double-stranded DNA (dsDNA) depends largely on the identities of the nearest-neighbor bases that determine this stacking *(7, 8)*. For oligonucleotide sequences in solution, a two-state (i.e., duplex and random coil) nearest-neighbor model reliably predicts the stability of DNA duplexes with Watson-Crick pairs *(9, 10)*. Sequence composition-dependent differences in homoduplex stability were first exploited by Lerman and Silverstein *(11)*, who detected sequence variations of DNA by use of DGGE.

The nearest-neighbor model was extended for heteroduplex stability to include parameters for interactions between the mismatches and neighboring bases *(10, 12, 13)*. Allawi and SantaLucia *(10)* found good agreement between predicted and observed thermodynamic parameters (ΔG_{37}, ΔH, and ΔS) of 45 single internal G·T mismatched oligonucleotides of different base composition and length. In addition, the magnitude of the difference in the melting temperature (T_m) of single-base mismatches vs single-base substitutions in homoduplexes was large in these oligonucleotides (i.e., 2–5 °C). The nearest-neighbor model also predicts the stability and thermodynamics of DNAs containing other internal single-base mismatches (C·T, A·C, G·A) *(14–16)*. A difference in the thermodynamic stability of heteroduplex DNA directly explains the improved efficiency of gel-based methods such as DGGE to detect sequence variations by comparison of a heteroduplex DNA with component homoduplex DNAs. This contrasts with the reduced sensitivity obtained by comparing a homoduplex DNA with the homoduplex DNA containing a single nucleotide substitution that is correctly Watson-Crick base paired.

Our experimental approach was to use heteroduplex melting in solution to detect single nucleotide sequence variations. The goal was to create a high-throughput system for the detection of sequence variants within polynucleotide PCR amplicons. Two enabling elements for this new approach are SYBR Green I (SYBR), a dye that fluoresces when bound to dsDNA, and the PE 7700 Sequence Detector, which enables simultaneous DNA melting and fluorescence quantification in a 96-well format. In this study, we used three naturally occurring SNP variants and a single nucleotide insertion/deletion to explore the effects of dye concentration, DNA concentration, and fragment length on the ability to detect SNPs by heteroduplex melting in solution by DNA melting analysis (DMA).

Materials and Methods

OLIGONUCLEOTIDE SYNTHESIS

An ABI 394 DNA/RNA Synthesizer (Perkin-Elmer) was used to synthesize the oligonucleotides used to create the 15- and 25-bp *HTR2A* fragments. We synthesized four oligomers with the variant nucleotide at positions 8 and 13, respectively. The two sense strands contained either a T or C. The two complementary strands contained either an A or G. The single-stranded DNAs were annealed to create pure homoduplex (perfect match) or pure heteroduplex molecules (single-base mismatch). To test the ability to detect a single base mismatch in a dsDNA mixture composed of both homoduplexes and heteroduplexes, the heteroduplex DNA was diluted with homoduplex DNA in a 1:1 molar ratio. The sequences of the fragments were as follows (mismatches are underlined): *HTR2A* (15 bp), 5'-TTAACTC<u>T/C</u>GGAGAAG-3'; and *HTR2A* (25 bp), 5'-TGACTTTAACTC<u>T/C</u>GGAGAAGC TAAC-3'.

PCR AMPLIFICATION OF HUMAN GENOMIC DNA

DNA fragments of 60 bp or greater were synthesized by PCR on genomic DNA templates. For DMA experiments using known SNPs, such as *HTR2A* 102T→C, we selected DNAs from individuals who were either homozygous for an allele (*HTR2A* 102T→C 102T/102T) or heterozygous (*HTR2A* 102T/102C). The total volume for the PCR reaction was 25 μL and contained 100 ng of genomic DNA, 0.25 mM dNTPs, 0.5 μM PCR primers, and 0.75 U of AmpliTaq Gold with appropriate buffer (Perkin-Elmer). PCR master mixtures were prepared daily. The denaturation and extension steps for all PCRs were at 95 °C for 15 s and 72 °C for 30 s in 30 cycles, respectively. The annealing temperature for the 60-, 78-, 100-, and 152-bp amplicons was 50 °C for 20 s. Each fragment was amplified using the following primers:

For the 60-bp fragment: forward, 5'-ACCAGGCTCTA-CAGTAATGA-3'; reverse, 5'-GTTAAATGCATCA-GAAGTGT-3'

For the 78-bp fragment: forward, 5'-TAAATGATGA-CACCAGGCTC-3'; reverse, 5'-CTGTCCAGTTAAA-TGCATCA-3'

For the 97-bp fragment: forward, 5'-CACCAGGCTC-TACAGTAATG-3'; reverse, 5'-GGAAAGGTTGGT-TCGATT-3'

For the 100-bp fragment: forward, 5'-TTAAATGATGA-CACCAGGC-3'; reverse, 5'-TGGTTCGATTTTCA-GAGTC-3'

For the 152-bp fragment: forward, 5'-GCTCAACTAC-GAACTCCCT-3'; reverse, 5'-TGAGAGGCACCCT-TCACAG-3'

For other known SNPs, the annealing step for the variants tested was 20 s. Annealing temperatures were 55 °C for *HTR2A* His452Tyr, 54 °C for *COMT* Val158Met, and 59 °C for *DRD2*-141C ins/del. For screening of unknown SNPs, the annealing step was performed at 60 °C for 20 s. PCR amplicons used for SNP screening by DMA were based on exon sequences of the *N*-methyl-D-aspartate receptor gene (*NR1*; GenBank Accession No. Z32773) and both exon and intron sequences from the acetylcho-
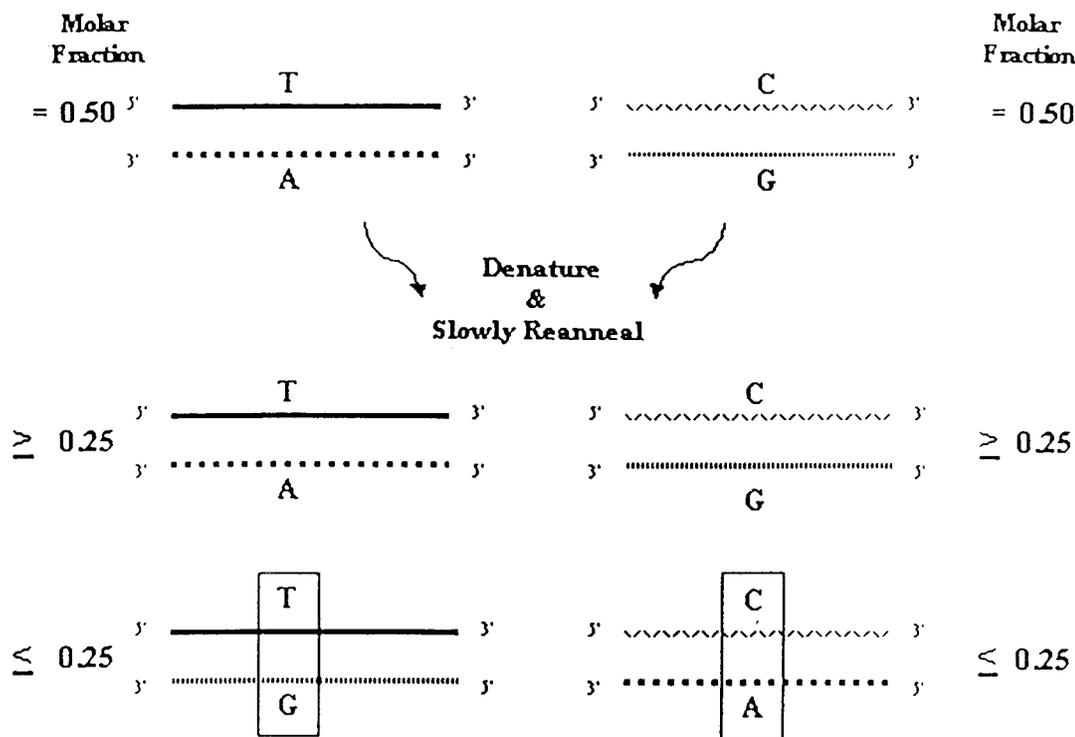
Fig. 1. Principle of DMA.

Denaturation and reannealing of heterozygous DNA leads to a mixture of four duplex DNAs. The dsDNA is composed of two different homoduplexes and two different heteroduplexes. Because of the thermodynamics of reannealing, the homoduplex DNA represents at least 50% of the total amount, whereas the heteroduplex DNA is at most 50% of the total duplexes. Mismatches in the heteroduplex DNA are represented by *open boxes*.

line receptor gene (*AChR*) α-subunit (GenBank Accession No. X02502). The primer sequences used to amplify these sequences were as follows:

> For *HTR2A*: forward, 5′-AGATGCCAAGACAACA-GATA-3′; reverse, 5′-ATTCACTCCGTCGCTATT-3′
> For *COMT*: forward, 5′-CTCATCACCATCGAGATCAA-3′; reverse, 5′-CCAGGTCTGACAACGGGTCA-3′
> For *DRD2*: forward, 5′-GTGTGGGTGGGAGCGCAGTG-3′; reverse, 5′-CCCCACCAAAGGAGCTGTA-3′
> For *NR1*: forward, 5′-CTAACACTCTTGCTCACACC-3′; reverse, 5′-ATCCTGTGTGGAGTGTGTAG-3′
> For *AChR*: forward, 5′-GCCCTGGTCCACACAAG-3′; reverse, 5′-TGGTCTCATCAAAGAAGCAA-3′

Following PCR amplification, each product was purified using either the Qiagen reagent set for PCR or gel extraction. The purified product was then denatured at 95 °C for 4 min and reannealed by slowly cooling to 60 °C over a period of 30 min to permit the formation of a mixture of homoduplex and heteroduplex molecules in the DNA amplified from heterozygous individuals.

SYBR FLUORESCENCE DETECTION OF DNA MELTING

The double strand-specific dye, SYBR Green I (SYBR), was obtained from Molecular Probes and used for DMA. The optimal excitation and emission spectra of SYBR are centered at 492 and 513 nm, respectively. SYBR is sup-

plied as a 10 000× concentrate by the manufacturer, with no molar concentration values or formula weights being supplied. The optimal concentration of SYBR used in our experiments was 3.6× dye (reduced from 10 000×), and dilutions were with 1× Tris-borate-EDTA buffer. The 3.6× SYBR dye concentration was optimal for concentrations of DNA between 20 and 100 ng in 15 μL. Fluorescence measurements and denaturation were accomplished using the PE 7700 Sequence Detector. Fluorescence signals were recorded approximately every 7 s over the entire time course of denaturation, which varied from 30 min to 4 h, gathering data for up to 96 samples at a time.

DATA ANALYSIS

Following DNA duplex denaturation and data acquisition, raw fluorescence data were exported to Microsoft Excel for statistical analyses. Fluorescence data from melting curves were converted into T_m by plotting the negative derivative of fluorescence vs temperature ($-dF/dT$ vs T). Melting point predictions were performed using MeltCalc software *(17)*.

VARIANT SCREENING USING dHPLC

Following PCR amplification, samples were denatured and reannealed as described above for DMA to enhance formation of DNA heteroduplexes. Samples were then
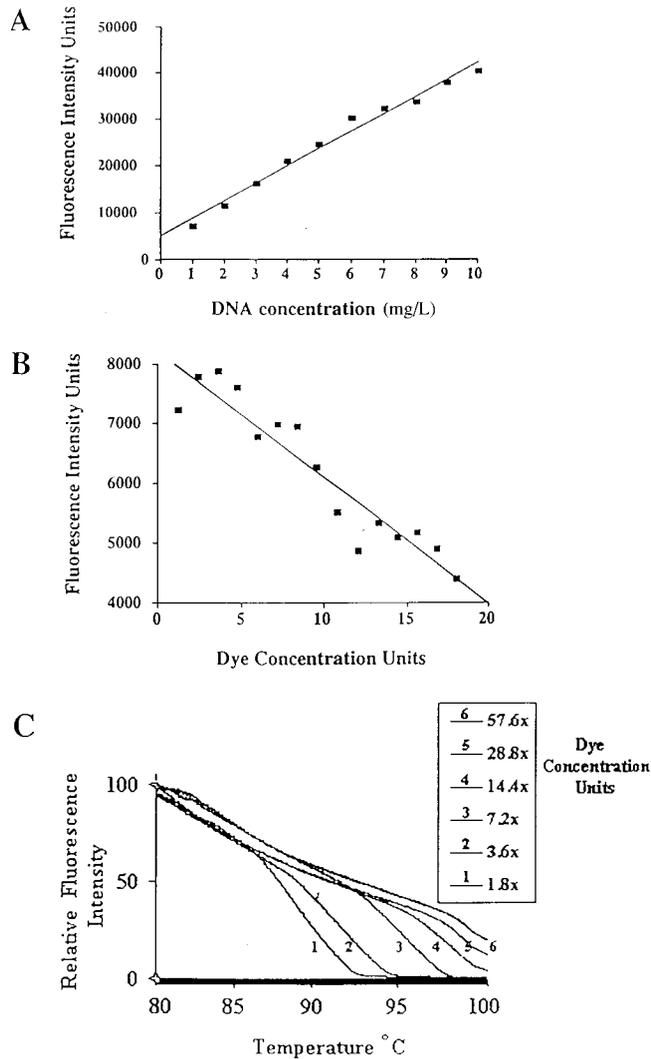
A

B

C

Fig. 2. Measurement of dsDNA concentration with SYBR using the PE 7700 Sequence Detector.

The dsDNA was the *HTR2A* gene 100-bp amplicon. (*A*), fluorescence intensity is proportional to dsDNA concentration at DNA concentrations of 1–6 mg/L and at a SYBR concentration of 3.6×. Slope of the line was 184, y-intercept was 5096, and r = 0.99. (*B*), SYBR quenches dsDNA fluorescence intensity signal. SYBR concentration was varied from 1.2× to 18×, with a dsDNA concentration of 1.5 mg/L. The slope of the line was −211, y-intercept was 8211, and r = 0.94. (*C*), effect of varying SYBR concentration (1.8× to 57.6×) on dsDNA melting, as measured by SYBR fluorescence intensity using the PE 7700.

processed using a Transgenomic dHPLC, consisting of a 96-well autosampler, column oven, pumps, degasser, variable wavelength ultraviolet detector, sample loop, and a PC-based data collection system. Before dHPLC, melting curves for PCR amplicons were simulated using the Transgenomic Wavemaker™ software to determine whether any significant shifts in T_m could be predicted for the amplicon. A Transgenomic DNASep column was used for separations. Buffers used on the column were as follows: buffer A, consisting of 10 mmol/L triethylammonium acetate (pH 7.4); and buffer B, consisting of 10 mmol/L triethylammonium acetate containing 250 mL/L acetonitrile. Loading buffer consisted of 80 mL/L aceto-

nitrile. T_ms and buffer gradients were determined using the Transgenomic melting temperature predictions software.

DNA SEQUENCE ANALYSIS OF CANDIDATE SNPs
Genomic DNA samples that were determined by DMA to contain a SNP were amplified using PCR, and the products were purified as described above to eliminate excess primer and genomic DNA. The sequencing reaction was performed in a 10-μL reaction containing 10 nmol/L one primer (forward or reverse), 4 μL of BigDye™ Terminator Cycle Sequencing reaction mixture (Perkin-Elmer), and 3 μL of the purified PCR product. The temperature cycle for the sequencing reaction consisted of 25 cycles of 10 s at 96 °C, 5 s at 56 °C, and 4 min at 60 °C. Purification of this reaction was performed using a gel filtration block (AGTC). The purified reaction products were dried under reduced pressure and resuspended in formamide and sequencing dye mixture. Products were resolved on an ABI 377 automated sequencer. The determined sequence was aligned and analyzed using the ABI AutoAssembler software.

Results

To determine the effect of SYBR concentration, DNA concentration, and fragment length on duplex DNA melting, we used a polymorphism located in the first exon of the 5-HT2a receptor (*HTR2A*) gene (102T→C; GenBank Accession No. X57830). These analyses were actually performed by comparing the melting characteristics of the DNA homoduplex to a mixture of homo- and heteroduplexes that naturally result from the amplification or synthesis of dsDNA and its denaturation and slow reannealing. As shown in Fig. 1, amplification, denaturation, and reannealing of heterozygous DNA will yield four dsDNA duplexes. The two homoduplexes constitute at least 50% of the reannealed dsDNA, and the two heteroduplex DNAs constitute at most 50% of the reannealed dsDNA. It is important to recognize that each one of the four dsDNA species formed during reannealing has different thermodynamic characteristics. For small DNA duplexes, G·T mismatches are less thermodynamically destabilizing than A·C mismatches *(10)*. However, the magnitude of the difference in melting thermodynamics of the two heteroduplex DNA species makes them readily distinguishable from either of the homoduplex dsDNAs.

An important limitation of our approach is that both the structure and concentration of SYBR are proprietary. SYBR is reported to bind to the minor groove of dsDNA (Molecular Probes Manual). Once bound and excited at a wavelength of 450 nm, it has a maximum emission signal at 513 nm.

CONCENTRATION OF DNA AND SYBR
Increasing the DNA concentration increases the T_m of duplex DNA in solution. Therefore, the effect of varying the DNA concentration and the ability to control for

Table 1. Physical characteristics of DNAs analyzed by DMA.

Gene	SNP	Amino acid substitution	Fragment size,[a] bp (%GC)	SNP nt[b] position within fragment	Predicted ΔT_m,[c] °C (homoduplex − heteroduplex)	Observed ΔT_m, °C (homoduplex − heteroduplex)
HTR2A	102T→C[d]		15 (40)	8	2.9; 5.4	4.9
	102T→C[d]		25 (40)	13	2.4; 5.8	3.5
	102T→C		60 (40)	30	0.7; 1.8	5.5
	102T→C		78 (40)	41	0.5; 1.3	5.0
	102T→C		97 (43)	31	0.4; 1.0	2.8
	102T→C		100 (41)	42	0.9; 1.5	1.4
	102T→C		152 (43)	68	0.2; 0.6	0.9
	1499C→T	His452Tyr	94 (44)	51	0.9; 1.3	2.6
COMT	1947G→A	Val158Met	110 (59)	66	0.6; 0.6	3.8
DRD2-141C	ins/del		133 (67)	43	NP	1.2
NR1[e]	3680C→T	Ala310Val	163 (60)	86	0.5; 0.7	2.2
AChR[e]	147G→T		167 (58)	131	0.2; 0.6	0.6

[a] Nonvariant fragment.

[b] nt, nucleotide; NP, no prediction from MeltCalc.

[c] Based on T_m predictions for each of two heteroduplexes using MeltCalc software of Schütz and von Ahsen (*17*).

[d] 15- and 25-bp fragments were chemically synthesized, and the remainder were PCR amplicons.

[e] SNPs discovered during this study.

variation in DNA concentration were critical to our ability to detect sequence variation by DMA. dsDNA concentrations could vary as a result of different efficiencies in DNA amplification.

To approach this issue, the extent of variation in initial dsDNA concentrations before melting was monitored. For these determinations, we used the *HTR2A* 100-bp amplicon. As shown in Fig. 2A, the DNA concentration could be measured on the ABI 7700 by monitoring SYBR fluorescence intensity, because this intensity is linearly proportional to dsDNA concentration. It was noted that higher concentrations of SYBR quenched the fluorescence signal (Fig. 2B) and that the relationship between SYBR and this quenching was again linear. Quenching by excess SYBR may be attributable to the presence of unbound SYBR. Therefore, it is important to use a constant concentration of SYBR if DNA concentrations are to be determined.

dsDNA quantification revealed that the concentration of a particular DNA amplicon was highly reproducible as determined by the SYBR fluorescence signal. The within-day imprecision (CV) for the *HTR2A* 100-bp amplicon from genomic DNA from an individual heterozygous for the T102C variant was 2.6% at a mean concentration of 1.98 g/L. Thus, amplicon concentrations were highly reproducible. Similar results were also obtained for different amplicons using different primer combinations (data not shown). Therefore, DNA concentrations could be reliably monitored in all of the samples before DMA, using the same device in which DMA was conducted, namely the ABI 7700.

Increasing the dsDNA concentration increased the T_m of the 100-bp *HTR2A* amplicon. In the presence of SYBR, as DNA concentration was increased from 0.5 mg/L to 10 mg/L, the melting curves were shifted to the right, as expected (data not shown).

The effect of different concentrations of SYBR on the temperature of dsDNA denaturation was evaluated (Fig. 2C) using the *HTR2A* 100-bp amplicon at a concentration of 2 mg/L in a 15-μL final assay volume. This was a convenient DNA concentration given that the typical yield of a PCR reaction was ~10 mg/L in a volume of 30 μL. The manufacturer recommends that SYBR, supplied as a stock, be used after diluting 1:10 000. We determined that the optimal dye concentration for a wide range of DNA concentrations was 3.6×. The use of lower SYBR concentrations led to low emission signals, but higher dye concentrations caused signal quenching. In addition, increasing the concentration of SYBR from 1.8× to 57.6× shifted the DNA melting curve to the right, indicating that by binding the dsDNA, the dye is stabilizing the DNA duplex, thereby increasing the temperature at which it is denatured. In fact, at higher SYBR concentrations relative to a constant DNA concentration, it could be seen that the dsDNA did not completely denature, as evidenced by the melting curves obtained using 28.8× and 57.6× SYBR (Fig. 2C), which are shifted far to the right.

ABILITY TO DETECT SINGLE-BASE MISMATCHES IN DNA FRAGMENTS OF DIFFERENT LENGTHS

Six dsDNA fragments of different lengths were either synthesized (15 and 25 bp) or PCR amplified (60, 80, 100, and 152 bp). All fragments had the same T→C polymorphism located in the middle of the same target *HTR2A* gene sequence. The effect of the single-base pair mismatches (G·T and A·C) was assessed by comparing the thermodynamic stability of the DNA heteroduplexes with homoduplex DNA (A·T). A 97-bp fragment was also amplified in which the mismatch was located 30 bp from the 5′ end (see Table 1).

When the 15- or 25-bp dsDNA fragment was subjected to a temperature increment of 2 °C/min, a readily observable difference in the rate of decrease in SYBR fluores-
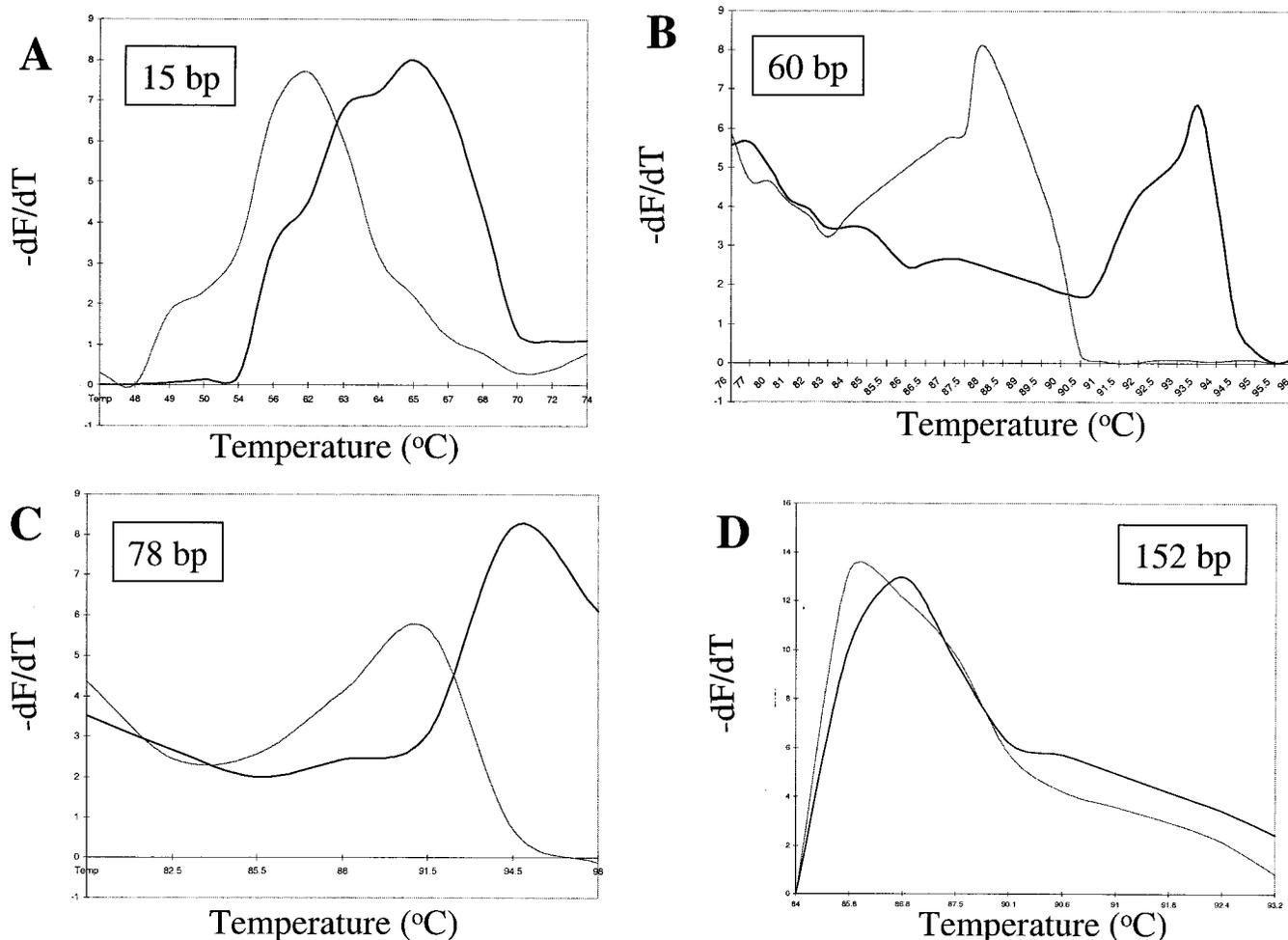
Fig. 3. Detection of single nucleotide mismatches in dsDNA fragments of different sizes.

Panels show the melting behavior of dsDNAs of increasing size containing the *HTR2A* 102T→C SNP. Data presented in *A* were produced using chemically synthesized 15-mer oligonucleotides, whereas the data presented in *B–D* are based on PCR amplicons. In each case, a heteroduplex/homoduplex mixture (see diagram in Fig. 1) was compared with the homoduplex dsDNA at the same concentration. Derivative melting profiles for homoduplexes are depicted with *thick lines*. The melting profile for the heteroduplex/homoduplex mixture is shown with a *thin line*. Peaks in the melting profile represent the T_m. DNA concentrations were 2 mg/L, and the SYBR concentration was 3.6×. Melting curves were acquired at melting rates of 2 °C/min for *A*, 0.5 °C/min for *B* and *C*, and 0.067 °C/min for *D*.

cence was observed for homoduplex dsDNA compared with the heteroduplex/homoduplex DNA mixture. The fluorescence data from each melting curve were converted into the T_m by plotting the negative derivative of fluorescence vs temperature ($-dF/dT$ vs T). A representative derivative plot for the 15-bp fragment is shown in Fig. 3A. The ΔT_m between the homoduplex DNA and the heteroduplex/homoduplex mixture was ~5 °C for the 15-bp fragment and 3.5 °C for the 25-bp fragment (Fig. 3A and Table 1). Each observed value was within the range of ΔT_m values predicted for that fragment (Table 1).

For longer dsDNAs, the rate of DNA melting was decreased to 0.5 °C/min for the 60- and 80-bp fragments and to 0.067 °C/min for fragments 97–152 bp in size. When this was done, the differences in the melting rates between homoduplex dsDNA and the heteroduplex/homoduplex mixture were again readily distinguishable. The observed ΔT_m for these fragments varied between 0.9 °C (152 bp) and 5.5 °C (60 bp; Fig. 3, B–D, and Table 1).

DETECTION OF ADDITIONAL SEQUENCE VARIANTS USING DMA

Two additional SNP variants and a single nucleotide insertion/deletion were also examined by DMA using amplicons 94, 110, and 133 bp in size. The SNPs, both of which produce amino acid substitutions, were *HTR2A* 1499C→T (His452Tyr; GenBank Accession No. X57830) and *COMT* 1947G→A (Val158Met; GenBank Accession No. Z26491) with fragment sizes of 94 and 110 bp, respectively. The insertion/deletion used was *DRD2*-141Cins/del (GenBank Accession No. X53502) in a 133-bp fragment (Fig. 4C). In all of these cases, the homoduplex/heteroduplex mixture produced derivative melting profiles that were distinguishable from the melting profiles for the homoduplex (Fig. 4). The homoduplex/heteroduplex mixture always had a lower T_m than that for the homoduplex (Fig. 4). The observed ΔT_m for each amplicon heteroduplex was 1.2–3.8 °C (Table 1).
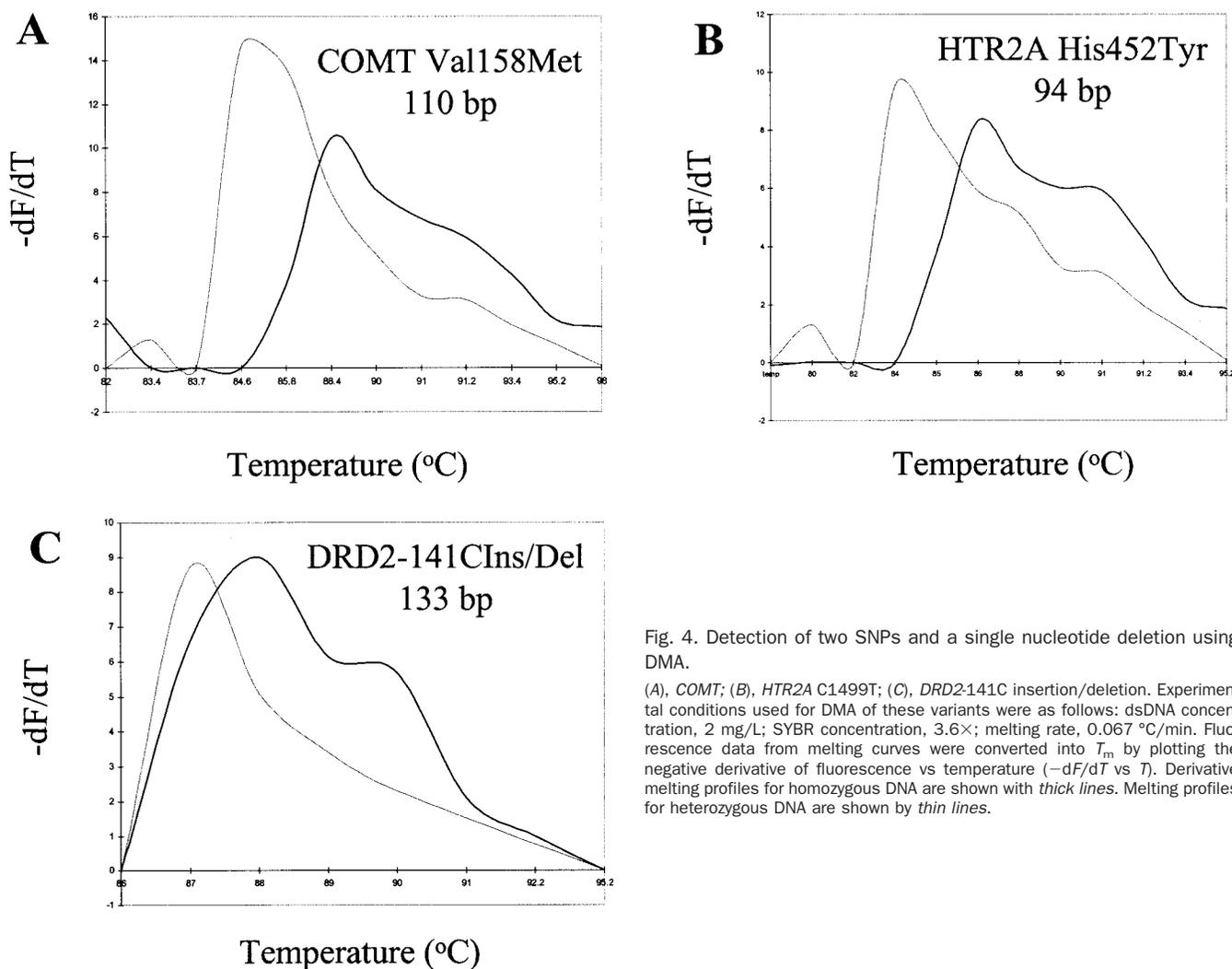
Fig. 4. Detection of two SNPs and a single nucleotide deletion using DMA.

(*A*), *COMT;* (*B*), *HTR2A* C1499T; (*C*), *DRD2*-141C insertion/deletion. Experimental conditions used for DMA of these variants were as follows: dsDNA concentration, 2 mg/L; SYBR concentration, 3.6×; melting rate, 0.067 °C/min. Fluorescence data from melting curves were converted into T_m by plotting the negative derivative of fluorescence vs temperature ($-dF/dT$ vs T). Derivative melting profiles for homozygous DNA are shown with *thick lines*. Melting profiles for heterozygous DNA are shown by *thin lines*.

SNP DISCOVERY USING DMA

The previous experiments allowed us to evaluate the merits of DMA using known SNPs. We wanted to use DMA to discover new SNPs and to validate our approach using previously known methods for SNP discovery and identification, namely dHPLC and direct DNA sequence analysis of PCR amplicons. To approach the issue of validation, we selected amplicon sizes that could also be easily screened by dHPLC. The PCR amplicons produced for these experiments were derived from portions of two different genes: (*a*) exon sequences of the *NR1* gene, and (*b*) exon and intron sequences from the *AChR* gene (see *Materials and Methods*). The two PCR amplicons were 163 bp (*NR1*) and 167 bp (*AChR*) in size and were amplified from genomic DNA prepared from six unrelated individuals. The PCR products were thermally denatured and allowed to reanneal, after which DMA was performed. The melting profile of each sample was followed, and the fluorescence data from melting curves were then converted into the T_m using the derivative melting profile ($-dF/dT$ vs T). Derivative melting data for representative

NR1 and *AChR* PCR amplicons are presented in Figs. 5A and 6A. One sample from the *NR1* gene amplification and one sample from the *AChR* gene amplification, each originating from a different individual, showed a shift in T_m when compared with the melting peaks obtained for the other *NR1* gene or *AChR* gene-derived samples. These results suggested the presence of heteroduplex/homoduplex DNAs. These melting profiles were repeated with similar results. The observed ΔT_ms for these fragments were 2.2 °C (163-bp *NR1*) and 0.6 °C (167-bp *AChR*; Table 1).

The two genomic DNA samples that produced the T_m shifts were then independently amplified for DNA sequence analysis. Direct DNA sequence analysis of these PCR products showed the presence of a SNP within the amplicon (Fig. 7). No other variants were detected in either of these PCR products. Direct sequencing of the *NR1* or *AChR* PCR products not showing a T_m shift indicated that they did not contain a SNP within the amplicon (data not shown). In addition, we confirmed that the new SNPs discovered in the *NR1* and *AChR* genes
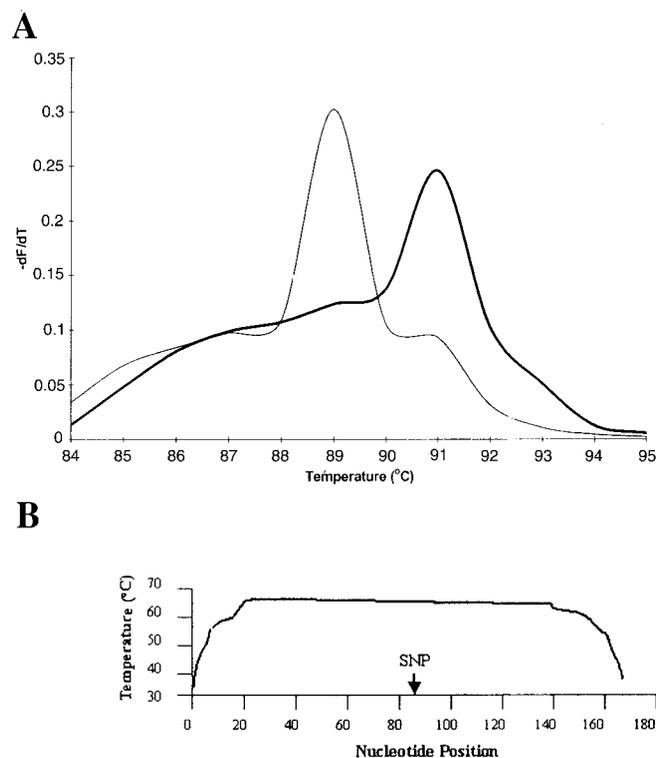
A

B

Fig. 5. Melting curve of a human *NR1* gene PCR amplicon used for detecting an unknown SNP.

(*A*), conditions used for DMA were as follows: dsDNA concentration, 2 mg/L; SYBR concentration, 3.6×; melting rate, 0.067 °C/min. Fluorescence data from melting curves were converted into T_m by plotting the negative derivative of fluorescence vs temperature ($-dF/dT$ vs T). The homoduplex melting curve is depicted by a *thick line*. The melting curve for the heteroduplex/homoduplex mixture is shown as a *thin line*. (*B*), predicted melting domains within the PCR amplicon. Position of the SNP was determined by DNA sequence analysis and is shown by an *arrow*.

A

B

Fig. 6. Melting curve of a human *AChR* gene PCR amplicon used for detecting of an unknown SNP.

(*A*), conditions used for DMA were as follows: dsDNA concentration, 2 mg/L; SYBR concentration, 3.6×; melting rate, 0.067 °C/min. Fluorescence data are presented as the negative derivative of fluorescence vs temperature ($-dF/dT$ vs T). Peaks in each curve represent the T_m. The homoduplex melting curve is depicted with a *thick line*. The heteroduplex/homoduplex mixture melting curve is shown by a *thin line*. (*B*), predicted melting domains within the PCR amplicon. Position of the SNP within the amplicon was determined by DNA sequence analysis and is shown by an *arrow*.

by DMA could be replicated using dHPLC (data not shown). It should be noted that each SNP was located in a region of the PCR amplicon that was predicted to behave as a single melting domain (Figs. 5B and 6B).

Discussion

In this study we have presented a rational method, DMA, for the detection of unknown SNPs. DMA utilizes the dsDNA-specific dye, SYBR, to detect differences in DNA melting thermodynamics between homoduplex dsDNA and heteroduplex/homoduplex dsDNA mixtures. DMA has been implemented in a highly efficient and sensitive system in which melting of synthesized oligomers or PCR amplicons is dynamically followed as a decrease in SYBR fluorescence. The method also allows for repeated analyses of the same DNA sample. The denatured DNA is simply reannealed, allowing the experiment to be repeated under different conditions if desired.

In principle, all dsDNA heteroduplex mismatches in amplicons 100 bp or less in size may be detectable by DMA. Our data showed that differences in melting profiles for homoduplex-heteroduplex DNA mixtures containing fragments 15–100 bp in length with single A·C and
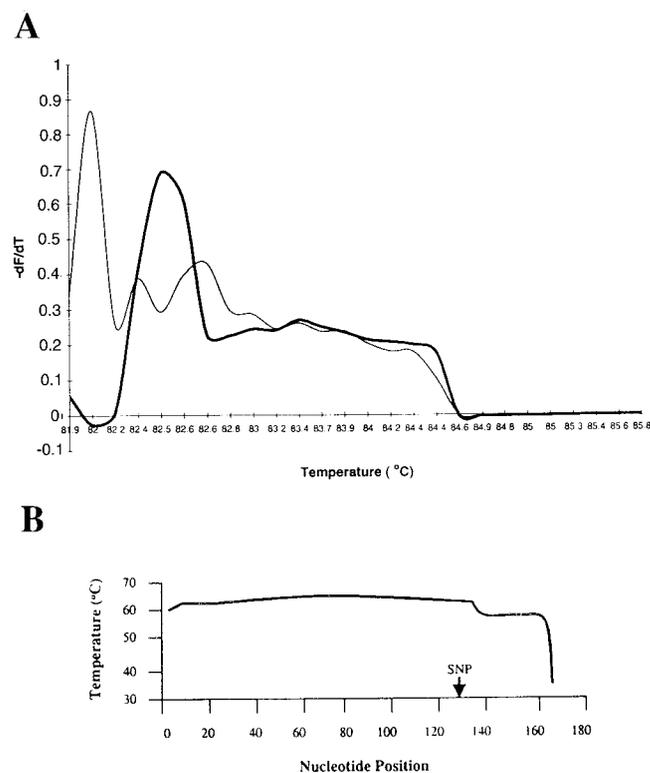
G·T internal mismatches were readily detectable by DMA. The observed ΔT_m for these different fragments varied between 1.4 and 5.5 °C (Table 1). For shorter DNA fragments containing either single A·C or G·T mismatches, A·C mismatches were the least stable thermodynamically, whereas G·T mismatches were the most stable *(15)*. Thus, DMA may allow direct access to entropy and enthalpy changes that result from single Watson-Crick mismatches. It should be noted that internal single-base mismatches could produce melting instability in larger fragments (>100 bp) and that our method was sufficiently sensitive to detect a melting difference in a homoduplex-heteroduplex mixture that contained either a single internal A·G mismatch or a single internal C·T mismatch (the *AChR* gene-derived fragment; Table 1).

Conceptually, DMA is simple. It quantifies the difference in heteroduplex stability while avoiding intervening steps, such as gels and columns, whose effectiveness depends on principles other than the thermodynamic properties of DNA. The conceptual simplicity of DMA leads to practical advantages over other high-throughput SNP detection methods. In SSCP, the SNP may not alter
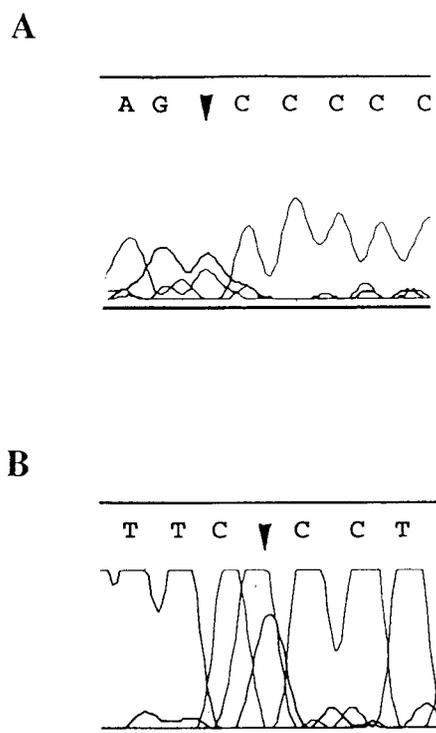
A

B

Fig. 7. Electropherograms showing the primary DNA sequence determined from a PCR product from an individual heterozygous for a SNP in the *AChR* gene (*A*) and individual heterozygous for a SNP located at position 3680 of the *NR1* gene (*B*).

(*A*), the overlapping peaks shown by the *arrowhead* (G and T) of the determined sequence correspond to nucleotide position 147 of the coding strand of the *AChR* gene. (B), the *arrowhead* indicates two overlapping peaks (A and G) from the noncoding strand of the *NR1* gene.

internal base pairing. In SSCP analysis, gel mobility is dependent on fragment size, the complex tertiary structures formed, and the way these conformations interact with the gel matrix. Different conformations may therefore have the same SSCP mobility. The lack of sensitivity of SSCP analysis led to a series of other SSCP approaches (e.g., the use of special gel constituents and SSCP analysis of multiple fragments containing the same sequence). The increased labor involved in these methods improves the sensitivity of SSCP analysis but also underlines its limitations. dHPLC is a relatively high-throughput method that is well suited to automation. However, as with SSCP analysis, the electrophoretic behavior of the DNA fragments can be erratic because we cannot yet predict the structure of the partially melted DNAs, nor can we predict their interaction with the immobile phase (i.e., reversed-phase columns). In dHPLC, the mobility of the partially melted DNA structure is highly dependent on the exact temperature and other conditions (e.g., pH or acetonitrile gradient) chosen.

DGGE most closely resembles DMA because the variant is detected based on the way in which the substitution alters the thermodynamic properties of heteroduplex or substituted homoduplex dsDNA. Like dHPLC and SSCP analysis, however, DGGE requires a shift in gel mobility,

which depends on many other variables. Finally, chemical and enzymatic mismatch cleavage both depend on molecular recognition of a mismatch site.

An understanding of the effects of sequence context on DNA melting has been largely developed using DNA oligomers <25 bp in length. For oligomeric DNAs, the thermodynamic parameters ΔT_m, ΔG, ΔH, and ΔS can be estimated using the output calculated by the MeltCalc program *(17)*, which is based on estimates of nearest-neighbor parameters for dsDNA with or without mismatches as described by Allawi and SantaLucia *(10, 14–16)*. However, the nearest-neighbor model has limitations because the melting of oligomers frequently, but not always, fits a two-state model (i.e., the DNA exists as either random coil or double helix). With longer DNA fragments, such as those used in our studies, this two-state model cannot be used because longer DNA fragments can have other unknown secondary structures, such as hairpin turns and loops. Water molecules and ions can also affect the relative stability of these structures *(18)*. The results from Table 1 support the idea that some deviations between mathematically predicted values and observed values of ΔT_m are also attributable to different structural characteristics of the longer DNA fragments.

In DMA, SYBR stabilizes DNA duplexes, and as illustrated in Fig. 2, doubling the dye concentration increased the T_m by 1–2 °C over large a range of dye concentrations. It is well known that both the DNA concentration and the time allowed for annealing or melting are critical factors for denaturation *(19)*. Our results show that DMA can be successfully used for variant detection provided that three factors, DNA concentration, transition time, and SYBR concentration, are all carefully controlled. In practice, only one of these three factors, i.e., DNA concentration, is likely to vary from one sample to the next. A low dsDNA concentration could simulate the melting behavior of heteroduplex DNA. Fortunately, our DNA yields were highly consistent because the amplifications were carried out consistently and under endpoint conditions, i.e., to >30 cycles, starting with >10 ng of genomic DNA template. These PCR conditions are fairly typical. Moreover, with DMA it is easy to detect significant amounts of variation in the initial dsDNA concentration using the initial fluorescence intensity as a quantitative measure.

The PE 7700 Sequence Detector (ABI Prism) was used to monitor the decrease in SYBR fluorescence signal over the course of dsDNA melting. With this particular instrument, it is possible to perform denaturation experiments and analyze the DNA melting curves in 96 samples at a time. The detector monitors the emission fluorescence signal approximately every 7 s, providing a very detailed and descriptive curve of DNA duplex melting behavior during the course of denaturation. With these detailed DNA melting curves, Watson-Crick mismatches in dsDNAs 100–150 bp in size were readily detected, with an observed ΔT_m of 1–5 °C.

Recently, melting analysis has been used for genotyp-

ing. Marziliano et al. *(20)* used a T_m assay to detect differences in the melting profile of a 132-bp PCR product having a TA insertion in the TATA box of the *UGT1A* gene promoter compared with the nonvariant reference promoter sequence. The melting profiles of the PCR products from individuals that were homozygous for either the deletion or the nonvariant sequence (reannealed amplicons composed of homoduplex DNA) were readily distinguishable from PCR products produced from heterozygous individuals (reannealed amplicons composed of heteroduplex DNA). It should be noted that the derivative melting profiles of the homoduplex samples and the heteroduplex samples generated single T_m peaks, similar to our own melting profiles (Figs. 5A and 6A).

In a separate study using melting curve analysis with a LightCycler™ (Roche Molecular Biochemicals), Aoshima et al. *(21)* detected differences in melting profiles between small PCR amplicons of 46 or 55 bp, the former having an internal 9-bp deletion, derived from a *CPS1* gene cDNA. However, the authors expected that the ΔT_m of a heteroduplex containing the 9-bp deletion and a homoduplex derived from a 100-bp fragment ($\Delta T_m = 0.6\ °C$) would make the melting difference undetectable with SYBR. However, we successfully detected SNPs in fragments >100 bp, including an unknown SNP in the *AChR* gene located within a 167-bp amplicon. For the *AChR* gene amplicon, the ability of DMA to detect a small melting temperature shift (observed $\Delta T_m = 0.6\ °C$) may have been aided by the location of the SNP, which was near a region of decreased stability (lower GC content; see Fig. 6B). In addition, we used a slower heating rate for larger amplicons than that used by Aoshima et al. *(21)*. Although additional comparative studies are certainly warranted to fully describe the power of DMA, the efficiency and sensitivity of this approach, as implemented on the PE 7700 Sequence Detector (and potentially the ABI 5700 and ABI 7900 HT), makes it highly suitable for large-scale detection of new sequence variants.

References

1. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 1998; 280:1077–82.
2. Fan JB, Chen X, Halushka MK, Berno A, Huang X, Ryder T, et al. Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. Genome Res 2000;10:853–60.
3. Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, Syvanen AC. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. Genome Res 2000;10: 1031–42.
4. O'Dell SD, Chen X, Day IN. Higher resolution microplate array diagonal gel electrophoresis: application to a multiallelic minisatellite. Hum Mutat 2000;15:565–76.
5. Jackson PE, Scholl PF, Groopman JD. Mass spectrometry for genotyping: an emerging tool for molecular medicine. Mol Med Today 2000;6:271–6.
6. Livak KJ. Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. Genet Anal 1999;14:143–9.
7. Borer PN, Dengler B, Tinoco I Jr. Stability of ribonucleic acid double-stranded helices. J Mol Biol 1974;86:843–53.
8. Freier SM, Sugimoto N, Sinclair A, Alkema D, Neilson T, Kierzek R, et al. Stability of XGCGCp, GCGCYp, and XGCGCYp helixes: an empirical estimate of the energetics of hydrogen bonds in nucleic acids. Biochemistry 1986;25:3214–9.
9. 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.
10. Allawi HT, SantaLucia J Jr. Thermodynamics and NMR of internal G·T mismatches in DNA. Biochemistry 1997;36:10581–94.
11. Lerman LS, Silverstein K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. Methods Enzymol 1987;155:482–501.
12. He L, Kierzek, R, SantaLucia J Jr, Walter AE, Turner DH. Nearest-neighbor parameters for G·U mismatches: [formula; see text] is destabilizing in the contexts [formula; see text] and [formula; see text] but stabilizing in [formula; see text]. Biochemistry 1991;30: 11124–32.
13. McDowell JA, Turner, DH. Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing. Biochemistry 1996;35:14077–89.
14. Allawi HT, SantaLucia J Jr. Nearest neighbor thermodynamic parameters for internal G·A mismatches in DNA. Biochemistry 1998;37:2170–9.
15. 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.
16. Allawi HT, SantaLucia J Jr. Thermodynamics of internal C·T mismatches in DNA. Nucleic Acids Res 1998;26:2694–701.
17. Schütz E, von Ahsen N. Spreadsheet software for thermodynamic melting point prediction of oligonucleotide hydration with and without mismatches. Biotechniques 1999;27:1218–22, 24.
18. Zieba K, Chu TM, Kupke DW, Marky LA. Differential hydration of dA·dT base pairing and dA and dT bulges in deoxyoligonucleotides. Biochemistry 1991;30:8018–26.
19. Wahl GM, Berger SL, Kimmel AR. Molecular hybridization of immobilized nucleic acids: theoretical concepts and practical considerations. Methods Enzymol 1987;152:399–407.
20. Marziliano N, Pelo E, Minuti B, Passerini I, Torricelli F, Da Prato L. Melting temperature assay for a *UGT1A* gene variant in Gilbert syndrome. Clin Chem 2000;46:423–5.
21. Aoshima T, Sekido Y, Miyazaki T, Kajita M, Mimura S, Watanabe K, et al. Rapid detection of deletion mutations in inherited metabolic diseases by melting curve analysis with LightCycler. Clin Chem 2000;46:119–22.