University of Utah Institutional Repository
Author Manuscript

Heterozygote PCR Product Melting Curve Prediction

Zachary L. Dwight ZACH.DWIGHT@PATH.UTAH.EDU
Department of Pathology
University of Utah
30 North 1900 East
Salt Lake City, UT 84132, USA

Robert Palais BOB.PALAIS@UVU.EDU
Department of Mathematics
Utah Valley University
800 West University Parkway
Orem, UT 84058, USA

Jana Kent JANA.KENT@PATH.UTAH.EDU
Department of Pathology
University of Utah
30 North 1900 East
Salt Lake City, UT 84132, USA

Carl T. Wittwer CARL.WITTWER@PATH.UTAH.EDU
Department of Pathology
University of Utah
30 North 1900 East
Salt Lake City, UT 84132, USA

Abstract

Melting curve prediction of PCR products is limited to perfectly complementary strands. Multiple domains are calculated by recursive nearest-neighbor thermodynamics. However, the melting curve of an amplicon containing a heterozygous single nucleotide variant (SNV) after PCR is the composite of four duplexes: two matched homoduplexes and two mismatched heteroduplexes. To better predict the shape of composite heterozygote melting curves, 52 experimental curves were compared to brute force *in silico* predictions varying two parameters simultaneously: the relative contribution of heteroduplex products and an ionic scaling factor for mismatched tetrads. Heteroduplex products contributed 25.7 +/- 6.7% to the composite melting curve, varying from 23-28% for different SNV classes. The effect of ions on mismatch tetrads scaled to 76-96% of normal (depending on SNV class) and averaged 88 +/-16.4%. Based on uMelt (www.dna.utah.edu/umelt/umelt.html) with an expanded nearest neighbor thermodynamic set that

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

includes mismatched base pairs, uMelt HETS calculates helicity as a function of temperature for homoduplex and heteroduplex products, as well as the composite curve expected from heterozygotes. It is an interactive web tool for efficient genotyping design, heterozygote melting curve prediction, and quality control of melting curve experiments. The application was developed in Actionscript and can be found online at http://www.dna.utah.edu/hets/.

Keywords: PCR, High Resolution Melting, Prediction, Software, Heterozygote, SNV

1 Introduction

High resolution melting analysis (HRMA) is a molecular diagnostic technology for genetic scanning and genotyping (Wittwer, 2010). Prediction of PCR product melting curves allows for better amplicon and primer design. Melting curve results can also be compared and validated against a theoretical model of DNA dissociation. Unfortunately, current models are only valid for perfectly-matched homoduplex products and do not account for the contribution of heteroduplexes that arise from heterozygotes. For kinetic reasons during hybridization and melting of heterozygotes (Gundry et al., 2003), the two heteroduplex products contribute less than the 2 homoduplex products to the composite heterozygote melting curves (Palais et al., 2005). uMelt HETS predicts the shape of melting curves for heteroduplex and homoduplex products, as well as the experimental composite heterozygote melting curve. The program implements an expanded nearest neighbor parameter set that includes mismatches (SantaLucia and Hicks, 2004), duplex-to-composite contribution factors, and recursive stability calculations requiring partition functions to calculate melting curves of both homoduplex and heteroduplex products.

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

2 Methods

uMelt HETS incorporates nearest neighbor thermodynamics. Similar to the original uMelt web application (Dwight et al., 2011), the standard 10 nearest neighbor enthalpy (ΔH) and entropy (ΔS) parameters were used with monovalent cation and Mg^{++} corrections. In addition, the thermodynamic values of mismatches (SantaLucia and Hicks, 2004) were included, a key structural difference between homoduplexes and heteroduplexes. The thermodynamic stabilities of tetrads were calculated and used recursively as described previously. To identify contributions of heteroduplex products, as well as the ionic effects of tetrads containing a mismatch duplex, brute force *in silico* comparisons were performed. Small amplicons bracketing a set of 52 forensic SNV sequences (Sanchez et al., 2006) containing all four SNV Classes (Liew et al., 2004) were used to compare experimental to calculated melting curves. For each of the 52 forensic SNVs, reactions were performed in triplicate. The triplicates were then averaged to identify a single representative curve for each SNV for comparison. All experimental melting curves were normalized from 0-100% helicity using the exponential method (Palais and Wittwer, 2009). PCR primer pairs are listed in Supplementary Table 1. Each 10 µl reaction contained 50 mM Tris-HCl pH 8.3, 2.0 mM Mg^{2+} (1.2 mM free Mg^{2+}), 200 µM each deoxynucleoside triphosphate (800 µM combined), 1X LCGreen+ (BioFire Diagnostics), 0.4 U KlenTaq1 (Ab Peptides) with 64 ng anti-Taq antibody (eEnzyme), 5 µg BSA, 0.5 µM primers and 50 ng genomic DNA. For each of the 52 SNVs, heterozygote melting curve predictions were generated by varying the collective heteroduplex contribution from 0-60% by 5% increments and the ionic scaling factor for mismatches from 40-100% by 2% increments, resulting in 360 melting curves. The area between the predicted and experimental curves, calculated as the 2D offset (Dwight,

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

2012) was plotted against the collective heteroduplex percentage and the ionic scaling factor. This identified a narrower range along each axis for a high-resolution minimization of the 2D offset from ~1000 *in silico* melting curves. Parameter minimization was also performed after shifting the predicted curves along the temperature axis (Gundry et al., 2003) to best overlay the experimental melting curves across their homoduplex portion (5-15% fluorescence). An example of the curve prediction and experimental comparison process is shown in Figure 1.

The optimized results of computational comparisons were implemented in an interactive and flexible web application. The user sequence and SNV are defined in a text box and a drop down selector populated with the alternate SNV bases. Melting curve predictions are provided in two tabs: genotype predictions and individual duplexes. The default view of the user interface (Figure 2) shows the product melting curves as genotypes (what is seen experimentally) and the alternate view (on a secondary tab) shows the individual contributing duplexes (Supp. Figure 1). Data can be downloaded as a text file and chart images can be exported in (.png) format. The application allows inputs of monovalent cation, free magnesium, and DMSO concentrations. Data resolution of the prediction can also be adjusted, along with the temperature range, to better represent the spectrum of high-resolution melting instruments that differ in the density of data points collected.

3 Results and Discussion

The area between predicted and experimental heterozygote melting curves was minimized by simultaneously optimizing 2 prediction parameters: the percent contribution of heteroduplexes to the composite melting curve and a cation scaling factor

UU IR Author Manuscript
UU IR Author Manuscript

INSTITUTIONAL REPOSITORY THE UNIVERSITY OF UTAH
University of Utah Institutional Repository
Author Manuscript

applied to mismatched tetrads. Absolute Tm differences were removed by overlay to focus on curve shape and required an average temperature shift of 2.8 +/- 0.8 °C. The discrepancy in Tm is larger than the reported average Tm difference (1.6 °C) for the nearest-neighbor parameter set used (SantaLucia and Hicks, 2004). Possible reasons include the absence of an adjustment for dye and the recursive nature of the uMelt algorithm, which becomes more effective with longer amplicons. The 52 SNV training set averaged ~45bp in length, resulting in less complex melting transitions with an enhanced focus on heteroduplex contributions. Curve overlay allows the simulations to focus on shape dominated by the heteroduplex contribution of the composite prediction. Initial analysis on 52 curve sets assuming a heteroduplex contribution of 50% and no ionic scaling gave areas of 53.9 +/- 32.4 (% Helicity x °C). As shown in Figure 3, the heteroduplex contribution shifts the melting region on the helicity axis and the ionic scalar shifts the heteroduplex region along the temperature axis.

After area minimization, the average heteroduplex contribution for all SNVs was 28.4 +/- 8.0 (% Helicity x °C) rather than 50% confirming a difference between homoduplex and heteroduplex formation and/or dissociation. The thermodynamic stabilities of the products and instrument melting rates determine melting behaviors and thus shape of a melting profile (Gundry et al., 2003) and would only contribute equally (50% homoduplex and 50% heteroduplex) if all duplexes were found in equal concentration. The other adjustment required to increase the fit between predicted and experimental curves was the magnitude of the ionic correction applied to the entropy of mismatch tetrads (SantaLucia et al, 2004). Across all 52 SNVs, experimental and predicted curve

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

shapes were closest when the ionic factor for mismatched tetrads was 88 +/- 16.4 % of matched tetrads. The thermodynamics of mismatched tetrads appear to be less influenced by ions than matched tetrads. If the areas are minimized using SNV Class specific parameters (Table 1), a further area decrease to 25.7 +/- 12.3 (% Helicity x °C) results. This reduction in area is statistically significant compared to the initial analysis using 50% heteroduplex contribution and no ionic correction ($p<0.001$ calculated by the two-tailed Mann-Whitney *U* test).

Interestingly, heteroduplex contributions differed between SNV classes (Table 1) revealing the highest total heteroduplex contribution in Class 4 (28.3%) followed by Class 1 (25.4%), Class 3 (25.3%), and Class 2 (23.8%). Ionic scaling factors also depended on SNV Class. Each class contains distinct tetrad groupings that produce differing thermodynamic stabilities (Liew et al., 2004), which can be more or less affected by the addition of ionic components to the buffer solution. The resulting ionic scaling factors indicate varying levels of sensitivity to cations due to the tetrad grouping and mismatches present. Class 3 was most sensitive to cations (scaling factor of 77%) followed by Class 1 (86%), Class 2 (91%) and Class 4 (97%).

Understanding and controlling the factors that influence PCR product melting curves amplified from heterozygotes is critical for scanning and genotyping assays that detect and identify different variants (Palais et al., 2005). Early work focused on optimizing heteroduplex formation (Jensen and Straus, 1993; Ruano and Kidd, 1992) analyzed on gels. Recent HRMA work shows that COLD-PCR can increase heteroduplex amplification for improved sensitivity (Daniels et al., 2012; Milbury et al. 2009) and

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

curve shape is dependent on the normalization method used (Dwight et al., 2012). Two important factors were identified and studied here: the relative heteroduplex contribution and the magnitude of the cation influence on entropy. Future challenges for prediction include multiple domains and/or multiple variants in a single PCR product, and the effects of buffer components. As shown in Figure 4, melting curve shapes are in good visual agreement with four random SNVs outside of the training set, but more samples would be necessary for a full validation set and statistical analysis. Buffer solutions containing a multitude of components may have competing effects and further complicate reducing the variation between experimental results and theoretical prediction. One potential improvement would be to use thermodynamic parameters developed for PCR conditions (cations and dyes), rather than to use generic correction factors for these components. Further, dye specific contributions are worth investigating as several differences dyes have been used for melt curve analysis. Nevertheless, this study is the first to predict melting curves of heterozygote PCR products. The software, uMelt HETS, is provided publicly without registration as a user-friendly web application with optimized parameters for each of the 4 different SNV Classes.

Acknowledgement
Canon US Life Science and BioFire Diagnostics provided financial support. We would like to thank Alexa Barnes for quality assurance and additional programming support. We would also like to thank Ling Xu from Canon US Life Science for her comparison data.

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

Conflicts of Interest

Carl Wittwer serves as Chairman of the Board for Biofire Diagnostics (Salt Lake City, UT) and holds equity interest. No other conflicts exist.

References

Daniels R, Ndiaye D, Wall M, McKinney J, Sene PD, Sabeti PC, Volkman SK, Mboup S, Wirth DF. 2012. Rapid, field-deployable method for genotyping and discovery of single-nucleotide polymorphisms associated with drug resistance in *Plasmodium falciparum*. *Antimicrob Agents Chemother*. **56**(6): p. 2976-86.

Dwight Z, Palais R, Wittwer CT. 2011. uMELT: Prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application, *Bioinformatics*, **27**, 1019-1020.

Dwight ZL, Palais R, Wittwer CT. 2012. uAnalyze: Web-Based High Resolution DNA Melting Analysis with Comparison to Thermodynamic Predictions. *IEEE/ACM Trans Comput Biol Bioinform.* **9**, 1805-1811.

Gundry CN, Vandersteen JG, Reed GH, Pryor RJ, Chen J, Wittwer CT. 2003. Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clin Chem*. **49**, 396-406.

Jensen, M.A. and Straus. N. 1993. Effect of PCR conditions on the formation of heteroduplex and single-stranded DNA products in the amplification of bacterial ribosomal DNA spacer regions. *PCR Methods Applic*. **3**, 186-94.

Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, Wittwer C. 2004. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem*, **50**, p. 1156-64.

Milbury CA, Li J, Makrigiorgos GM. 2009, COLD-PCR-enhanced high-resolution melting enables rapid and selective identification of low-level unknown mutations. *Clin Chem*, **55**(12): p. 2130-43.

Palais RA, Liew MA, Wittwer CT. 2005. Quantitative heteroduplex analysis for single nucleotide polymorphism genotyping, *Anal Biochem*, **346**, p. 167-75.

Palais R, Wittwer CT. 2009, Mathematical algorithms for high-resolution DNA melting analysis. *Methods Enzymol*, **454**: p. 323-43.

Ruano G, Kidd KK. 1992. Modeling of heteroduplex formation during PCR from mixtures of DNA templates. *PCR Methods Applic*. **2**, 112–116.

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

Sanchez JJ, Phillips C, Borsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedoa A, *et al.* 2006. A multiplex assay with 52 single nucleotide polymorphisms for human identification, *Electrophoresis*, **27**, 1713-1724.

SantaLucia, J, Hicks D. 2004. Thermodynamics of DNA Structural Motifs, *Annu Rev Biophys Biomol Struct*, **33**, 415-440.

Wittwer, CT. 2010. Making DNA Melting Useful, *Clin Chem*, **56**, 1500-1501.

UU IR Author Manuscript
UU IR Author Manuscript

INSTITUTIONAL REPOSITORY THE UNIVERSITY OF UTAH
University of Utah Institutional Repository
Author Manuscript

Figure Legends

Figure 1. Graphical representation of the fitting procedure shows the steps necessary to minimize the 2D Offset between theoretical and experimental heterozygous melting curves. Panel A : Two experimental homoduplex product curves are imported. Panel B : Two heteroduplex product curves are predicted. Panel C : Using the four products, a heterozygous composite is created using the parameters. Panel D : The predicted curve is overlaid on the experimental curve and 2D Offset is recorded. The smallest 2D Offset is selected as 'best-fit'.

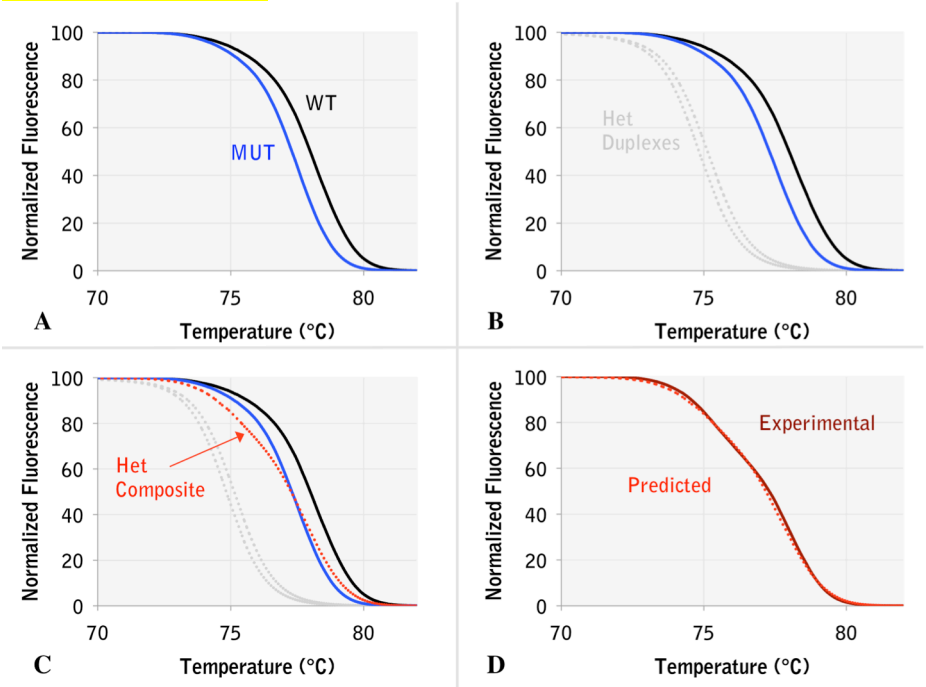


Figure 2. uMelt HETS user interface displaying predicted melting curves and derivative peaks of genotypes.

UU IR Author Manuscript
UU IR Author Manuscript

INSTITUTIONAL REPOSITORY THE UNIVERSITY OF UTAH
University of Utah Institutional Repository
Author Manuscript

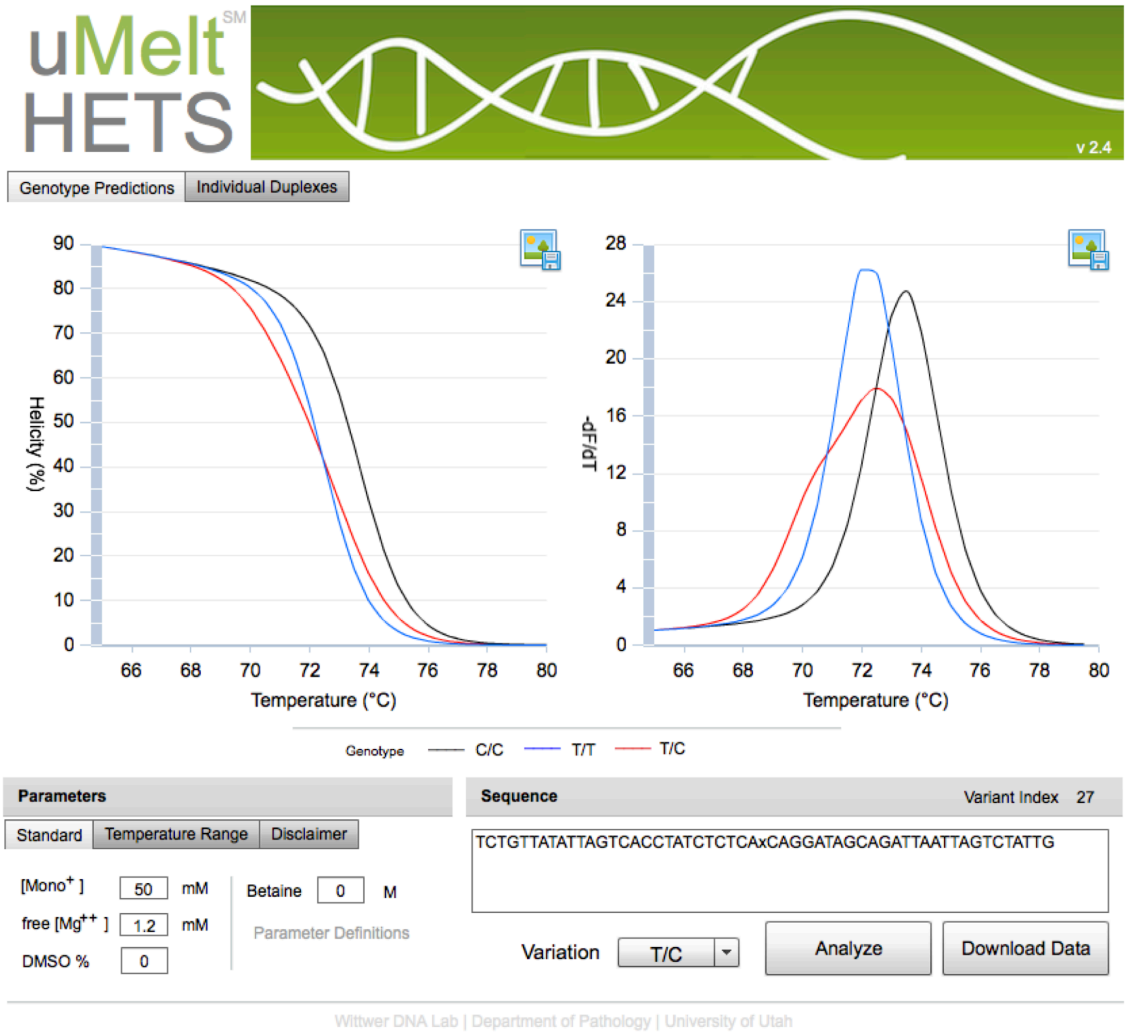
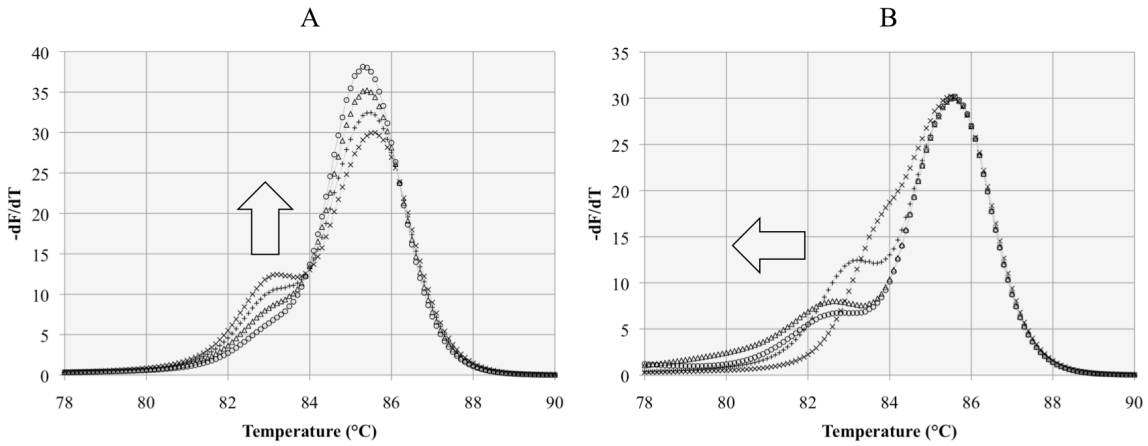


Figure 3. The heteroduplex contribution parameter, increased by 5% from 10%-25%, changes the amplitude but not the position of the heteroduplex melting region (panel A). The cation scalar (%), decreased by 10% from 100% to 70%, moves the predicted heteroduplex melting region across the temperature axis (panel B).



UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

Figure 4. Predicted heterozygote melting curves (grey) and experimental results (black) show good agreement in shape and peak heights for all SNV Classes. Panel A (rs653568) represents Class 1, Panel B (rs1126670) represents Class 2, Panel C (rs644617) represents Class 3, Panel D (rs566476) represents Class 4. Panel A had a 2D Offset (the area between curves) of 37 (% x °C) after overlay, accompanied by Panel B with 44 (% x °C), Panel C with 41 (% x °C), and Panel D with 31 (% x °C). The average absolute helicity difference at each temperature between experimental and predicted curves for Panel A, B, C and D was 9.2%, 4.2%, 4.1%, and was 6.6% respectively.

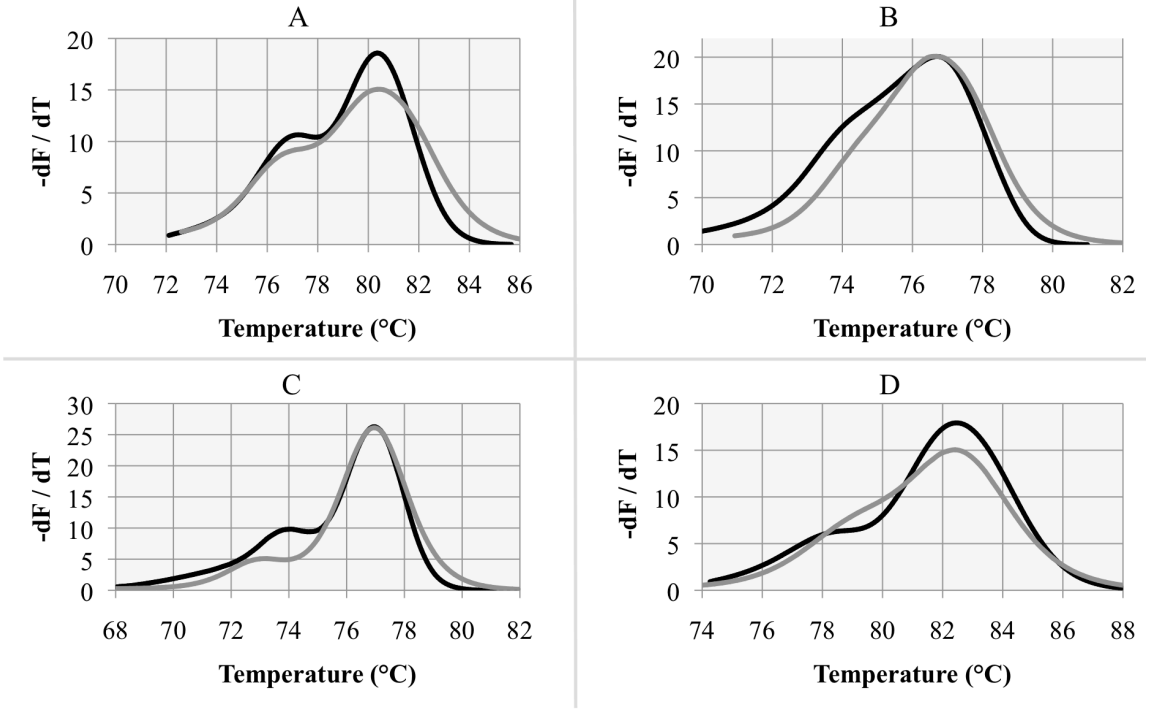


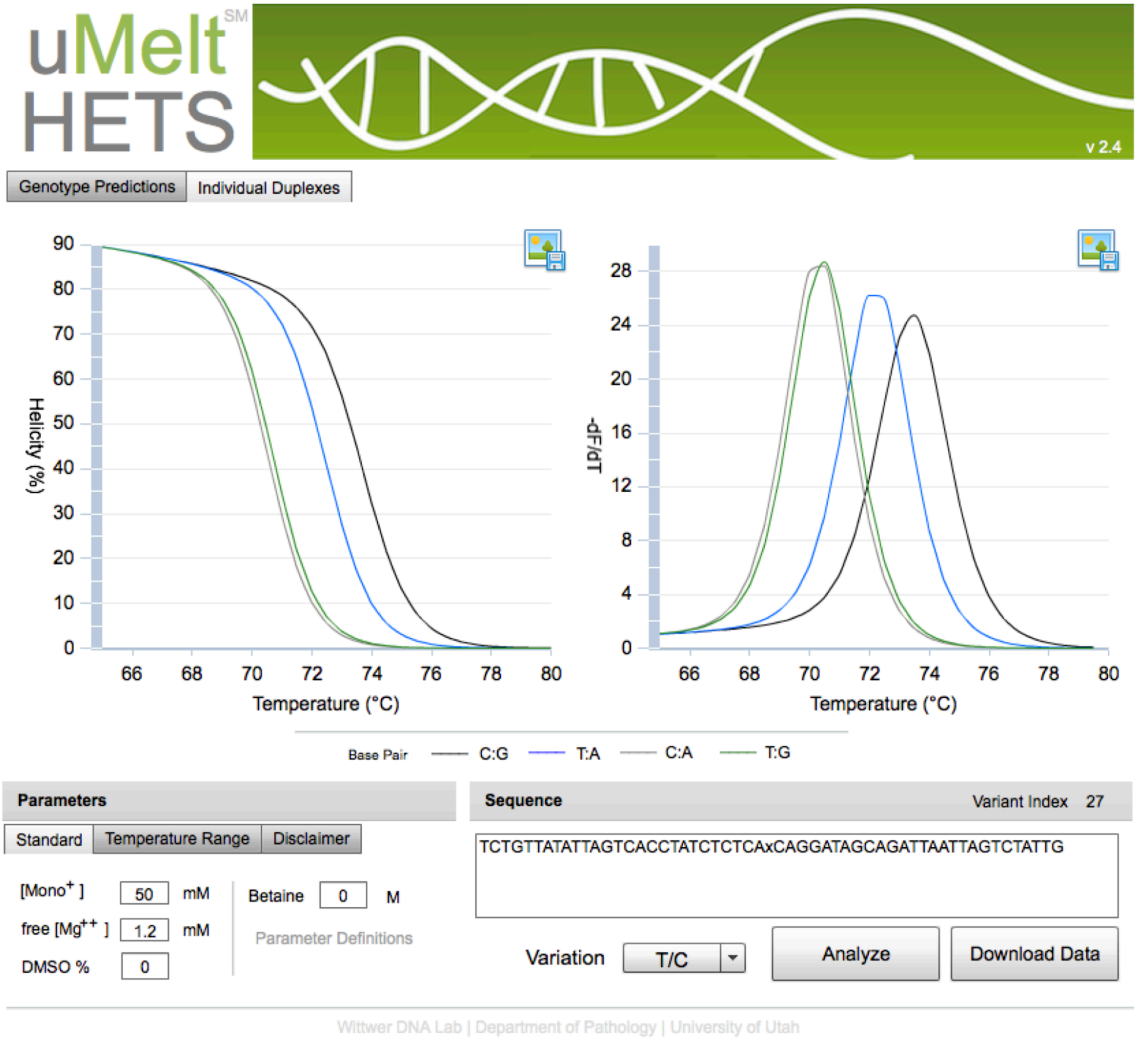
Table 1. Best fit parameters by SNV Class that minimize the area between experimental and theoretical melting curves of heterozygotes after PCR.

SNV Class	# SNVs	Total Heteroduplex (%)	Ionic Scalar (%)
1 (C/T or A/G)	37	25.4 +/- 6.6	86 +/- 15.7
2 (C/A or G/T)	5	23.8 +/-4.9	91 +/- 12.1
3 (C/G)	3	25.3 +/- 2.3	77 +/- 39.3
4 (T/A)	7	28.3 +/- 8.6	97 +/- 5.4
Total	**52**	**25.7 +/- 6.7**	**88 +/- 16.4**

UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

Supplementary Figure 1. uMelt HETS user interface displaying predicted melting curves and derivative peaks of individual duplexes.



UU IR Author Manuscript
UU IR Author Manuscript

University of Utah Institutional Repository
Author Manuscript

rs#	SNP	Forward Primer Sequences	Reverse Primer Sequences
rs1490413	A/G	CTGATGTGGGTTCTTTGC	TGAGGCCAGCCAGT
rs876724	C/T	CCACTGCACTGAAGTATAAGT	TTAGCAGAGTGTGACAAAAAA
rs1357617	A/T	TCCCAAGCTGAATTTGGG	GCTGATAAGAAACATGACCAAG
rs2046361	A/T	TCTATGAACGATCATTTTCAAAATTAAAATT	TGTTGACACTTCACCTTCT
rs717302	A/G	AAAGGCATATCGTATTAACTGTG	AACCTAATGACAGACGTTCA
rs1029047	A/T	CAGCAAAAAAAGTAAGAATTCAAGATGG	GGCTAAAGAAAAGAAAAAAGATATGAGG
rs917118	C/T	AAGATGGAGTCAACATTTTACAAG	GATGACTGAGGTCAACGAG
rs763869	C/T	AGGATGTTTGTTTATATTATTTCTAACTCA	CTACTCCCTCATAATGTAATGC
rs1015250	C/G	GAAAAGAACCCAGGTGTTTTAT	CTTATGAAAAATCACTGGGACA
rs735155	A/G	GAGAGGGAGACCTGCT	ACCGAATTCAACGGGAA
rs901398	C/T	CAAACTAGCTGAATATCAGCC	GCAAGAAACTAATGTTAGCTATCAA
rs2107612	A/G	CTAATTATGTGTTTTTTCTAAATCATATTGTC	ACATTATTCAACTGTTTTGGAGAAAA
rs1886510	C/T	GTAAATTTTAGTAATTCTTAAAAGAATAAGCAATAT	TGGATTTTCACAACAACACTTG
rs1454361	A/T	GGAAATACACCCTGAGCTG	CTTTTCAGTATCCATTTAGAAACAAC
rs2016276	A/G	ACAGTCCACACATTCAGC	GCTGAGAGAGAGAAAAATGTG
rs729172	A/C	AAGGCTCCTCTGCAGA	CTCTTGCGGTTACATACATT
rs740910	A/G	AGTAAGGTGAGTGGTATAATCATA	GCTTTATTTTGTTTTGCTTTTTACAA
rs1493232	C/A	TCTTTTGGGTGCTAGGC	TTTATTGTGAGGCCTGTTTATTTT
rs719366	C/T	CTTTTCCTCCTCCCATTCTA	CCCCAGAGGAGTAGC
rs1031825	A/C	TTAACCTATTAAATTTTAATGAGTATTTTATTTATCTA	CTTTCAAGTATGCTCGGG
rs722098	A/G	CTGTTGACAGTAATGAAATATCCTT	CAGCACATCCAAAATTTAAATCCTT
rs733164	A/G	GCCATCCCACTTGGAA	CTCAGGAATGTCAGGCAA
rs826472	C/T	TCTGTTATATTAGTCACCTATCTCTC	CAATAGACTAATTAATCTGCTATCCT
rs2831700	G/A	GCTAAACTATTGCCGGAGA	CTGTCTTTGGCCTTCTAAC
rs873196	C/T	AAATCCCAAGTGCTGCC	AATCAGGCATGTTCACATTACA
rs1382387	G/T	GAGAAACACCTGAACTTTCA	GCCCACTGCAGGG
rs2111980	A/G	CATCTTGGCAGCATCCTT	CACCAGGCTGGAGC
rs2056277	C/T	GGAGACAGGCATGAATGAG	TTCCCTGTCTTGAACTCC
rs1024116	A/G	TGTTCTAATAAAAAGGATTGCTCA	GTAAATACAAAGGCATACTTTATTAAGT
rs727811	A/C	CGGAACTTCAACGACTT	GCTGGGAGATGCAGATGAT
rs1413212	A/G	CAGTGGTGGAGCATGG	GGAGACATTTGTTCATATAAGTGAAAT
rs938283	C/T	CTAGTACGTTAGATGTGACC	GTAAGGGTCCTATCTCCAAA
rs1979255	C/G	CGAAAGTTCTTCTCCTATGGA	GAACGTTGGAACTCTTGC
rs1463729	A/G	CAGCATACACTCATAGCCA	GGTGGGAATGGCTGC
rs2076848	A/T	CACCACCAGAAATCAGGG	CCGCTTCAGGTCCATC
rs1355366	A/G	GGAGGCCTCGAGGAT	ACTGGCTTTGGCAAGTC
rs907100	C/G	TTCATGATGCCCTGGCA	CCAGTTGGAGCCTTCTT
rs354439	A/T	GTATCTCTCTCATGTATCACATTCC	GTGAATGATATTCAGAATATTGTGCTTA
rs2040411	A/G	TCTCTGTATTTTCTTACTCTAAGTG	CTATGAAAACCAGTTCTCATGAAAT
rs737681	C/T	TGTGAGGCCATCTCCAC	TGGGCCCAGGGATG
rs2830795	A/G	GGACACACCATTTTATTGTCTAAA	AATCAGCAATATAATAGGACTTCTTTG
rs251934	C/T	GCAGTGAGGCTTTTAAGTAG	GTAGTAGATATCTGGCTGTCC
rs914165	A/G	GGGCACCAAAGAGGG	GCTGCCTGAGGGTG
rs10495407	G/A	GGTTGCATTGGATTCTCATTGAA	GGAAAAACAACAGAATGGAATAGG
rs1360288	C/T	AGAGGAGACCTGGGAAG	AGGGGATGCAGCCAG
rs964681	C/T	TGTGAGGAACGTCGAC	GTGGCCCCAAATGCC
rs1005533	A/G	GCAAGAGCCGTGGAAT	TGAACAGCGGCGAC
rs8037429	C/T	ACCTCCATAGTAATAATGTAAGAGTT	CAGAAATAACTCTTGCGCTA
rs891700	A/G	TTCCATTCTTTTTTTTTTGAAGCC	ACACTCCTTAGAACTATGCAA
rs1335873	A/T	GTATGTATTGTTGGCCGTG	TACCTAGCTATGTACTCAGTA
rs1028528	A/G	GAAAGGTCCTTACTCGACAT	GCGGATCTGCACACA
rs1528460	C/T	GAGATCAATATTTAGCCTTAACATATTTAAG	AACAAAACTTTTACTGTAACATTTTTC

UU IR Author Manuscript
UU IR Author Manuscript