Single nucleotide polymorphism genotyping by two colour melting curve analysis using the MGB Eclipse™ Probe System in challenging sequence environment

Yevgeniy S. Belousov,[1] Robert A. Welch,[2] Silvia Sanders,[1] Alan Mills,[1] Alena Kulchenko,[1] Robert Dempcy,[1] Irina A. Afonina,[1] David K. Walburger,[1] Cynthia L. Glaser,[2] Sunita Yadavalli,[2] Nicolaas M.J. Vermeulen[1] and Walt Mahoney[1]*

[1]Epoch Biosciences, 21720 23rd Drive SE, Bothell, WA 98021, USA
[2]Core Genotyping Facility, Advanced Technology Center, National Cancer Institute, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
**Correspondence to:* Tel: +1 425 482 5153; Fax: +1 425 482 5550; E-mail: nvermeulen@epochbio.com

Date received (in revised form): 9th January 2004

Abstract

Probe and primer design for single nucleotide polymorphism (SNP) detection can be very challenging for A-T DNA-rich targets, requiring long sequences with lower specificity and stability, while G-C-rich DNA targets present limited design options to lower GC-content sequences only. We have developed the MGB Eclipse™ Probe System, which is composed of the following elements: MGB Eclipse probes and primers, specially developed software for the design of probes and primers, a unique set of modified bases and a Microsoft Excel macro for automated genotyping, which ably solves, in large part, this challenge. Fluorogenic MGB Eclipse probes are modified oligonucleotides containing covalently attached duplex-stabilising dihydrocyclopyrroloindole tripeptide (DPI_3), the MGB ligand (MGB™ is a trademark of Epoch Biosciences, Bothell, WA), which has the combined properties of allowing the use of short sequences and providing great mismatch discrimination. The MGB moiety prevents probe degradation during polymerase chain reaction (PCR), allowing the researcher to use real time data; alternatively, hybridisation can be accurately measured by a post-PCR two-colour melt curve analysis. Using MGB Eclipse probes and primers containing modified bases further enhances the analysis of difficult SNP targets. G- or C-rich sequences can be refractory to analysis due to Hoogsteen base pairing. Substitution of normal G with Epoch's modified G prevents Hoogsteen base pairing, allowing both superior PCR and probe-based analysis of GC-rich targets. The use of modified A and T bases allows better stabilisation by significantly increasing the T_m of the oligonucleotides. Modified A creates A-T base pairs that have a stability slightly lower than a G-C base pair, and modified T creates T-A base pairs that have a stability about 30 per cent higher than the unmodified base pair. Together, the modified bases permit the use of short probes, providing good mismatch discrimination and primers that allow PCR of refractory targets. The combination of MGB Eclipse probes and primers enriched with the MGB ligand and modified bases has allowed the analysis of refractory SNPs, where other methods have failed.

Keywords: *MGB Eclipse, modified base, genotyping, SNP, AT-rich, GC-rich, automated, melt curve*

Introduction

Several fluorogenic nucleic acid assays (Molecular Beacons™, TaqMan®, Invader®, Scorpion Primers etc) and assay formats — real-time polymerase chain reaction (PCR), end-point analysis, melting curve analysis — have been developed recently for single nucleotide polymorphism (SNP) detection.[1] Yet certain types of SNPs, like mismatches in G–C-rich targets or G–T substitutions in A–T-rich targets, or any G–T mismatches, remain difficult to detect.[2–4] The MGB Eclipse system significantly increases the number and composition of SNP targets that can be analysed. The solution to this problem came through the marriage of molecular biology and chemistry. G- or C-rich sequences (in particular areas of four or more consecutive Gs) can be refractory to analysis, due primarily to Hoogsteen base pairing.[5] Substitution of normal G with

modified G (Figure 1a) breaks Hoogsteen base pairs, allowing both PCR- and probe-based analysis of GC-rich targets. The use of our modified A and T allows better stabilisation by significantly increasing the T_m of oligonucleotides containing the said modified bases. Modified A creates an A-T base pair that has a T_m just below that of a G-C base pair, and Epoch's modified T creates a T-A base pair that has a T_m that is about 30 per cent higher than the unmodified T-containing base pair. Together, the presence of MGB and modified bases permits the use of short probes, providing good mismatch discrimination and primers that allow PCR of targets that are refractory due to high AT and GC content. The MGB Eclipse probes are stable to 5′-nuclease degradation, and their signal is hybridisation-based only. By contrast, the MGB ligand-free Molecular Beacon probes show signal generation by both 5′ nuclease cleavage and hybridisation-based mechanisms.[6] The use of the MGB ligand to improve SNP detection of traditional probes has previously been reported for TaqMan assays[7–9] and MGB Eclipse Systems.[6]

The MGB moiety at the 5′ end of MGB Eclipse probes effectively blocks the exonuclease activity of Taq, preserving the intact probe (Figures 1b and c). This allows amplified DNA targets to be accurately genotyped by analysis of the real-time data or by our preferred method, a fluorogenic melt curve analysis. After the completion of PCR, both MGB probes (each with a different fluorescent reporter dye) are annealed to the amplified SNP-containing sequence. The DNA duplexes are denatured over a time course and the decrease in the fluorescent signal of each probe is measured. Comparison of the two-colour melting curves allows differentiation of sequence variants of the target. This melt analysis, following the real-time data analysis, provides further conformation of the genotype call. In concert, the real-time data, the combination of two-colour melt curve analysis, the modified bases and the MGB moiety provide a rich system to specifically detect challenging SNPs. A comparison of real time and melting curve formats is described, and examples of genotyping are discussed.

Materials and methods

Synthesis of fluorogenic MGB probes and PCR primers

Fluorogenic MGB Eclipse probes were synthesised on an MGB-quencher solid support using commercially available reversed 5′-DNA phosphoramidites, as described elsewhere.[6] This approach was taken due to the insolubility of the MGB-quencher and the lack of a corresponding phosphoramidite derivative. Fluorescent reporter groups (fluorescein and 5′-tetrachlorofluorescein) were introduced using the corresponding phosphoramidites. A conventional glycol linker was introduced between the DNA sequence and the fluorescent dye to increase the distance between the fluorescent dye and quencher. MGB Eclipse probes and primers containing modified G

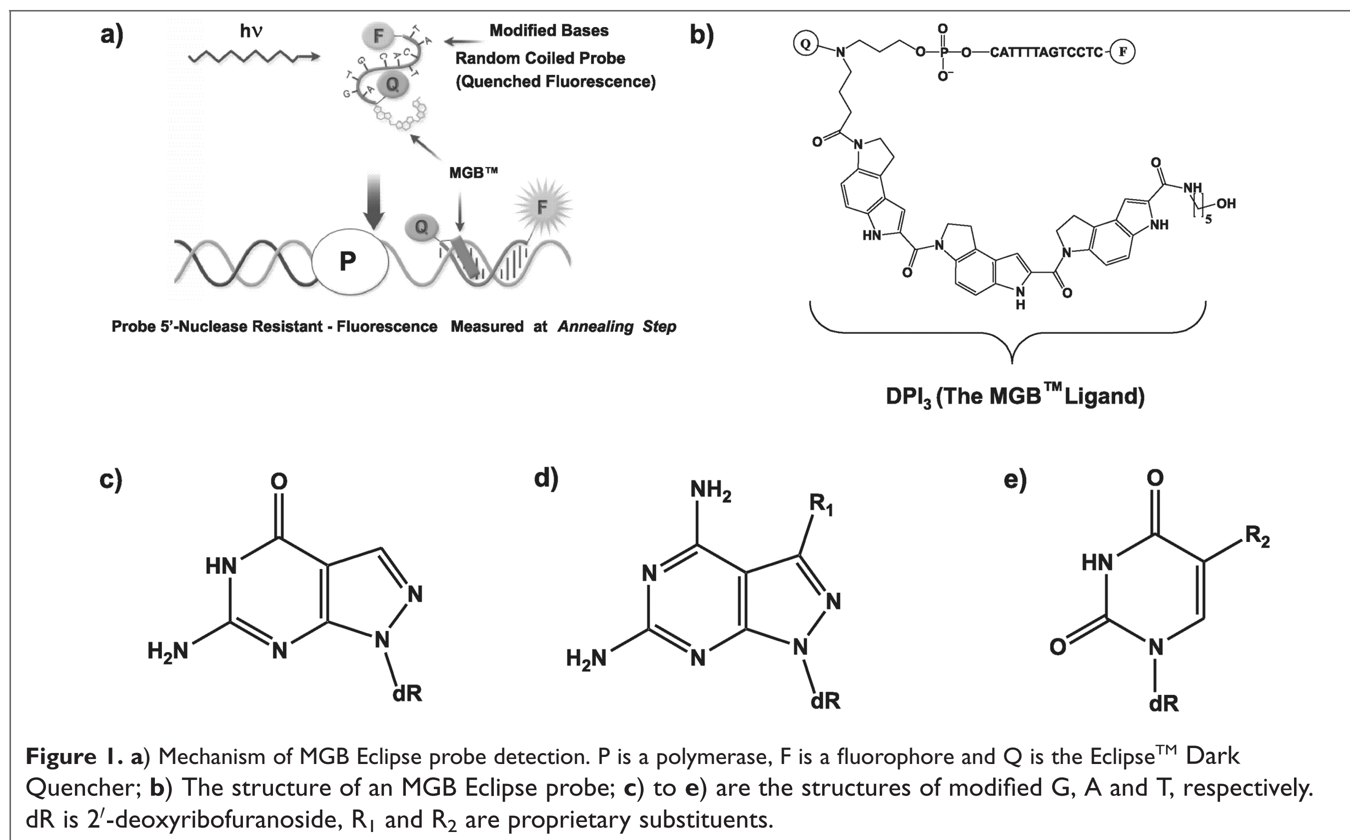


Figure 1. a) Mechanism of MGB Eclipse probe detection. P is a polymerase, F is a fluorophore and Q is the Eclipse™ Dark Quencher; **b**) The structure of an MGB Eclipse probe; **c**) to **e**) are the structures of modified G, A and T, respectively. dR is 2′-deoxyribofuranoside, R_1 and R_2 are proprietary substituents.

(Super G™, Epoch Biosciences, Bothell, WA) were prepared as described elsewhere,[6] using a dimethylformamide (DMF) protecting group. MGB Eclipse probes and primers containing modified A and T (Figure 1d and e; Super A™, Super T™, Epoch Biosciences, Bothell, WA) were prepared using conventional DNA synthesis protocols, the probes were synthesised using reversed phosphoramidites and the primers synthesised using forward phosphoramidites (synthesised by methods described in WO 0164958). MGB Eclipse probes were purified by reverse phase HPLC, dried and re-dissolved in 1× TE buffer. Concentrations were determined by measuring the 340 nm absorbance of the DPI_3 chromophore in the MGB Eclipse probe[6] and were formulated as a 20× solution (4 μM of each probe). Primers were formulated as a 20× mixture composed of an excess primer (40 μM) and a limiting primer (2 μM) to ensure the excess synthesis of the target strand for the complementary MGB Eclipse probe.

Probe and primer design

MGB Eclipse probes and primers were designed to SNP-containing target sequences (150 bp up and downstream of the SNP of interest) using the MGB Eclipse™ Design Software 3.0 (www.epochbio.com/products/MGBEclipse_Software.htm). This program utilises thermodynamic parameters and nearest neighbour parameters that were determined for the MGB moiety, the Eclipse Dark Quencher and all the modified bases. The software was run in the Express Mode, and in each case the best choice was used for further study.

Templates

One hundred and two unrelated Centre d'Etude du Polymorphisme Humain (CEPH) DNA samples were obtained from the Coriell Institute of Medical Research (http://locus.umdnj.edu/) after specifying that the DNA samples were to be used for research purposes only. A list of the templates used is available at http://snp500cancer.nci.nih.gov.

Real time PCR using MGB Eclipse probes

Real time PCR was conducted on either an ABI Prism® 7900 Sequence Detection System (SDS) (Applied Biosystems, Foster City, CA), or on a MJ Research PTC-200 Peltier Thermal Cycler (Waltham, MA).[10] On both instruments, 50 cycles of a three-step PCR (95°C for five seconds, 58°C for 20 seconds and 76°C for 30 seconds) profile were run after an initial 2 minutes at 95°C. If necessary, fluorescent data were collected at 58°C with an ABI 7900 SDS. Commercially available 2× Jump Start™ Taq Ready Mix™ for Quantitative PCR with 2 mM final Mg^{++} concentration (Sigma Catalog #D 74403) supplemented with JumpStart Taq Polymerase (Sigma Catalog #90 4184) to a final amount of 0.37 U/μl was used. The final concentration of both probes was 0.2 μM; the concentration of limiting primer was 0.1 μM and excess primer was 2 μM. Each 5 μl reaction solution contained 10 ng of genomic DNA lyophilised in 96- or 384-well plates with a SPD 1010 SpeedVac® (ThermoSavant, NY) prior to reaction set-up. Routinely, 102 CEPH DNA samples were tested in triplicates using a 384-well plate. A Biomek® 2000 Laboratory Automation Station (Beckman Coulter, USA) was used to set up PCR reactions.

Genotyping analysis using fluorogenic melt curves and MGB Eclipse probes

After completion of PCR, the plate was transferred to the ABI PRISM 7900 SDS (if PCR was performed in an MJ Research Cycler), and the instrument was set for dissociation curve analysis using fluorescein and tetrachlorofluorescein detection. The thermal profile was set using an initial denaturing temp of 95°C for 30 seconds, an annealing temp of 30°C and a final temperature of 80°C. The ramp rate was set to 10 per cent, which is ~1 degree per ten seconds. After reaching the final temperature, the collected data were saved for analysis. Melt curve data can be graphically visualised as the first derivative of the melt curve over temperature. For automated genotyping, the data are exported to Microsoft® Excel (Microsoft Corporation, Redmond, WA) and analysed using the MGB Eclipse Melt Macro available from Epoch Biosciences (www.epochbio.com).

Results and discussion

Design of MGB Eclipse probes and primers

SNP detection using fluorogenic probes and PCR requires robust amplification and exquisite ability to distinguish between alleles. Unfortunately, many interesting and important SNPs can be buried in folded or aggregated regions of DNA, making analysis problematic.[11] G-C-rich regions of DNA are especially difficult to amplify and probe, as G bases are prone to forming secondary structures such as G-tetrads (Hoogsteen base pairing) and parallel stranded duplexes.[12–15] Hybridisation assays can be carried out at higher temperatures to melt out complex secondary structures, but this requires probes that bind to the desired DNA strands at these elevated temperatures and often results in mismatches not being detected such as the T-G mismatch (the hardest mismatch to detect). Long probes (20–30 nt) can be used for A-T-rich sequences, but these frequently require very stringent hybridisation conditions to allow the detection of SNPs in DNA duplexes with very small melting temperature differences.[9,16] We have previously shown that standard probes lacking the MGB moiety don't discriminate well, if at all, if the SNP is located in regions with high AT content.[9]

The use of fluorogenic MGB probes has helped solve the problem of genotyping A-T-rich sequences and this technology has been adopted for use in both TaqMan MGB and MGB Eclipse Systems. For MGB Eclipse, a probe T_m of 59–63°C (with 2.0 mM Mg^{++} PCR buffer and probe concentrations of 0.2 μM) is used as a default setting in the MGB Eclipse Design Software. This T_m is desirable, since it ensures efficient hybridisation-triggered fluorescence at the

middle of the dissociation temperature range (40–80°C), but does not block PCR at the primer extension temperature of 76°C. As described below, actual T_m — as measured by fluorogenic melting curve analysis — is usually lower than calculated, since the concentration of the single strand DNA (ssDNA) target is variable.

Genotyping assays with fluorogenic MGB probes require robust PCR, so design of primers is also a crucial issue. Primer design software is commercially available, and sophisticated algorithms can be used to avoid coiled DNA targets, self-complementary primers (primer-dimers), long A-T-rich primers or other potential problems. Since PCR efficiency declines with length of amplified sequence,[17] short amplified fragments (amplicons) are desirable. Amplicons of 80–200 bp improve PCR efficiency and thus provide a stronger signal. Optimum performance of MGB Eclipse probes is obtained when there is an excess of the complementary DNA target strand. Probe and primer designs for a typical SNP containing the DNA target (*selectin E* gene, SELE-02; http://snp500cancer.nci.nih.gov) are displayed in Figure 2. Note that a modified G analogue had to be used in one of the probes, to avoid self-association.

Comparison of real time and melting curve SNP assays

The power of a melt curve is in the examination of the signal reported at a range of temperatures for each probe. This allows easy discrimination of the signal for each probe. Any probe bound to a mismatch will necessarily be less thermodynamically stable (usually with T_m about 5–10°C lower) than the same probe bound to a perfect match. By melt curve analysis, a high level of discrimination over a range of temperatures can be utilised to make genotype calls. Genotype calling is highly accurate, since unique melting curve patterns are observed for wild-type, mutant and heterozygous alleles, using two differently labelled MGB Eclipse probes specific for wild-type and mutants, respectively.

The advantage of melting curve analysis over endpoint fluorescence read is that sample-to-sample variations in PCR yield and/or template DNA amount and quality do not affect the T_m of the samples with the same genotype and consequently do not affect genotyping calls. By contrast, endpoint detection makes calls based on final fluorescence intensity after PCR, which strongly depends on the quality of the samples and PCR yield. In some cases, these differences in fluorescence can lead to the wrong genotype call.[10]

The power of genotyping using melting curve analysis with two differently labelled MGB Eclipse probes, specific for each allele, is demonstrated by a comparison of real time PCR and melting curve assay results for the SELE-02 SNP, a C to T substitution, as shown in Figure 3. The assay was tested against 102 individual human genomic DNA samples. As shown, even difficult to detect SNPs can be discriminated by melting curve analysis. Note that the mismatched tetrachlorofluorescein-labelled MGB Eclipse probe hardly shows up in the melting curve (Figure 3d), since the hybrid melts at such a low temperature. The mismatched fluorescein-labelled MGB Eclipse probe does show up (T_m difference is 7°C), but is easily discriminated to the perfect match (Figure 3b). By contrast, real time PCR the analysis can be ambiguous, since only the strength of the fluorescent signal is measured. As shown with the tetrachlorofluorescein real time data, the allele 2 and heterozygous curves are close together, which makes genotype assignment difficult (see scatter plot in Figure 3e, samples 1 and 2). The melting curves, however, have 'match' T_ms for the questionable samples only in the Fb channel, which corresponds, by definition, to a homozygous 'T' genotype. The results were in 100 per cent concordance with sequencing data (http://snp500cancer.nci.nih.gov/snp.cfm).

MGB Eclipse probes with modified A and T bases for genotyping A-T-rich targets

As described above, MGB Eclipse probes dramatically improve genotyping of A-T-rich targets, since the MGB

GTAAATTTTCTACGGTAGGTGCCAGCCACAGCTATACATTAGCATCAAGGTTTAGGATAGG
TGGGCCTGAAACTCAGATTTGGAACTTTGTAAAACACTTCAGTTAAAGTCTATGGCACTCTG
TAGGACTGCTGATGTCTCTGTTGCACACTGTTGATTCTAAAATCAAAGGCACTCAGTATAAG
CACAGTAATAGTCCTCC[C/T]CATCATGCTTTGTATTTTCCGTAGCTGCCTGTACCAATACAT
CCTGCMGTGGCCACGGTGAATGTGTAGAGACCATCAATAATTACACTTGCAAGTGTGACCC
TGGCTTCAGTGGACTCAAGTGTGAGCAAAGTAAGTCTGGTTCTYGCCTCTTTCTTCACTTGA
GATGGTAGCACCATCTCACGTCCTAGCTGGC

Forward primer AGG CAG CTA CGG AAA ATA CAA A
Reverse primer GAT GTC TCT GTT GCA CAC TGT T
Probe 1 MGB-Q-TGA TG**G** gGA GGA CT-Fa
Probe 2 MGB-Q-ATG ATG **A**GG AGG ACT-Fb

Figure 2. Partial sequence of the *selectin E* gene (gi:4506870; *rs3917410* (*SELE-02*). The SNP is shown in square brackets. The primer and probe sequences are underlined and shown below. Q is the Eclipse Dark Quencher. The alleles are in bold and underlined in the probe sequences. 'g' is modified G. Probes are complementary to the shown sequence. Fa and Fb are fluorescein and tetrachlorofluorescein, respectively.

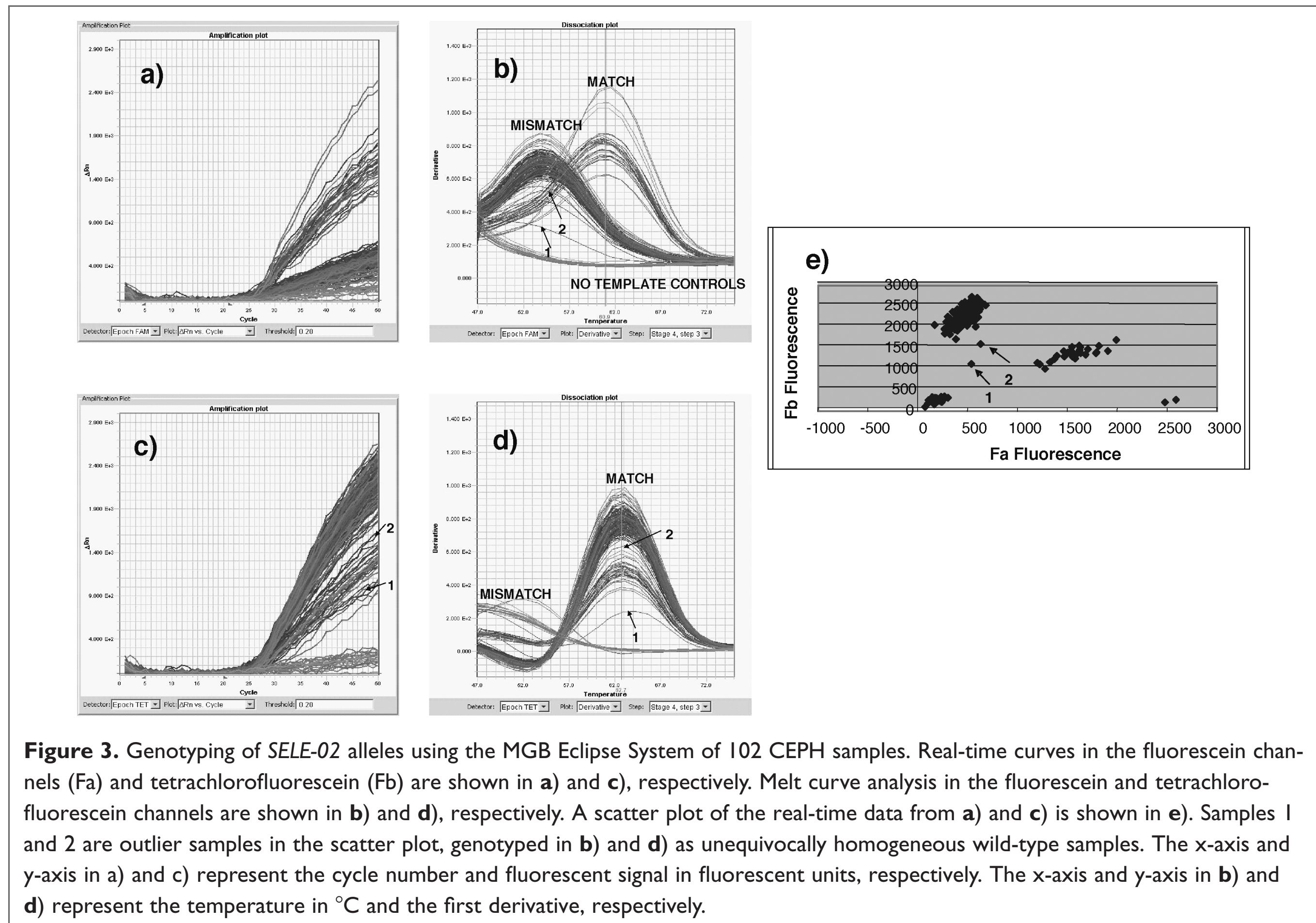


Figure 3. Genotyping of *SELE-02* alleles using the MGB Eclipse System of 102 CEPH samples. Real-time curves in the fluorescein channels (Fa) and tetrachlorofluorescein (Fb) are shown in **a**) and **c**), respectively. Melt curve analysis in the fluorescein and tetrachlorofluorescein channels are shown in **b**) and **d**), respectively. A scatter plot of the real-time data from **a**) and **c**) is shown in **e**). Samples 1 and 2 are outlier samples in the scatter plot, genotyped in **b**) and **d**) as unequivocally homogeneous wild-type samples. The x-axis and y-axis in a) and c) represent the cycle number and fluorescent signal in fluorescent units, respectively. The x-axis and y-axis in **b**) and **d**) represent the temperature in °C and the first derivative, respectively.

probes are shorter and have better mismatch discrimination. Still, very A-T-rich SNP sites can require long (>18 bases) MGB probes for fluorogenic PCR assays. MGB Eclipse probes also have an optimum length requirement beyond that required for mismatch discrimination. We and others have shown that optimal signal and discrimination for the 'hybridisation-triggered fluorescence' mechanism is achieved with probes ~15 bases in length.[9,18] We have largely solved the problem of A-T-rich targets by incorporating modified A and T bases into MGB Eclipse probes and primers. The *NAT1* — N-acetyltransferase 1 (arylamine N-acetyltransferase) gene (http://snp500cancer.nci.nih.gov) is an example of a difficult target; a portion of the 3′-UTR of the NAT1 gene sequence is shown in Figure 4. Attempts to make a working assay without modified bases failed due to poor discrimination (data not shown).

The sequence shown in Figure 4 is especially difficult, since there are two closely positioned SNPs within an 8 bp region of DNA. Since the region is A-T rich, long primers and MGB probes would be required, and overlap of the binding regions would prevent the analysis of each SNP. As shown in Figure 4, using modified A and T in MGB Eclipse probes and primers

ATCCAGCTCACCAGTTATCAACTGACGACCTATCATGTATCTTCTGTACCCT
TACCTTATTTTGAAGAAAATCCTAGACATCAAATCATTTCACCTATAAAAAT
GTCATCATATATAATTAAACAGCTTTTTAAAGAAACATAACCACAAACCTTT
TCAAATAATAATAATAATAATAATAA[A/T]AAATGT[A/C]TTTTAAAGAT
GGCCTGTGGTTATCTTGGAAATTGGTGATTTATGCTAGAAAGCTTTTAATGT
TGGTTTATTGTTGAATTCCTAGAAAAGTTTTATTGGTAGATGAGTAAATAAA
ATATTGTAAAAAAACTTATTGTCTATAAAGTATATTAAAACATTGTTGGCTA
ATATAATTTGAAAAAAAGTGGTTTTTTGGAAGACTTAGGATATTATGGTGCT
ACATAATTTTTCCTCGATGCTCTCT

SNP1 [A/T]
Forward primer CACCAATTTCCAAGATAACCACAGG
Reverse primer tAGACAtCAAATCattTCACCTA
Probe 1 MGB-Q-ACATTT**t**TtattaT-Fa
Probe 2 MGB-Q-ACATTT**A**TtattaT-Fb

SNP2 [A/C]
Probe 1 MGB-Q-CtttAAAA**G**ACattt-Fa
Probe 2 MGB-Q-CtttaAAA**T**ACattt-Fb

Figure 4. A portion of the *NAT-1* sequence (gi:27754152). The two SNPs are shown in square brackets. The underlined primer sequences are the same for both SNPs and are shown below with corresponding probes. Small letters represent modified bases. Design was made for the antisense strand. Alleles are bold in the probe sequences.

allowed two SNP systems to be successfully prepared without overlap of the adjacent SNPs. MGB Eclipse probes to the first [A/T] SNP (Figure 5a) were successfully designed using two modified A and four modified T bases, thus shortening the probes to 14 bases and preventing overlap with the second [A/C] SNP (Figure 5b). MGB probes to the second SNP used two modified A and six modified T bases, thus shortening the probes to 15 bases. As can be seen by the fluorogenic melt curves, both assays can be used to discriminate between homozygous and heterozygous samples. The data clearly show that genotyping of both of these difficult SNPs is possible. (Figure 5).

MGB Eclipse probes with modified G bases for genotyping G-C rich targets

As described earlier, G-C-rich regions of DNA are especially difficult to amplify and probe, since G bases are prone to forming secondary structures such as G-tetrads and parallel stranded duplexes.[12–15] We have reported that formerly inaccessible G-rich probes can be easily prepared using modified G, an analogue of the naturally occurring dG base. Modified G (Super G, also called pyrazolopyrimidine G, PPG, or 8-aza-7deaza-guanine) hybridises with a similar affinity to natural dG, but shows little propensity to form unwanted secondary structures or aggregates.[19] Many of these 'non-ideal' complexes can arise from hydrogen bonding (Hoogsteen bonds) to the N-7 position of guanine. Modified G removes these complexes by exchanging the nitrogen at the 7-position with the carbon at the 8-position of guanine.

The major advantage of modified G in MGB Eclipse probe design is that previously inaccessible G-rich probes can be used to interrogate C-rich targets. Normally, C-rich probes are used to analyse G-C-rich targets, but this can be problematic, since the complementary G-rich target strand

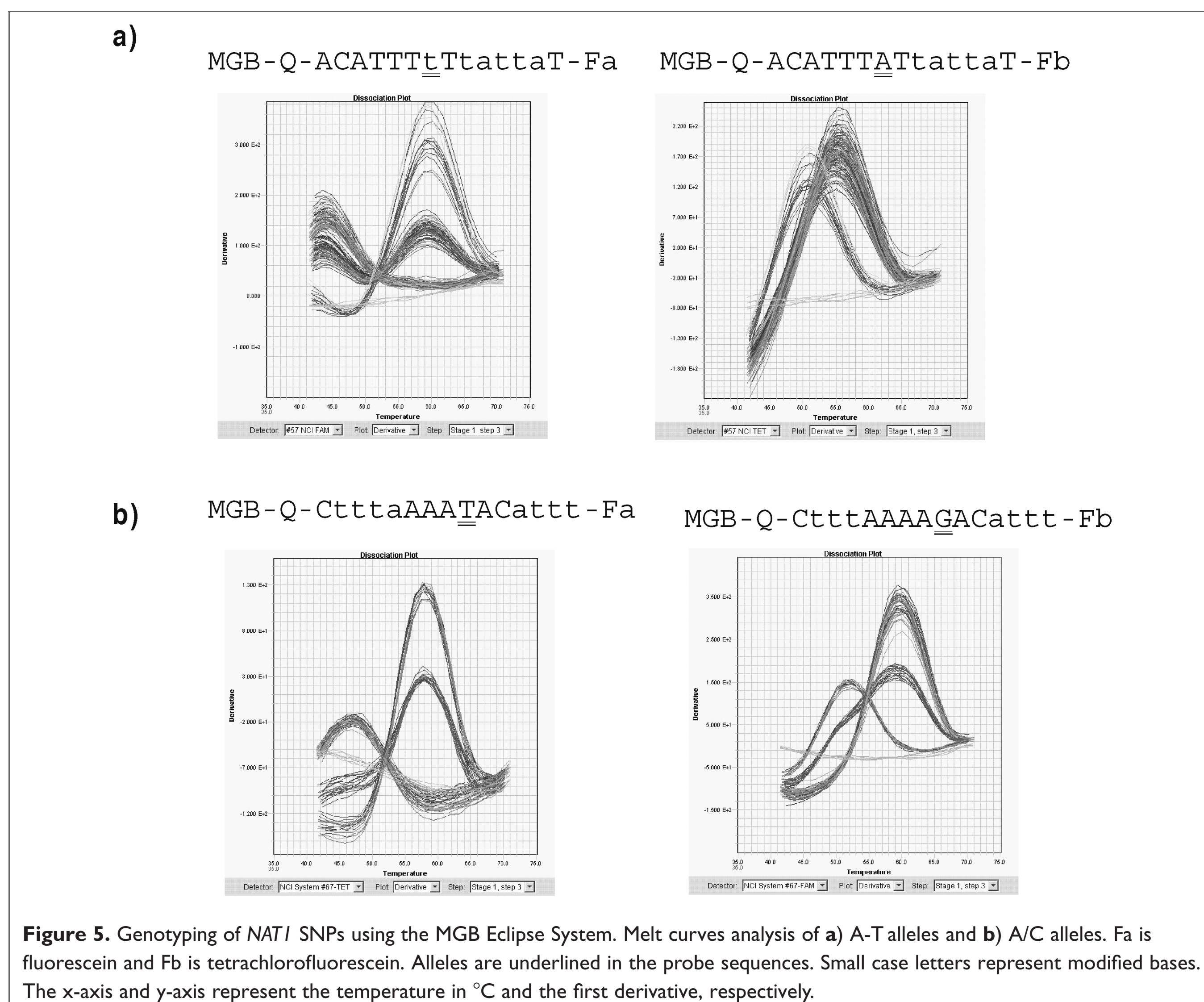


Figure 5. Genotyping of *NAT1* SNPs using the MGB Eclipse System. Melt curves analysis of **a**) A-T alleles and **b**) A/C alleles. Fa is fluorescein and Fb is tetrachlorofluorescein. Alleles are underlined in the probe sequences. Small case letters represent modified bases. The x-axis and y-axis represent the temperature in °C and the first derivative, respectively.

can fold or aggregate, thus increasing the assay failure rate. The ability to probe C-rich targets with G-rich probes thus improves assay success rate. An example of a G-C-rich target sequence where the use of modified G is required is presented in Figure 6a. An example of one such problematic sequence is found in the *RAD23B* gene in IVS 5 (http://snp500cancer.nci.nih.gov), since this SNP is located in a DNA region where 12 out of 15 bases are G-C. Choosing the C-rich strand as the target requires the use of extremely G-rich MGB Eclipse probes. The successful designs shown in Figure 6 used probes with a G-content of 11/14 bases (fluorescein probe) and 9/14 bases (5′-tetrachlorofluorescein probe). By substituting two or three modified G bases into these MGB Eclipse probes, DNA synthesis was successful and this difficult SNP was genotyped, as shown in Figures 6b and 6c. Attempts to make a working assay without modified G substitution failed.

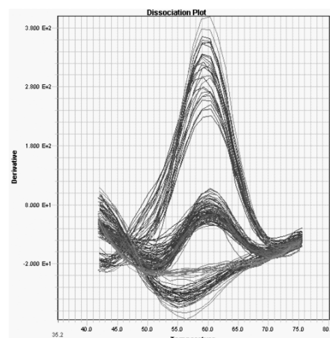
Automated genotyping by melting curve analysis on the ABI PRISM 7900

Automated melting curve analysis was successfully applied to high throughput detection of more than 200 human SNPs. A panel of 102 human genomic DNAs was challenged with each system, as described above. A 100 per cent call rate and 100 per cent concordance with sequencing data was obtained (http://snp500cancer.nci.nih.gov). Typically, after completion of the PCR amplification, a melt curve is performed. After the resultant melt data are imported into Excel, the MGB Eclipse Melt Macro, probe T_m and signal threshold values are automatically determined from the melting curve data (in the absence of the manual user intervention described below), and genotypes called and graphed for each sample. The vertical lines z_1 and z_2 in Figures 7a and 7b represent the T_ms for the A and C alleles, respectively. Similarly, the signal thresholds in each channel are represented as double horizontal lines y_1 and y_2 in the fluorescein and tetrachlorofluorescein channels, respectively. The macro looks in the fluorescein channel for the highest fluorescent signal, which represents the T_m of the probe complementary to the wild-type homozygous signal, indicated as a vertical line (z_1) in Figure 7a. An acceptable fluorescein T_m window is set automatically with the default setting of $\pm 3°C$ T_m. A signal in this window is assigned a value of 1. Any signal outside of this window is assigned a value of zero. Similarly, in the tetrachlorofluorescein channel, the macro looks for the highest signal which represents the T_m of the homozygous mutant allele, shown as a vertical line (z_2) in Figure 7b. Again, an acceptable tetrafluorescein T_m window is set automatically as $\pm 3°C$ T_m. A signal in this window is also

a)

TATCTATAACCGCAATTTCCCTTTGCCTTGTAACCTATTTATAGGTT
CCAGGGATCAAACCAGAGGGCATCTTTGGGGGACCATTACCTACT
ACATATATGTTGTAATTTATACTGTTACTCATCTTTGTATTCCCAGC
ATAGTAGTWCCTGAAATGTTGTATACATGAATCARTAAATGTAWAG
AGAATGCTTATTTATTAAATGTTCCTTTTTTTCCCCTCC(A/G)CCCTC
CCTTTTTAGTGACGGGTCAGTCTTACGAGAATATGGTAACTGAGAT
CATGTCAATGGGCTATGAACGAGAGCAAGTAATTGCAGC

b) MGB-Q-GgGAGgGTGGAGgG-Fa



c) MGB-Q-AAGgGAGgGCGGAG-Fb

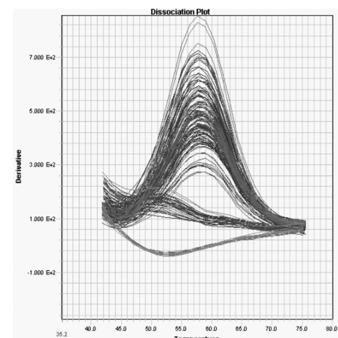
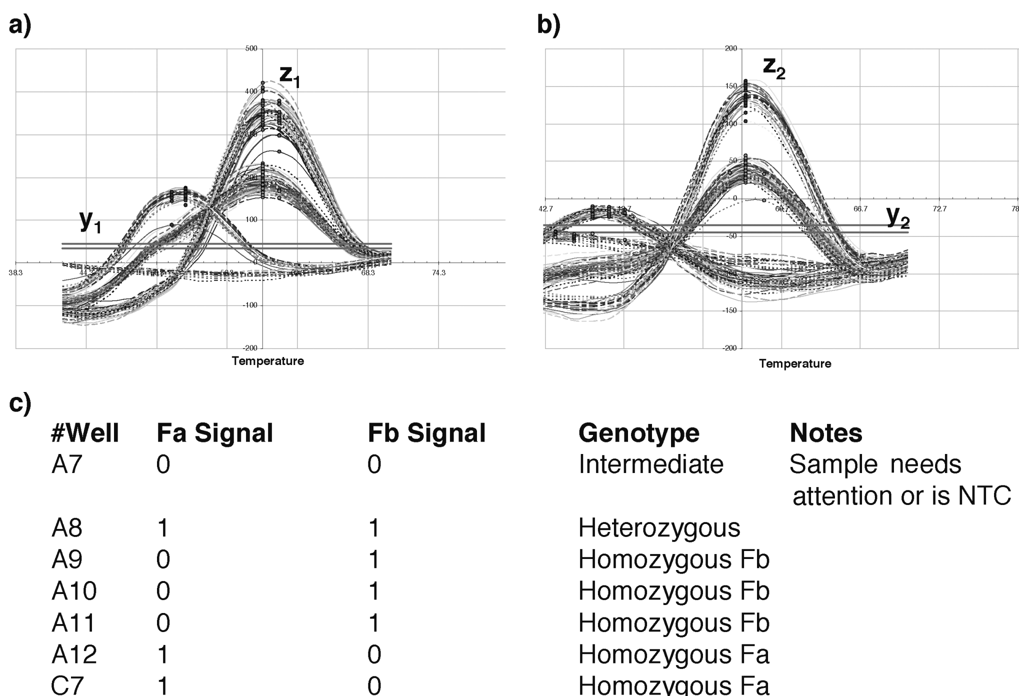


Figure 6. a) A portion of the sequence of the RAD23B-02 SNP (gi:12871592). The SNP is shown in parenthesis. Primer sequences are underlined. **b**) and **c**) MGB Eclipse data with two probes labelled with fluorescein specific for the A-allele and 5′-tetrachloro-fluorescein specific for the G-allele, respectively. The alleles are underlined in the probe sequences. 'g' is modified G. The x-axis and y-axis represent the temperature in °C and the first derivative, respectively.



#Well	Fa Signal	Fb Signal	Genotype	Notes
A7	0	0	Intermediate	Sample needs attention or is NTC
A8	1	1	Heterozygous	
A9	0	1	Homozygous Fb	
A10	0	1	Homozygous Fb	
A11	0	1	Homozygous Fb	
A12	1	0	Homozygous Fa	
C7	1	0	Homozygous Fa	

Figure 7. Semi-automated genotyping using the MGB Eclipse Melt Macro and genotyping of *NAT-1* A/C alleles. **a**) Shows the melting curves in the fluorescein channel and **b**) shows the melting curves in the tetrachlorofluorescein channel. The vertical lines z_1 and z_2 are manually set for the T_ms for the A and C alleles, respectively. Similarly, the thresholds in each channel are manually set and are respectively set as double horizontal lines y_1 and y_2 in the fluorescein and tetrachlorofluorescein channels, respectively. **c**) A subset of automated genotyped alleles from a 384-well plate is listed as an example. Sample in well A7 is no template control (NTC); samples in other wells contain different human genomic DNAs.

assigned a value of 1, while any signal outside the window is set as a zero value. The melt curves of samples can now be typed automatically, as shown in Figure 7c.

A sample with a value of 1 only in the fluorescein channel is typed as a homozygous allele Fa, while a value of 1 in only the tetrachlorofluorescein channel represents a homozygous allele Fb. Values of 1 in both channels represent a heterozygous sample (Figure 7c). A zero value in both channels may be a no template control (NTC), as is the case in Figure 7 (well A7) or indicates that the sample needs manual attention, as PCR may have failed. The macro normally sets probe T_m and signal threshold parameters automatically.

Acknowledgments

This project has been funded in part with Federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. NO1-CO-12400 and the National Institute of Allergy and Infectious Diseases, Grant Number 1R43-A1052905-01A1. The content of this manuscript does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organisations imply endorsement by the US Government. The publisher or recipient acknowledges the right of the US Government to retain a non-exclusive, royalty-free licence in and to any copyright covering the article. Michael W. Reed is acknowledged for his assistance in the preparation of the manuscript. We also thank Vladimir Gorn, Irina Shishkina and Zinaida Sergueeva for the synthesis and purification of the probes and primers.

References

1. Didenko, V.V. (2001), 'DNA probes using fluorescence resonance energy transfer (FRET): Designs and applications', *Biotechniques* Vol. 31, pp. 1106–1121.
2. Shi, M.M. (2002), 'Technologies for individual genotyping: Detection of genetic polymorphisms in drug targets and disease genes', *Am. J. Pharmacogenomics* Vol. 2, pp. 197–205.
3. Watt, J.R. and Davis, P.W. (1996), 'Kinetics of G-quartet-mediated tetramer formation', *Biochemistry* Vol. 15, pp. 8002–8008.
4. Peyret, N., Seneviratne, P.A., Allawi, H.T. and SantaLucia Jr., J. (1999), 'Nearest-neighbor, thermodynamics and NMR of DNA sequences with internal A·A, C·C, G·G and T·T mismatches', *Biochemistry* Vol. 38, pp. 3468–3477.
5. Mohanty, D. and Bansal, M. (1994), 'Conformational polymorphism in telomeric structures: Loop orientation and interloop pairing in d(G4TnG4)', *Biopolymers* Vol. 34, pp. 1187–1211.
6. Afonina, I.A., Reed, M.W., Lusby, E. *et al.* (2002), '5′ Minor groove binder-conjugated DNA probes for quantitative DNA detection by hybridization-triggered fluorescence', *BioTechniques* Vol. 32, pp. 940–949.
7. de Kok, J.B., Wiegerinck, E.T., Giesendorf, B.A. and Swinkels, D.W. (2002), 'Rapid genotyping of single nucleotide polymorphisms using novel minor groove binding DNA oligonucleotides (MGB probes)', *Hum. Mutat.* Vol. 19, pp. 554–559.

8. McGuigan, F.A.A. and Ralston, S.H. (2002), 'Single nucleotide polymorphism detection: Allelic discrimination using Taqman', *Psychiatr. Genet*. Vol. 12, pp. 133–136.
9. Kutyavin, I.V., Afonina, I.A., Mills, A. *et al*. (2000), '3′-Minor groove binder-DNA probes increase sequence specificity at PCR extension temperatures', *Nucleic Acids Res*. Vol. 28, pp. 655–661.
10. Afonina, I., Belousov, Y., Metcalf, M. *et al*. (2003), 'Single nucleotide polymorphism detection with MGB Eclipse™ assays', *J. Clin. Ligand Assay* Vol. 25(3), pp. 268–275.
11. Southern, E.M., Case-Green, S.C., Elder, J.K. *et al*. (1994), 'Arrays of complementary oligonucleotides for analyzing the hybridization behavior of nucleic acids', *Nucleic Acids Res*. Vol. 22, pp. 1368–1373.
12. Henderson, E., Hardin, C.C., Walk, S.K. *et al*. (1987), 'Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine–guanine base pairs', *Cell* Vol. 51, pp. 899–908.
13. Marsh, T.C., Vesenka, J. and Henderson, E. (1995), 'A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy', *Nucleic Acids Res*. Vol. 23, pp. 696–700.
14. Lu, M., Guo, Q. and Kallenbach, N.R. (1993), 'Thermodynamics of G-tetraplex formation by telomeric DNAs', *Biochemistry* Vol. 32, pp. 598–601.
15. Hardin, C.C., Watson, T., Corregan, M. and Bailey, C. (1992), 'Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG)', *Biochemistry* Vol. 31, pp. 833–841.
16. Kutyavin, I.V., Lukhtanov, E.A., Gamper, H.B. and Meyer, R.B. (1997), 'Oligonucleotides with conjugated dihydropyrroloindole tripeptides: Base composition and backbone effects on hybridization', *Nucleic Acids Res*. Vol. 25, pp. 3718–3723.
17. Applied Biosystems, ABI PRISM® 7900HT Sequence Detection User Guide. Foster City, CA 2002, USA: Applied Biosystems.
18. Parkhurst, K.M. and Parkhurst, L.J. (1996), 'Detection of point mutation in DNA by fluorescence energy transfer', *J. Biomed. Optics* Vol. 1, pp. 435–441.
19. Kutyavin, I.A., Lokhov, S.G., Afonina, I.A. *et al*. (2002), 'Reduced aggregation and improved specificity of G-rich oligodeoxyribonucleotides containing pyrazolo[3,4-d]pyrimidine guanine bases', *Nucleic Acids Res*. Vol. 30, pp. 4952–4959.