2008

Development And Forensic Application Of Dye Probe Fluorescence Resonance Energy Transfer For Improved Detection Of Changes In Dn

Micah Halpern
University of Central Florida

DEVELOPMENT AND FORENSIC APPLICATION OF DYE PROBE FLUORESCENCE RESONANCE ENERGY TRANSFER FOR IMPROVED DETECTION OF CHANGES IN DNA SEQUENCE

by

MICAH HALPERN
B.S. Florida Institute of Technology, 1997

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Sciences
in the Department of Chemistry
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term
2008

Major Professor: Jack Ballantyne

ABSTRACT

Discovering, screening, and associating changes in DNA sequence are important to a broad range of disciplines and play a central role in Forensic Science. The typical types of changes include sequence variations [single nucleotide polymorphisms (SNP)] and length variations [short tandem repeats (STR)]. The steps for forensic DNA sample processing are similar for both types of changes but diverge at the point of detection. A number of approaches are being explored for SNP genotyping while STR analysis primarily consists of size-based analysis by capillary electrophoresis.

Limitations exist for all current detection methods that pose significant impacts to forensic analysis. Bi-allelic SNPs result in three possible genotypes with a minimal amount of information generated per marker. Limitations for SNP analysis are due to the inability to amplify a suitable number of SNP markers from low DNA content samples to provide an appropriate level of discrimination. Multi-allelic STR markers are currently the marker of choice for forensic typing but a variety of experimental artifacts are possible that consist of either biology or technology related causes. Molecular genotyping methods developed across other disciplines have potential to alleviate some of these shortcomings but no current approach is capable of genotyping both SNP and STR loci with a single chemistry. The need for a more effective, efficient, and generalized approach led to development of a unique method called Dye Probe Fluorescence Resonance Energy Transfer (dpFRET) and determination of its suitability for forensic analysis.

The development phase of the research consisted of synthetic testing to establish proof of concept for the chemistry followed by polymerase chain reaction (PCR) based assays to demonstrate real world applications. Following successful development, the boundaries and limitations for the technology were established (sensitivity, allelic dropout, mixed samples) and efforts were made to improve the approach. In the process, parallel testing for other fields including molecular pathology and conservation biology were incorporated to explore potential widespread application of this new approach.

The overall goal of this project was to develop and explore the limitations for a unique approach to genotyping both SNPs and STRs. A majority of the work involved development of the method itself with the ultimate objective of application for forensic science. The focus of this project was to address and alleviate some of the shortcomings of current approaches that result in potential limitations for forensic analysis. It is expected that future applications of this technology might impact a wide range of disciplines to aid in discovery, screening and association of changes in DNA sequence.

ACKNOWLEDGEMENTS

I would like to thank Dr. Jack Ballantyne for giving me the opportunity to work in his lab, and for his guidance, teaching, patience and support. I would also like to thank Dr. Glenn Cunningham and Dr. Barry Fookes for serving as members of my committee and I thank them for their input and patience.

My graduate experience was intense and challenging and would not have been possible without the constant support and encouragement of my parents, Jack and Joan Halpern, and I can never thank them enough or adequately show my appreciation.

I would also like to recognize Midwest Research Institute for their ongoing support in developing both my ideas and my scientific career. In addition, I would also like to thank Dr. Jean Dubach for providing me with a strong foundation in molecular biology and whose contagious enthusiasm made me interested in the field. I would also like to thank the Johnson County Criminalistics Laboratory, National Center for Forensic Science, Dartmouth School of Medicine and Brookfield Zoo for providing both samples and advice on applying my ideas to the fields of Human Forensics, Molecular Pathology and Conservation Biology.

Finally, I must thank my wife Adriana Halpern for her never-ending support and infinite patience that has truly allowed me to pursue my dreams.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYSMS/ABBREVIATIONS

bp—base pairs

dI—deoxyinosine

DNA—Deoxyribonucleic Acid

dpFRET—Dye Probe Fluorescence Resonance Energy Transfer

FET—Fluorescence Energy Transfer

FRET—Fluorescence Energy Resonance Transfer

INDEL—Insertion/Deletion

Mhc—Major Histocompatability Complex

NTC—No Template Control

PCR—Polymerase Chain Reaction

RFLP—Restriction Fragment Length Polymorphism

SNP—Single Nucleotide Polymorphism

STR—Short Tandem Repeat

TdT—Terminal deoxynucleotidyl transferase

Tm—Melting Temperature

VNTR—Variable Number Tandem Repeat

CHAPTER ONE: INTRODUCTION

Discovering, screening and associating changes in DNA sequence have importance across a broad range of disciplines including not only forensics but medicine, ecology and molecular biology to name a few. The typical types of alterations in DNA sequence that can be observed include sequence variations often termed single nucleotide polymorphisms (SNP) and length variations referred to in forensics as short tandem repeats (STR). There are a number of other terms used to define length variations [variable number tandem repeat (VNTR), insertion/deletions (INDEL), microsatellites, etc.] but the general principle is the same. SNPs are defined as a variation between samples at a particular site within a sequence (i.e., CCT to TCT). STRs are a form of INDEL that arises when one or more nucleotides is added or subtracted from a sequence (i.e., CCT to CT). STRs are an example of a specific type of INDEL often attributed to polymerase slippage (Nadir *et al.* 1996) or defective DNA replication repair (Ellegren 2004) and consist of repeating units of 1-6 base pairs (i.e., CAGCAG). Examining both these types of changes has become central to a number of disciplines due to the ability to utilize these alterations to look at everything from identity to disease.

The first utilization of DNA changes for human identity studies began with a case involving the murder of two young girls in England in the early 1980s (Wambaugh 1989) and the application of an early form of DNA fingerprinting developed by Sir Alec Jeffreys (Jeffreys *et al.* 1985). Dr. Jeffreys developed an approach based on repeat regions (VNTRs) that were examined using restriction fragment length polymorphism (RFLP). This was followed by the introduction of a

variety of techniques including single-locus probe and multi-locus probe RFLP methods and more recently polymerase chain reaction (PCR) based assays. The introduction of PCR into human identity testing made it possible to utilize tiny amounts of sample to yield a useful DNA profile and has heralded the next step in the evolution of forensic science.

No matter the approach for detection, the steps for DNA sample processing in forensic science are similar for both genotyping approaches (STRs and SNPs). Interestingly enough, the forensic approach is similar and based on the same fundamental principles and methods used for medical diagnostics and genetic mapping. Briefly, after the biological material is collected from the scene, DNA is extracted and typically quantitated followed by PCR amplification and final detection of the informative DNA marker. It is at the point of detection that current approaches diverge depending on whether sequence (SNP) or length (STR) variations are to be examined. The current accepted method for detection of STRs in forensic applications is electrophoresis (primarily by capillary) with other options for size separation [mass spectrometry (Butler *et al.* 1999), array based hybridization (Pourmand *et al.* 2007, Kemp *et al.* 2005 and Radtkey *et al.* 2000] under development. SNP detection encompasses a much larger list of potential approaches (Gut 2001 and Kwok 2001). A few of the primary SNP typing methods that have been more accepted for forensic studies includes approaches such as minisequencing (Tully *et al.* 1996), Taqman (Lareu *et al.* 2001) and pyrosequencing (Andréasson *et al.* 2002). Developments in other fields for screening of both sequence (SNPs) and size (STRs) variations have the potential to contribute to advancing approaches in forensic science by alleviating or avoiding issues posed by current approaches.

SNPs are appealing for forensic applications for a number of reasons including small PCR product size for degraded DNA samples, potential for multiplexing and automation and simplified analysis to name a few. SNP markers are generally bi-allelic with two possible alleles resulting in three possible genotypes. This means that a minimal amount of information can be generated per marker for SNPs. It has been estimated that approximately 50-100 SNP markers would be required to match the discriminatory power of 10-16 STR loci (Gill *et al.* 2004). The result is the inability to simultaneously amplify a suitable number of SNP markers from low DNA content samples. The ultimate solution for SNP typing would be an approach that could genotype multiple changes per reaction hence reducing the impact on sample consumption. Although SNPs are appealing markers for forensic applications, multi-allelic STR markers are currently the marker of choice for forensic typing.

The most discriminatory markers currently used in forensic laboratory analysis are the extensively validated collection of STRs comprising the CODIS loci. The standard approach for analysis of these markers is multiplex amplification followed by capillary electrophoresis size separation. There are a number of issues that are important for obtaining accurate genotyping results of STRs as detected by capillary electrophoresis. Varieties of known experimental artifacts are possible and consist of either biology or technology related causes. The most common biological artifact are stutter peaks, incomplete 3'(A) nucleotide additions and tri-allelic patterns. Technology related artifacts are due to matrix failures, dye blobs, voltage spikes and sample contaminants. The solution to a majority of these issues would be an approach that either

minimizes or is not subject to these artifacts. As STR analysis plays a central role in forensic analysis, any potential solution to these shortcomings would significantly impact genotyping analysis.

Technologies developed across other disciplines to screen for these changes have potential to alleviate some of the difficulties associated with current forensic approaches. These technologies are almost exclusively PCR-based and fall under the major categories of hybridization based, enzyme based, post-amplification detection and different forms of DNA sequencing. The goal of the research detailed in this thesis focuses on an advancement specifically within the first category of hybridization. Within this category, developments aimed at discovering and identifying DNA changes can be classified under two major sub-categories of (1) generic DNA intercalator techniques and (2) strand specific hybridization.

The first subcategory within hybridization comprised of generic methods utilizes DNA intercalating dyes that exhibit increased fluorescence when bound to double stranded DNA. These fluorescent moieties include SYBR, SYTO and a host of other well characterized dyes (Gudnason *et al.* 2007). The primary application of end point melting curve analysis using these dyes focuses on SNP genotyping with no significant work to date on utilizing this approach for STR genotyping. Typical application of this approach focuses on the goal of either discriminating PCR artifacts (i.e., primer dimer) from specific amplicon or SNP genotyping. Genotyping methods solely using intercalating dyes have shown a somewhat low level resolution between amplicons with similar sequence (Herrmann et al. 2006). More recent development for

higher resolution screening has focused on using more proprietary dyes (LC Green) and advances in data analysis (Reed *et al.* 2004). Although somewhat limited in their ability to resolve many different types of changes in DNA between samples, the major benefit to this hybridization based approach is the cost savings associated with minimized reagent requirements and reduced design constraints.

The second subcategory within hybridization is composed of strand specific methods that utilize additional nucleic acid reaction components (beyond generic dyes) to monitor the progress of amplification reactions. The most typical added reaction component is some form of oligonucleotide probe designed in or around the sequence of interest. These methods often use fluorescence energy transfer (FET) as the basis of detection. One or more nucleic acid probes are labeled with fluorescent molecules, one of which is able to act as an energy donor and the other of which is an energy acceptor molecule. These are often referred to as a reporter molecule and a quencher molecule respectively. The donor molecule is excited with a specific wavelength of light which falls within its excitation spectrum and subsequently it will emit light within its fluorescence emission wavelength. The acceptor molecule is then excited at the emitting wavelength of the first molecule by accepting energy from the donor molecule by a variety of distance-dependent energy transfer mechanisms. A specific example of FET is Fluorescence Resonance Energy Transfer (FRET). Generally, the acceptor molecule accepts the emission energy of the donor molecule when they are in close proximity (e.g., on the same, or a neighboring molecule) with the distance of separation termed the Forster distance. The basis of FRET detection is to monitor the changes at the acceptor emission wavelength caused by

separation of the two moieties. There are two commonly used types of FRET probes, those using hydrolysis of nucleic acid probes to separate donor from acceptor, and those using hybridization to alter the spatial relationship of donor and acceptor molecules.

Hydrolysis probes (Figure 1) are commercially available as Taqman probes and are primarily used for either quantitation or SNP genotyping. There has been no described example of application of this approach for genotyping STR loci. Hydrolysis probes consist of DNA oligonucleotides that are labeled with donor and acceptor molecules. The probes are designed to bind to a specific region on one strand of a PCR product. Following annealing of the PCR primer to this strand, Taq enzyme extends the DNA with 5' to 3' polymerase activity. Taq enzyme also exhibits 5' to 3' exonuclease activity. TaqMan probes are typically protected at the 3' end to prevent extension. If the TaqMan probe is hybridized to the product strand, the Taq polymerase enzyme will subsequently hydrolyze the probe, liberating the donor from acceptor as the basis of detection. The signal in this instance is cumulative, the concentration of free donor and acceptor molecules increasing with each cycle of the amplification reaction. This approach is typically used for quantitation and more recently has been adapted for SNP detection on an assay specific basis.

As opposed to hydrolysis probes, hybridization probes (Figure 2) are available in a number of forms and are not consumed during detection. Molecular beacons are an example of oligonucleotides that have complementary 5' and 3' sequences such that they form hairpin loops. Terminal fluorescent labels are in close proximity for FRET to occur when the hairpin structure

is formed. Following hybridization of molecular beacon to a complementary sequence the fluorescent labels are separated, so FRET does not occur, and this forms the basis of detection. Another approach to using hybridization probes utilizes a pair of labeled oligonucleotides commonly known as dual hybridization probes. These hybridize in close proximity on a PCR product strand bringing donor and acceptor molecules together so that FRET can occur. Variations on this approach can include using a labeled amplification primer with a single adjacent probe. As opposed to dye based hybridization, hybridization probes have shown good success with obtaining high levels of resolution for SNP genotyping (Bernard *et al.* 2000) but suffer from other shortcomings.

The use of either dual hybridization probes or molecular beacons requires labeling with two fluorescent molecules which subsequently increases the cost involved in using these approaches. In addition, both methods require the presence of a reasonably long stretch of known sequence so that the probe/probe pair can bind specifically in close proximity to each other. This can be a problem in some applications, where the length of known sequences that can be used to design an effective probe may be relatively short. Furthermore, the use of pairs of probes involves more complex experimental design whereby the genotype is a function of the denaturation of both probes and requires careful design parameters often limited by sequence identity.

The most significant shortcoming to all current forms of discovering and screening changes in DNA, whether by dye or probe, is the lack of application of hybridization based approaches for genotyping different types of DNA changes (SNP and STR) with a single chemistry. It was this

need for a more effective, efficient and generalized approach that was the driving force behind the development of a unique approach called Dye Probe FRET (dpFRET).

Dye Probe FRET (dpFRET)

The most optimal approach to discovery and screening of multiple variations would be to combine the reduced cost and ease of use of generic intercalating dyes with the resolution and increased sensitivity of hybridization probes. This combination has been proposed by others but has not yet been explored to its fullest potential. A less sophisticated approach has been demonstrated by genotyping with unlabeled probes both post PCR (Zhou *et al.* 2005) as well as inclusion within the amplification reaction (Zhou *et al.* 2004). This approach required additional analysis and was not able to discriminate all potential alleles. Additionally, it was not demonstrated that this same approach could be used for STR typing.

An integrated system utilizing FRET between an intercalating dye and a probe labeled with a single fluorophore has been reported previously. Howell et al. demonstrated a basic application of the approach which showed a dramatic increase in signal intensity when compared with standard intercalating dye and FRET approaches (Howell *et al.* 2002). The same technology is also summarized specifically for studying changes in DNA hybridization (Howell 2006). Takatsu et al. describes a related approach based on labeled nucleotide incorporation followed by dye fluorophore FRET detection (Takatsu *et al.* 2004). These studies had yet to identify and demonstrate the true potential of this approach for genotyping a number of different types of changes in DNA (SNP and STR) and its application to forensics and individual identification.

dpFRET SNP Genotyping

The current approach for genotyping SNPs using dpFRET (Figure 3) entails: (1) generation of template for probe hybridization by asymmetric PCR, (2) hybridization of a fluorophore labeled probe in the presence of a DNA intercalating dye, and (3) standard melt curve analysis. Many different dye/fluorophore combinations are possible, but for purposes of illustration an example will be presented that uses SYBR Green I as the donor dye (Excitation 490, Emission 520) and Texas Red as the acceptor fluorophore (Excitation 590 and Emission 620). The dye is excited at a wavelength of 490 nm, emits at a wavelength of 520 nm which is transferred via FRET to the fluorophore on the labeled probe and reemitted at 620 nm. Figure 4 shows excitation and emission wavelengths for SYBR Green I and Texas Red, region of FRET between the two molecules and the dual emission signal generated by both dyes. Filter band widths used to provide the excitation signal or measure the emission signal were provided with the real-time PCR system.

One of the unique beneficial outcomes of this approach is the generation of two melt peaks. The peak at the lower melt temperature is a result of the signal from the FRET probe and the peak at a higher temperature is the result from the melting of the amplicon itself. The amplicon melt peak is generated by fluorescence of intercalated SYBR Green I at the tail end of the SYBR Green I emission spectrum (light hashed region in Figure 4). This secondary melt peak provides a positive signal for amplification of specific product and can be used to distinguish non-specific signal occasionally generated by the probe for > 45 cycle amplification reactions.

The dpFRET approach has been tested for its ability to detect and differentiate between single and multiple SNPs within a target region using a single labeled probe (Figure 5). It is this ability to multiplex SNP detection that has the potential to minimize the amount of sample that is consumed per genotyping reaction. Due to melting behaviors of DNA, positioning of the signal generating molecule (fluorophore) at the distal end of the probe rather than throughout the strand as in the case of intercalating dyes capitalizes on the "end effects" seen for DNA melting. These effects referred to as either "end effects" or "end fraying" can propagate several base pairs into the duplex as demonstrated by nuclear magnetic resonance experiments (Leijon et al. 1992). Melting models that have been developed that incorporate considerations for the preference of melting to initiate from the ends of a short duplex have shown statistically improved data fits (Doktycz et al. 1995). A potential conclusion can be drawn from existing studies that intercalating dyes that bind across the duplex have reduced resolution of signal differences due to SNPs at different positions as a result of a signal blending effect that does not capitalize on these "end effects" for melting analysis. It is hypothesized that in the case of dpFRET and other probe based detection chemistries that end effects contribute to the enhanced resolution seen for these approaches.

dpFRET STR Genotyping

A significant strength to the dpFRET approach is that the same technology can be applied to the typing of not only SNPs but STRs as well. In a similar manner as SNP detection, template is generated for a region of interest by asymmetric PCR followed by hybridization with an allele specific probe. The allele specific probe contains a defined number of repeats and results in a

match/mismatch based analysis. Application of dpFRET to STRs results in two potential melt peaks for the probe consisting of either a match (higher Tm) or a mismatch (lower Tm) with the number of repeats contained within the target. Mechanisms for probe hybridization are depicted in Figure 6 for both a simple (single core repeat) and complex (multiple core repeats) locus. An example of the experimental results using dpFRET for STR analysis is shown in Figure 7. Success for genotyping of both homozygotes and heterozygotes is possible using this approach.

This approach for STR genotyping provides significant benefits over standard size separation based analysis. No additional manipulation beyond a standard melt curve is required significantly reducing the time to results. The only additional costs are labeled probes which are significantly less than reagents required for fragment analysis. Less sample manipulation is required and the protocol is highly amendable to microfluidic and automated platforms. Most importantly, the objective analysis can be automated and does not suffer from the same potential artifacts as CE analysis.

The goal of the research described in this thesis was to develop and explore the limitations for this unique approach to genotyping SNPs and STRs. A majority of the work involved development of the method itself with the ultimate objective of application for forensic science. The real focus of this project was to address and alleviate some of the shortcomings of current approaches that result in potential limitations for forensic analysis. As the approach is a general method for genotyping, additional applications (molecular pathology and conservation biology) were tested and explored. It is expected that future applications of this technology might impact a

wide range of disciplines beyond forensic science to aid in discovery, screening and association of changes in DNA sequence.

CHAPTER TWO: MATERIALS AND METHODS

Synthetic SNP Testing

Detection and analysis of single and multiple SNPs was tested using two synthetic template libraries composed of a variable human sequence and a variable animal species sequence library.

Sequence corresponding to positions 14925-14974 of the Cambridge human mitochondrial genome (J01415) and mutated templates were synthesized, purified by standard desalting and concentrations were standardized by a commercial source (Integrated DNA Technologies). The mutated templates consisted of representatives for every possible single point mutation within the 30 bp central core region comprised of positions 14935-14964. The variable animal species template library contained sequence corresponding to the same position of the Cambridge human mitochondrial genome from a number of animal species as listed in Table 1 and was generated by the same commercial source. Non-variable 10 bp sequences flanking the variable regions were also included in each template to avoid potential problems associated with incomplete synthesis such as N-1 templates. All sequences for both libraries are listed in Appendices C and D.

Both template libraries were evaluated by standard melt curve analysis with human reference probe sequences (30 bp: ACGTCTCGAGTGATGTGGGCGATTGATGAA, 21 bp:

TCGAGTGATGTGGGCGATTGA, 15 bp: GTGGGCGATTGATGA) labeled at the 3' terminus with a Texas Red-X NHS Ester. The fluorescent probe was commercially synthesized, HPLC purified and quantity standardized by a commercial source (Integrated DNA Technologies). Hybridization reactions contained 1X SYBR Green I Master Mix (Bio-Rad), 50 uM template and 5 uM labeled probe and were subjected to the following thermal protocol on an IQ5 real-time thermal cycler (Bio-Rad): 95 degrees for 1 minute, 25 degrees for 1 minute and incremental increase of 0.2 degrees to a final temperature of 95 degrees. A standard excitation filter of 490 nm (30 nm bandwidth) was coupled with a 620 nm (20 nm bandwidth) emission filter placed in the appropriate corresponding position of the emission filter wheel.

Terminal deoxynucleotidyl transferase (TdT) Probe Labeling

In an effort to reduce the cost of probe synthesis, in-house labeling of synthetic oligonucleotides was tested using TdT (New England Biolabs) and ChromaTide Texas Red-12-dUTP (Invitrogen). The same oligonucleotide sequence used for synthetic probe testing (ACGTCTCGAGTGATGTGGGCGATTGATGAA) was commercially synthesized, purified and quantitated (Integrated DNA Technologies) and used as template for TdT labeling. The following were combined: 200 uM oligonucleotide, 1X NEB buffer 4, CoCl2 (5 mM), Texas Red-12-dUTP (1 mM) and 60 units of terminal transferase. The reaction was incubated overnight at 37 degrees and terminated by incubation at 70 degrees for 10 minutes. The TdT labeled probe was purified by chromatographic separation of unincorporated fluorophore nucleotides using a DyeEX kit (Qiagen). Labeled probe was tested against the variable human sequence template library and melted as previously described.

Inosine Probes

Artificial manipulation of duplex melt temperatures was tested by incorporation of the nucleotide analogue inosine at different positions within the human reference probe sequence. Two hybridization probes were commercially synthesized, purified and quantitated (Integrated DNA Technologies). One probe contained an inosine at position 30 within the sequence and a second probe at positions 28, 29, and 30. Both probes were fluorescently labeled with TdT as previously described to test impacts of inosine on duplex melting characteristics. Fluorophore labeled inosine probes were tested against the variable human sequence template library and melted as previously described.

Assay Design, Amplification and Probe Hybridization

SNP Species Identification—*Cytochrome B*

Published sequences (NCBI) of *Cytochrome B* for multiple animal species were aligned using MegAlign (Lasergene) and regions of conservation were used to manually design primers according to standard practice. Optimal primer sequences used for dpFRET testing were CYTB 0088F Mix: 5'-TCCGCATGATGAAAyTTyGGnTC-3' and CYTB 0438R Mix: 5'-GTGGCCCCTCAGAAdGAyATyTG-3'. Previously extracted and quantitated genomic samples derived from blood for multiple animal species were provided by Brookfield Zoo (Brookfield, IL). Previously extracted and quantitated genomic samples for human and ferret species was provided by the National Center for Forensic Science (Orlando, FL). All quantitation was verified using Picogreen and supplier recommended protocols (Invitrogen). Asymmetric

PCR reactions contained 1X SYBR Green Mastermix (BioRad), 500 nM forward primer and 15 nM reverse primer. The protocol used for asymmetric amplification included an initial denaturation at 95 degrees for 3 minutes followed by 40 cycles of 95 degrees for 10 sec, 59 degrees for 40 sec. This was immediately followed by 40 cycles of 95 degrees for 10 sec, 56 degrees for 40 sec. Following amplification, the reaction was supplemented with 5 uM of commercially synthesized human reference probe (previously described) and subjected to melt curve analysis using a 0.5 degree incremental increase in temperature on an IQ5 real-time PCR platform (Bio-Rad).

SNP Individual Identification—*Mhc DRB*

Published sequences (NCBI) of *Mhc DRB* for multiple animal species were aligned using MegAlign (Lasergene) and regions of conservation were used to manually design primers according to standard practice. Optimal primer sequences for dpFRET testing were UNIV_MHCdr_3F Mix: 5'-ACGGsACsGAGCGGGTG-3' and UNIV_MHCdr_3R: 5'-CACCCCGTAGTTGTGTC-3'. Previously extracted and quantitated genomic samples derived from blood for two families of captive Humboldt Penguins (*Spheniscus humboldti*) were provided by Brookfield Zoo (Brookfield, IL). Quantitation was verified as previously described. Asymmetric PCR reactions containing 1X SYBR Green Mastermix (BioRad), 100 nM forward primer and 500 nM reverse primer were amplified using the following thermal protocol: Initial denaturation at 95 degrees for 3 minutes followed by 40 cycles of 95 degrees for 10 sec, 63 degrees for 40 sec. This was immediately followed by 40 cycles of 95 degrees for 10 sec, 59 degrees for 40 sec. Following amplification, the reaction was supplemented with 5 uM of

commercially synthesized Texas Red fluorescently labeled probe: UNIVdr 0245 (ATAACCAAGAGGAGTCCGTGCGCTTCGACAGCGA/3'TR), UNIVdr 0273 (5'TR/AGCGACGTGGGGGAGTACCGGGCGGTGACGGAGCTGGG), UNIVdr 0309-3'TR (GGGCGGCCTGATGCCGAGTACTGGAACAGCCAGAAGGA/3' TR), UNIVdr 0340-3'TR (CAGAAGGACCTCCTGGAGCAGAGGCGGGCCGCGGTGGA/3' TR), HUMdr 0509-3'TR (GGCTGAGGTGGACACGTACTGCCGA/3' TR) and HUMdr 0536-3'TR (CACAACTACGGGGTGGTGACCCCTTTCACT / 3'TR). Reactions were subjected to melt curve analysis using a 0.5 degree incremental increase in temperature on an IQ5 real-time PCR platform (Bio-Rad). Amplicons generated for dpFRET testing were also sequenced using standard dideoxy sequencing according to manufacturers protocols (Applied Biosystems) for comparison to dpFRET results.

STR Individual Identification—TPOX and D3S1358

Human TPOX and D3S1358 primer sequences from the PowerPlex 16 kit (Promega) were commercially synthesized (Integrated DNA Technologies) and tested against CE genotyped samples derived from buccal swabs provided by the Johnson County Crime Laboratory (Olathe, KS). Primer sequences included: TPOX F (5'-GCACAGAACAGGCACTTAGG-3'), TPOX R (5'-CGCTCAAACGTGAGGTTG-3'), D3S1358 F (ATGAAATCAACAGAGGCTTGC) and D3S1358 R (ACTGCAGTCCAATCTGGGT). The thermal protocol used for PCR amplification consisted of the following: Initial denaturation at 95 degrees for 3 minutes followed by 40 cycles of 95 degrees for 10 sec, 59 degrees for 30 sec and 72 degrees for 30 sec followed by 40 cycles of 95 degrees for 10 sec, 57 degrees for 30 sec and 72 degrees for 30sec. Following

amplification, each reaction was supplemented with 5 uM of commercially synthesized allele specific Texas Red labeled probe and melted as previously described using a 0.5 degree incremental increase in temperature on an IQ5 real-time PCR platform (Bio-Rad). Probes consisted of the following basic structure:

TPOX [GAACCCTCACTG $(AATG)_N$ TTTGGGCAAATAAACGCTGACAAG]

D3S1358 [TGCATGTATCTA $(TCTG)_N$ $(TCTA)_N$ TGAGACAGGGTCTTGC]

The number of core repeats (N) corresponded with each allele tested.

Sensitivity and Allelic Dropout—*Mhc DRB* and TPOX

Human genomic samples used for STR individual identification testing were also used to determine assay sensitivity and potential for allelic dropout. Both homozygote and heterozygote samples were tested using protocols previously described for *Mhc DRB* and TPOX. Samples were re-quantitated using Picogreen and manufacturers protocols (Invitrogen) and diluted ten fold from 5 nanograms (approximately equivalent to 1000 genomic copies) to 500 femtograms (approximately equivalent to 0.1 genomic copies) in water using ten fold dilutions. Amplification and melt curve analysis was performed as described previously.

Mixed Sample Testing—TPOX

Laboratory generated mixes of human genomic samples were used to determine the potential to detect multiple STR genotypes within a mixed sample. Following quantification of material obtained from the Johnson County Crime Laboratory (described previously), 1 nanogram samples from a homozygote, heterozygote and an individual lacking a TPOX eight repeat allele

were mixed in different combinations to examine the ability to detect changes in allelic concentrations within a sample.

Following laboratory generated mix testing, samples provided by the Dartmouth School of Medicine were tested for application to “Real World” samples. Samples were originally obtained for a previous study on chimerism in bone marrow transplant patients. Multiple cell fractions (donor, recipient, monocytes, granulocytes, peripheral blood and bone marrow) were sampled following treatment to monitor the success or rejection of the transplanted tissue. If transplant recipient genotype is detected in any of the cell fractions this dictates the need for additional testing and alters treatment. Genotypes generated by standard protocols used in forensic analysis (Beckman Coulter CEQ 8000 and ID kit) were supplied by Dartmouth School of Medicine for comparison to dpFRET STR genotyping. Samples were analyzed using dpFRET as previously described for the TPOX locus and results compared to current accepted protocols.

CHAPTER THREE: RESULTS

Synthetic Single SNP Testing—Template Variation

Results for the variable human sequence template library testing using a 30 bp probe are shown in Figure 8. dpFRET results are shown in the top panel which depicts the melt temperature for each positional change within the template tested with a fluorophore labeled probe. Error bars of ± 0.4 degrees are labeled for each data point to account for thermal block variation. The range for an exact match (reference template) is highlighted across the graph. The bottom panel represents similar testing with an unlabeled probe (standard intercalating dye melt analysis) to explore fluorophore effect on melting temperature. The 30 bp 3' fluorophore labeled probe resulted in discrimination of any change at any position except for mutations in the template complementary to probe nucleotides 30, 29, and 1. In contrast, the unlabeled probe was unable to discriminate mutations at multiple positions both distal and internal within the template (probe nucleotides 26, 22, 13 and 1). It is also important to note that the melt point graph is similar between labeled and unlabeled probes with the labeled probe displaying more significant variation from the reference for most points.

To understand effect of probe size, 21 and 15 bp fluorophore labeled probes were also tested and showed similar results with finer resolution at the ends of the template using the dpFRET approach. The fluorophore labeled 21 bp probe (Figure 9 top panel) was indistinguishable from the reference for template mutations complementary to probe positions 21 and 1 and showed no effect due to template mutation in flanking sequence. Similar melting protocols using an

unlabeled probe (Figure 9 bottom panel) resulted in melt temperatures indistinguishable from the reference for mutations at multiple positions (probe nucleotides 17, 8, 4, 3, 2 and 1). Additionally, an effect was seen for mutations in upstream sequence flanking the unlabeled probe (probe nucleotides +1, +3 and +4). The fluorophore labeled 15 bp probe (Figure 10) resulted in differential melt temperatures from the reference for all mutations except probe nucleotide 15 with a minor difference due to a flanking mutation (probe nucleotide -12). An unlabeled 15 bp probe was not tested.

Synthetic Multiple SNP Testing—Species Variation

All synthetic animal species templates showed reduced melt temperatures compared to the human reference sequence when hybridized with a human probe sequence (Figure 11). A few templates are listed with the number of SNPs in parenthesis to illustrate the range of sequence divergence. In general, increased number of SNPs within the template tended to reduce the melt temperature as would be expected. Four species templates (Skate, Aardvark, Dogfish and Dugong) did not produce melt curves when tested with a human probe sequence. All these templates had > 10 SNPs. It should also be noted that closely related Orangutan sequence showed a differential melt temperature and contained only a single SNP. Unlabeled probe was not tested against the animal species library.

TdT Probe Labeling

Melt curve results for a 3' fluorophore TdT labeled 30 bp oligonucleotide probe showed no significant differences from commercially synthesized probe (Figure 12). Melt temperatures for

both commercially synthesized probe and TdT labeled probe were within the limits of experimental error (± 0.4 degrees) for each template within the human variation library. It should also be noted that labeling efficiency of the enzyme was extremely low and did not provide significant amounts of reagent for sample testing.

Effect of Inosine on Probe Hybridization

Addition of inosine at variable positions had a significant effect on probe melting for all templates within the variable human sequence library (Figure 13). All templates for both probe treatments (substitution of one or three inosines) showed reductions in melt temperature. In order to account for these reductions and highlight any positional effects, melt temperatures for all templates were altered by a 2 degree increase for the single inosine probe and 6 degrees for the three inosine probe. Following modification of the data, the probe treatment with one inosine at the 3' (position 30) end of the probe showed a significant difference from an unmodified probe for all three mutations at position 30 of the template with only a slight difference at position 5 downstream of the modified residue. The probe treatment with inosine at positions 30, 29 and 28 showed a marked difference in melt temperature from the unmodified probe at positions complementary to the inosine residues (positions 30, 29, 28), at positions adjacent (positions 27, 26, 25, and 24) and at positions distal (positions 7 and 5) to the modified residues.

SNP Haploid Locus Testing (*Cytochrome B*)

Melt curve results for real world species testing is shown in Figure 14 with melt curve values listed in Table 2. All species tested resulted in two melt peaks indicative of both amplification

and probe hybridization except python which produced only an amplification peak. Published python species sequences differ by > 10 nucleotides from the human reference probe sequence used for testing. The no template control resulted in a broad non-specific probe melt peak but did not exhibit an amplicon peak characteristic of production of an amplification product.

SNP Diploid Locus Testing (*Mhc DRB*)

Paternity results for real world testing of two known Humboldt Penguin families are shown in Figure 15. A sequence alignment for the amplification products produced using the universal *Mhc DRB* PCR assay is listed at top of the figure. Differences relative to sequence for the H960336 individual are listed using standard degenerate nucleotide base codes (i.e., Y=C or T, R=A or G, etc.). All melt temperatures generated by dpFRET analysis were converted to allele designations of either A, B or C for presentation purposes. Paternity results previously established by Brookfield Zoo through both Southern blot analysis and zoo keeper records for the two families are depicted at the bottom of the figure. Previously established paternity agreed with results generated by dpFRET analysis.

SNP Assay Sensitivity and Allelic Dropout

The limit of detection using dpFRET for SNP analysis was 5 picograms (approximately 1 genome equivalent) for both homozygote and heterozygote samples (Figure 16). Fluorescent signal showed no decrease for less concentrated samples and no allelic dropout was observed for the heterozygote. Both 500 femtograms (approximately 0.1 genome equivalents) and the no

template control showed non-specific probe interaction as evidenced by a broad probe melt peak with neither sample resulting in a peak indicative of specific target amplification.

STR Simple Locus Testing (TPOX)

dpFRET analysis of the TPOX locus for samples provided by the Johnson County Criminalistics Laboratory showed identical results to genotype data previously generated by the crime lab using standard capillary electrophoresis detection (Figure 17). dpFRET melt curves for each allelic probe are shown.

STR Complex Locus Testing (D3S1358)

Similar to STR simple locus testing, dpFRET analysis of the D3S1358 STR complex locus resulted in similar although not identical results. When analyzed by size, complex STR loci can result in the same size profile for alleles that do not contain the same sequence. This is due an equivalent change (an addition to one core repeat with a deletion in the second core repeat) that cannot be differentiated based on size. Discrepancies for some samples were seen when analyzed by dpFRET due to the sequence based analysis of the approach that was able to detect this type of difference between alleles. As this complicated the comparison between dpFRET and standard approaches, an example of the results generated by dpFRET are provided in Figure 18 to illustrate this potential phenomenon. Two individuals both typed as homozygotes and containing 17 repeats resulted in differential patterns (17’ homozygote and 17, 17’ heterozygote) when analyzed by dpFRET. Additional studies using cloning and sequencing is currently underway to verify these results.

STR Assay Sensitivity and Allelic Dropout

Preliminary results to determine the limit of detection using dpFRET for STR analysis was 50 picograms (approximately equivalent to 10 genomic copies) for both homozygote and heterozygote samples (Figure 19). It is important to note that fluorescent signal showed no decrease for less concentrated samples and no allelic dropout was observed for the heterozygote.

STR Mixed Sample Testing

Artificial Mix

Artificial mixtures of homozygote and heterozygote samples tested with an 8 repeat allelic probe resulted in fluorescent match and mismatch signal intensity changes approximately equivalent to the concentration of allele within the sample (Figure 20). The first mix composed of a homozygote and heterozygote (left panel) contained approximately 3X the amount of target allele (8 repeats) compared to non-target allele (10 repeats) and resulted in a significantly higher match peak signal intensity. It should be noted that the match and mismatch peak fluorescent intensities are not directly correlated with sample allelic content (match ~ 170 RFU, mismatch ~ 80 RFU). The second mix (middle panel) contained an equal proportion of target and non-target allele and resulted in approximately equivalent fluorescent intensities for the match (~ 110 RFU) and mismatch (~ 90 RFU) peaks. The third mix (right panel) was composed of 3X non-target allele and resulted in markedly higher mismatch peak signal intensity. Similar to the first treatment, peak height intensity did not correlate with sample allelic content (match ~ 90 RFU, mismatch ~ 130 RFU).

Real World Mix

dpFRET analysis for samples from two bone marrow transplant cases provided results similar to analysis by capillary electrophoresis (Figure 21). Case 1 (top panel) resulted in all cellular fractions displaying donor genotype for both alleles (8 and 12) tested. This was in agreement with results generated by capillary electrophoresis that detected 90-95% donor for all fractions. dpFRET testing for case 2 (bottom panel) resulted in donor genotype for all cellular fractions except granulocytes which showed a mix of both donor and recipient at approximately a 1:1 ratio. This result was in agreement with previous capillary-based testing that showed a 50% contingent of donor genotype within this sample. Additional cases were tested (data not shown) and showed similar results to Case 1. Additionally, all blinded donor and recipient allelic assignments generated by dpFRET analysis were in agreement with previously established genotypes.

CHAPTER FOUR: DISCUSSION

The goal of the research described in this thesis was to develop a unique technology for SNP and STR genotyping and explore its potential for use in forensic science. The overall objective was to minimize or alleviate the weaknesses posed by current approaches that potentially limit forensic analysis. In the process, parallel testing for other fields was incorporated to further explore the limitations posed by the developed technology and potential widespread application. When first beginning development, it is important to control as many variables as possible. Synthetic template testing was used initially to establish proof of concept for the chemistry thus removing the necessity to account for variables upstream in the process. Once the chemistry itself proved successful, the next step was to develop PCR based assays that could be utilized for general screening purposes but more importantly could test "real world" application of the approach. Following development of the complete protocol, it was necessary to establish the boundaries within which the technology was applicable. Following these successes, the next progression was extension of the approach to other DNA changes which led to application and testing for STR genotyping. Essentially, development proceeded from simple markers (SNPs) to more complex markers (STRs) with the overall goal of application to forensic science.

SNP Development and Testing

SNP Synthetic Templates

Hybridization based genotyping of changes in DNA often depend on oligonucleotide melting temperature (Tm). The Tm of duplex DNA is defined as the temperature where one-half of the

nucleotides are paired and one-half are unpaired (Wetmur 1991). Tm can be predicted using a variety of formulas with the most accurate being the thermodynamic nearest neighbor model (SantaLucia 1998). The nearest neighbor model is based on the assumption that probe hybridization energy can be calculated from enthalpy and entropy of all nearest neighbor pairs, including a contribution from each dangling end (Bommarito *et al.* 2000). Dangling ends (also known as "end effects" or "end-fraying") account for the effects seen when a shorter probe is bound to a target with flanking sequence (Hayes *et al.* 1970 and Lane *et al.* 1997). Various interactions contribute to probe/template stability, but it has been demonstrated that melting of the complex is initiated at the ends of the duplex (Doktycz et al. 1995). The results of this project suggest that this dangling end effect provides dpFRET with a higher level of resolution as compared to an intercalating dye. The goal of synthetic SNP genotype testing was to test this hypothesis and determine optimal probe design and performance limitations.

The first phase of development for dpFRET SNP genotyping involved determination of the effect of probe size on resolution. Initial testing used a synthetic library of templates that encompassed any potential change at every position complementary to the probe sequence. The most obvious result for all probe sizes tested (30, 21, and 15 bp) demonstrated that this approach is not currently capable of assigning a unique melt temperature to every potential change (position and nucleotide mutation). This is evidenced by different mutations at different positions sharing the same melt temperature. However, the capability was demonstrated for producing a differential melt temperature relative to a perfect match with the probe sequence. In other words, a mutation at two different locations within the sequence can potentially produce

the same melt temperature, but that temperature is almost always lower than a perfect match between the probe and reference sequence. Mutations at the ends (5' and 3') of the template were indistinguishable from the reference sequence for larger (30 and 21 bp) probes most likely due to inadequate "end effects" potentially due to the size of the probe. A reduction in probe size to 15 bp produced a differential melt temperature from the reference sequence for all mutations at all positions. The potential to manipulate probe melting and produce a unique melt temperature for all changes relative to a reference for larger probes is explored further in the section on incorporation of inosine.

The most likely explanation for the effect of higher resolution with a reduction in probe size is a decrease in the amount of energy required to break the bonds between the probe and template. A smaller oligonucleotide requires less energy and hence a base mismatch will have a more intense effect on melting temperature of smaller sequences. It also might be the case that end effects are amplified proportionally with decreasing probe size. In its current state, dpFRET can be applied for SNP discovery with follow-on sequencing for determination of the exact position and mutation. For purposes of SNP screening, it may be necessary to take into account design considerations for discrimination of certain targeted changes. Overall, current testing suggests that probe size should be limited to 15-30 bp depending on application and desired level of resolution.

For both the 30 and 21 bp probes, dpFRET showed higher resolution for internal template changes than SYBR Green I (intercalating dye) alone. This result lends credibility to the

hypothesized end-effects theory. It appears that when utilizing an intercalating dye internal mismatches are averaged out across the template as it melts. Any single mismatch is averaged with all matching nucleotides across a template producing a lower signal to noise ratio due to increased noise. By localization of the differential melting signal to the end of the hybrid complex (fluorophore labeled end), the effect is more significant because FRET can only occur across a limited distance. So, signal differences contributed from the mismatch remains constant, but the noise produced by dye intercalated at a distance is minimized. This would have the potential to increase the signal to noise ratio providing a higher level of resolution. Although data has been generated for one particular 30 bp sequence (variable human template library) and appears to support the intended approach, additional testing with a range of synthetic template sequences should be undertaken to further lend support for this hypothesis.

The second phase of development for dpFRET SNP genotyping was aimed at testing the limits of resolution for detecting multiple SNPs within a template sequence. Many other hybridization based genotyping systems are unable to genotype more than a single SNP per assay design. One of the benefits of dpFRET that could contribute to solving the SNP multiplexing dilemma encountered by forensic analysis is the ability to detect multiple changes within one template with a single probe design. In an effort to test the limits of this approach, a template library was synthetically generated that encompassed one to twelve SNPs in varying configurations based on a region of Cytochrome B sequence for a number of animal species. The reference and complementary probe sequence were based on human Cytochrome B with the intended application for animal species genotyping.

Results indicated the ability to detect as many as nine collective mutations within a 30 bp sequence. Beyond nine base pairs, the probe and template were not able to hybridize in a manner sufficient to intercalate dye and donate signal to the fluorophore probe for genotyping. Hence, even with 30% divergence between the probe and template, a melt signal was still generated. Similar to probe size testing on the variable human sequence library, all probe/template complexes showed a reduced melt temperature compared to the reference human sequence but were unable to classify all templates as unique due to insufficient resolution. In other words, it was possible to tell human from any other species, but the approach could not differentiate between species. This is most likely due to the fact that multiple mutations at variable positions can have the same destabilizing effect on the DNA duplex and hence would not produce a unique melt temperature. Similar to single mutation testing, additional synthetic sequence testing would also provide support for this hypothesis.

One potential contribution to forensic analysis for genotyping multiple SNPs is human versus non-human species identification. The most common method employed by forensic laboratories is amplification and direct DNA sequencing of mitochondrial Cytochrome B sequence using a universal pair of primers (Branicki *et al.* 2003 and Parson *et al.* 2000). Molecular markers other than Cytochrome B have been commonly used for broad species identification in phylogenetics including rRNA genes (Balitzki-Korte *et al.* 2005), cytochrome oxidase I (Savolainen *et al.* 2005), and various other mitochondrial and nuclear gene combinations (Bellis *et al.* 2003). Molecular approaches other than DNA sequencing including nested PCR, RFLP, and

fluorophore based detection have also been explored for forensic application to this type of analysis (Guha *et al.* 2005, Lopez-Andreo *et al.* 2005, Guglich *et al.* 1994, Blackett *et al.* 1992 and Murray *et al.* 1995). Such methods have failed to be adopted by the forensic community due to inability for universal application, lack of accuracy or efficiency, sensitivity, or additional shortcomings. Results have demonstrated that with further development, the dpFRET approach for SNP genotyping has potential for application to not only the issue of human versus non-human source attribution but could potentially contribute to mitochondrial DNA analysis, nuclear and Y-chromosome SNP genotyping and a number of other analyses.

For widespread application, an optimal SNP genotyping system should be capable of producing a unique signal for any change in a cost effective manner. These two requirements prompted follow-on developmental experiments to enhance resolution and reduce the cost for dpFRET analysis. In order to produce a unique melt temperature for any change within a template (30 bp region) using dpFRET, this would require altering the melting behavior of the probe/template hybrid complex. Most chemical additives (PEG, Urea, DMSO, Betaine, etc.) act on DNA in a sequence independent manner (Spink *et al.* 2007). In other words, there is an equal shift in melt temperature for all templates. After initial testing (data not shown) using some of these additives, it was discovered that a more sequence dependent approach was necessary. A commonly used nucleotide analogue, Inosine, was explored for its ability to alter melt temperatures with the goal of producing a unique signal for any mutation.

DeoxyInosine (dI) is a naturally occurring base that, while not truly universal, is less destabilizing than mismatches involving the four standard bases. Hydrogen bond interactions between dI and dA, dG, dC, and dT are weak and unequal, with the result that some base-pairing bias does exist in the following manner: dI:dC > dI:dA > dI:dG > dI:dT (Case-Green *et al.* 1994). It was hypothesized that this base pairing bias could differentially affect melting behavior of local and distal portions of the probe/template duplex. In other words, a mutation from C to T at one position would bind inosine in a weaker manner, and affect the melting of the nearest neighbors. It was conceptualized that incorporation of inosines at the distal end was most likely to provide this effect based on previously developed models (Watkins *et al.* 2005). Watkins et al. showed that Inosine has the potential to alter the melting behavior of a probe in different ways based on number and location of inosine bases within the probe, probe sequence, and template nearest neighbor sequence. Based on these models, it is hypothesized that through locating Inosine bases in a sequence dependent fashion, this has the potential to manipulate probe/template melting and subsequently provide a unique temperature for any change within a template. Although this would require extension of current inosine melt models, preliminary results have demonstrated that single and multiple insertions of Inosine within the probe sequence are able to alter the melting behavior of corresponding template mutations, nearest neighbor mutations as well as distal mutations. With additional experimentation and model development, inosine has potential to provide an approach for generating unique signals for all SNP changes using dpFRET.

Cost effectiveness is the other important aspect for widespread application of a SNP genotyping system. The dpFRET approach capitalizes on the use of intercalating dye whose benefit is significantly reduced cost compared to dual labeled fluorescent probes. However, there is still a need for a single fluorophore labeled probe. Due to the limited amount of probe required to generate a signal, commercial probe synthesis is cost effective for screening many samples. However, in cases where sample throughput is low but many different sequences would need to be tested per sample (such as mutation scanning) that would require the use of many different probes, an alternative approach is required. As opposed to commercial synthesis, enzymatic probe production provides a potential cost effective alternative.

Terminal transferase (TdT) is a template independent polymerase that catalyzes the addition of deoxynucleotides to the 3' hydroxyl terminus of DNA molecules. TdT can be used to incorporate a fluorophore labeled nucleotide at the 3’ end of an oligonucleotide probe. The FRET combination tested used Texas Red as the acceptor fluorophore that was incorporated by TdT as a dUTP. Unfortunately, TdT labeling has been shown to be extremely inefficient at incorporation of this particular fluorophore (Igloi 1996). This may potentially be due to steric hinderance of the active site for the enzyme using this particular fluorophore/nucleotide combination. Although it was shown that TdT produced probe was capable of reproducing data generated with a commercially produced probe, inefficiency of incorporation limits the amount of labeled probe produced hence limiting the cost effectiveness of the approach. Additional chemical end labeling strategies were tested (ULYSIS - data not shown) and also proved unsuccessful. Future

studies should examine alternative fluorophore/nucleotide combinations that show higher incorporation efficiencies by TdT.

Continued development of dpFRET will facilitate reaching the goal of a cost effective SNP genotyping system capable of screening and discriminating any mutational change. Inosine probe melt manipulation and cost effective TdT labeling were both steps towards a unique concept of "Relative Sequencing." This concept is detailed in Figure 22 and consists of the following approach. First, a set of probes is designed against a reference sequence of interest. Each probe encompasses approximately 30 base pairs of sequence and the full complement of probes would stretch across and encompass the sequence of interest. For example, if one were interested in looking for SNPs in 270 base pairs of human mitochondrial Dloop (control region) sequence, nine 30 bp probes would be designed that covered the region of interest. Testing of multiple samples with each probe would produce a melt temperature either matching or lower than the reference sequence. Any probes that produced a lower Tm would signify the presence of a SNP relative to the reference sequence at that probe position. With the current state of dpFRET, follow-on sequencing would be needed to identify the exact mutation and position of the SNP. For applications to forensic analysis, the victim would serve as the reference sequence with potential perpetrators tested using the described approach. The significant contribution of this approach would be the ability to screen a multitude of potential samples at a significantly reduced cost compared to standard sequencing. All samples matching the victim could be disregarded and the focus could be placed on probative samples for full laboratory analysis. Other fields could benefit from a similar approach. For example, "Relative Sequencing" could

be useful for screening large numbers of clinical samples for either SNP discovery (i.e., a change in a gene promoter) or screening following identification of a candidate SNP. With further development and a relativistic approach to analysis, dpFRET has potential to significantly impact approaches for SNP genotyping.

SNP "Real World" Samples

Synthetic testing was used to develop and define the limits (probe size, assay optimization, etc.) of dpFRET SNP genotyping. The next stage comprises development of a complete protocol that incorporates PCR amplification of target sequences. Initial development focused on a haploid marker (mitochondrial *Cytochrome B*) to minimize melt curve complexity. This was followed by development for diploid marker (nuclear *Mhc DRB*) testing to explore the ability of the assay to discriminate two different alleles within the same individual. Both phases of development incorporated unique primers designed for this project based on alignments of published sequence for multiple species. Results to date have shown the assays to be successful for amplification of multiple species with potential utility in a number of fields including forensics for species and individual identification. Following amplification optimization, dpFRET SNP genotyping assays were designed and tested for both haploid and diploid markers.

The most obvious result from Haploid Cytochrome B species testing was a lower resolution of amplicon melt peaks (Range = 5 degrees) as compared to the melt peaks generated by a dpFRET probe (Range = 32 degrees). Unique identification of all species was possible through analysis of probe melt peaks in contrast to melt peaks generated from the PCR amplicons themselves which

were unable to resolve most species identifications. Similar to the results obtained by synthetic probe size testing, the reduction in size for the signal generating strand from amplicon to probe size resulted in much higher resolution. Also similar to synthetic species testing, the real world python sample showed no probe peak due to sequence divergence beyond the 10 bp or 30% limit. It should also be noted that the no template control (NTC) resulted in a broad probe melt peak but displayed no amplicon peak. This phenomenon has been reproduced in follow-on development and it was found that that the source of this peak is primarily due to excess probe concentration. It is hypothesized that the probe forms a probe/probe dimer that produces a signal at a significantly reduced melt temperature and characteristically results in no amplicon peak. Optimization of probe concentration has shown to alleviate this effect. Two melt peaks (probe and amplicon) is a unique benefit to the dpFRET approach that provides an inherent amplification control that can be used to further qualify results.

The amplicon peak is most likely due to the strong fluorescent signal generated by SYBR Green I whose emission tail end falls within the detection bandwidth. In other words, not only does the probe/template hybrid duplex contribute signal from FRET, but fluorescent signal is also derived from the intercalation of dye by the amplicon itself. It is this additional amplicon signal that can be used as a qualification of positive amplification in a manner similar to standard intercalating dye melt curve analysis. It was this additional information provided by the amplicon peak that led to the classification of the probe/probe dimer signal as noise. It is also important to note that the amplicon melt peaks produced early in development for species testing included a small shoulder peak. It was discovered through follow-on testing and optimization that this shoulder

was due to a minor population of unlabeled probe resulting from incomplete synthesis that participated in amplification and production of a secondary specific product. Testing consisted of both reductions in probe concentrations and addition of small amounts of EDTA with the probe post-amplification (data not shown). Both approaches were successful by limiting reagents necessary to produce the anomalous result. The method of reducing probe concentration lowered the minor population of unlabeled probe below the point at which it could contribute to amplification. The addition of EDTA chelated magnesium required by the Polymerase enzyme as a cofactor for amplification. Optimization and development for *Cytochrome B* species testing provided valuable contributions that were then applied to a more complex diploid locus.

The diploid *Mhc DRB* marker has been used for a number of applications including phylogenetic (Chardon *et al.* 1999) and biomedical studies (Doxiadis *et al.* 2001). It was determined that this marker could also have potential for paternity analysis in a broad range of animal species. With this goal in mind, a universal assay for amplification was optimized. Preliminary dpFRET testing of this marker used human sequence as a reference for probe designs and samples from known Humboldt Penguin families for testing. Paternity results had already been established by Brookfield Zoo using Southern Blot hybridization and Jeffreys' VNTR probes and confirmed by Zoo keeper observations. Southern analysis is a labor intensive process and prompted efforts for development of dpFRET as an alternative. An optimal approach would be use of a single probe set for SNP genotyping designed against a single reference. In this case, probes designed against human sequence were used to test samples from another species. It was logical to assume that as

long as divergence remained below ~ 30% as demonstrated by synthetic testing, this approach would be feasible.

Results from human probe testing showed that probes were able to hybridize to a different species sequence (Humboldt Penguin) and still maintained the ability to resolve differences between individuals. In other words, although differential SNPs from the human design existed that were conserved among all Penguin samples tested, differences between Penguin individuals could still be resolved. For regions of the amplicon that were more highly divergent from human, species specific probes were designed against a Humboldt Penguin reference sequence which showed better resolution than human probe sequences for heterozygote individuals. Overall, dpFRET was successful for analysis at a diploid locus.

Results from the *Mhc DRB* experiments provided three important pieces of information that helped further define the limit and application of this approach. First, the capability to resolve multiple alleles (heterozygote) within a single sample with a single assay was demonstrated. As opposed to other approaches (Taqman), dpFRET has the potential to minimize the amount of sample consumed for testing. Second, by cross-species application of one set of probe sequences the flexibility of both assay design and broad application was also demonstrated. This is particularly important in fields like conservation biology where current studies require assays specific to each and every animal tested. Finally, it was demonstrated that data equivalent to current approaches could be generated using dpFRET that requires significantly less resources with faster time to results. Future development should include testing on additional species

which would further support this approach to contribute to a number of different fields including not only forensics, but conservation and population biology as well.

The need to define the limits of sensitivity of a genotyping approach is important in many fields, most particularly forensics. Results for human dpFRET SNP genotyping demonstrated a preliminary established detection limit of a single copy for both homozygote and heterozygote samples. Although this result will need to be confirmed with additional testing, it is not surprising based on the amplification protocol. The current version of dpFRET uses 50 to 80 cycles of amplification depending on the intended approach. Unpublished claims of reliable single copy detection have been made using similar numbers of cycles involving other chemistries (http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents /generaldocuments /cms_039235.pdf). It is hypothesized that low levels of sensitivity are possible due to probe based detection of specific amplicon that is essentially “blind” to generation of non-specific product that can significantly impede size based analysis. Signal generation produced by probe hybridization in dpFRET is capable of capitalizing on this same approach. Only non-specific product with less than 30% divergence will produce a probe melt peak with dpFRET analysis. An example of this phenomenon is demonstrated in Figure 23. Products generated for *Cytochrome B* probe testing were analyzed by agarose electrophoresis. In addition to the 350 bp specific product, multiple non-specific products are also amplified that showed no signal during previous genotyping with dpFRET. These preliminary results are significant for fields like Forensics and Molecular Pathology were sensitive detection are typically required for samples with low target concentration.

Overall, development and testing of dpFRET for SNP genotyping resulted in marked success with implications for limitations to the approach. Preliminary work demonstrated the approach is robust for low copy number detection within a sample with no apparent allelic dropout. The approach proved successful at genotyping both a haploid and diploid locus and displayed highly flexible design strategies capable of detecting single or multiple SNPs using a single assay. These results suggest this approach has potential to contribute to advancing the use of SNPs in forensic analysis by providing the capability to genotype multiple changes per reaction that subsequently minimizes the amount of sample consumed per test. Future development efforts aimed at this goal should also focus on obtaining higher levels of resolution and reduced reagent costs through previously described alternative approaches. Following completion of development of dpFRET for SNP genotyping, the next step was to examine whether the same approach could be applied for genotyping repetitive sequences known as STRs.

STR Development and Testing

STRs are composed of repetitive sequences and can also be referred to as microsatellites, variable number tandem repeat and a host of other terms. The application of these markers is based on the concept that a genotype can be generated based on the number of repetitive core sequences that varies between individuals. The greatest advantage to these markers is their ability to produce multiple alleles per marker providing more information per test than biallelic SNPs. STRs are the accepted marker of choice for forensics, conservation biology, and more recently are being valued in clinical studies for the ability to monitor progression of cancer

(De Schutter *et al.* 2007 and Halachmi *et al.* 2007) and aid in monitoring transplant success (Lion 2003). Information is typically generated by amplification followed by sizing of the alleles by capillary electrophoresis. Not only is this approach subject to a number of artifacts, but also requires specialized equipment and a high degree of training to generate genotypes. An alternative approach to interrogate these highly informative markers would provide a significant step in advancing analyses using STRs.

STR sequences can vary in core repeat content but typically have a similar overall structure. Conserved flanking sequences are used to amplify a repetitive region composed of a core repetitive section. This core repeat can be composed of either a single sequence (simple repeat) or multiple sequences (complex repeat) with additional SNPs potentially present within the repetitive and/or flanking sequences. With the success of dpFRET for SNP genotyping, the next logical step in development consisted of applying the same approach for STR genotyping. Initial concepts were aimed at using the core repeat sequence as a probe followed by measurement of signal intensity to differentiate the number of repeats (data not shown). After extensive testing, it was concluded that this approach was unsuccessful due to irreproducible probe hybridization. It was determined that what was required was a presence/absence based analysis. It was this alternative approach that ultimately proved successful for dpFRET genotyping of STR loci.

Following extensive testing of different design strategies, a basic strategy emerged. The locus was divided into three regions composed of a “reporter flank,” “core repeat region,” and “anchor flank.” It was hypothesized that the anchor flank would be designed with a higher Tm than the

fluorophore labeled reporter flank. This would hopefully favor hybridization of the anchor region, followed by hybridization of the core repeat region and finally the reporter flank. Upon denaturation, a higher melt signal would be generated for a perfect match versus an imperfect match permitting the correlation of sample allele content with the number of repeats contained within the probe. If the probe were to encounter a mismatch with the template sequence, the result would be imperfect hybridization with the reporter region of the probe resulting in decreased signal intensity and more importantly a lower melting temperature. This would primarily be due to the reduction in bonding energy of the template/probe complex similar to what was seen with SNP mutation testing. A number of different designs were tested for varying lengths of both the reporter and anchor flanks (data not shown). Shorter flanks resulted in partial melt peak separation between a matched and mismatch template. Ultimately, a calculated Tm difference of approximately 10 to 15 degrees between the reporter flank sequence and anchor flank sequence proved to be a good indicator of successful probe design for STR genotyping. Following completion of a probe design strategy, extensive testing of both a simple and complex locus encompassed the next stage of development.

Assay development for STRs followed a similar approach as dpFRET SNP testing by moving from a more simple (TPOX) to a complex (D3S1358) locus. TPOX is one of the loci typically employed for individual identification as part of the collection of loci known as CODIS and is located on chromosome 2 within intron 10 of human thyroid peroxidase gene. Validated primer sequences from the Promega PowerPlex kit were used to remove any ambiguity potentially generated by in-house designs. Following brief optimization for 80 cycle amplifications (data not

shown), probes were designed using the previously established strategy for the most common alleles (8-12 repeats). Blinded samples previously genotyped using standard forensic protocols (capillary electrophoresis) were provided by the Johnson County Criminalistic Lab and analyzed by the dpFRET STR approach. dpFRET produced the same genotypes for TPOX and in a fraction of the processing and analysis time required for the current approach.

Similar design strategies and testing were used for the complex locus D3S1358 which is located on chromosome 3, not known to be located within a coding region and is also one of the core loci within CODIS. Results were equally successful and provided higher resolution than current approaches. This demonstrated that dpFRET analysis can be accomplished using existing primer designs and amplification strategies for both simple and complex loci. It is important for forensic applications that the only difference between dpFRET and current protocols is that allele detection is accomplished by hybridization as opposed to size-based genotyping.

Although it produces equivalent results, dpFRET does not suffer from many of the shortcoming of size-based detection of STRs by capillary electrophoresis. Current approaches are subject to a number of biological and technological artifacts. Biological artifacts include stutter peaks, 3' (A) additions and tri-allelic patterns. These all result in additional peaks on a chromatogram that complicate analysis of single source samples. Although none of these effects were evident in any testing to date, it would be hypothesized that only stutter peaks could potentially be detected as mismatched signal by dpFRET analysis. 3' (A) additions and tri-allelic patterns theoretically would not alter probe hybridization due to lack of interaction with the probe. As dpFRET does

not utilize capillary electrophoresis, it is not subject to any of the technological artifacts seen with current approaches. With the possible exception of stutter peaks, dpFRET has the potential to provide forensics and other fields with a more rapid objective analysis of STR loci that alleviates the weaknesses inherent in the current detection approach.

Results for D3S1358 locus testing suggests that dpFRET has the potential to provide higher resolution of complex STR markers than capillary electrophoresis generated profiles. A complex locus with more than one core repeat has the potential to generate the same size product with different alleles. For example, D3S1358 17 and 17' are different alleles but cannot be differentiated by size. Results generated using 17 and 17' specific dpFRET probes were able to differentiate between these two different genotypes due to differential probe hybridization (see Figure 18—Individual 11). Additional support is required to prove this hypothesis. Future development for complex loci will necessitate the cloning of amplification products followed by sequenced verification of sample allelic content.

No matter the approach for STR genotyping, allelic dropout is an important consideration for analysis of trace level samples. This phenomenon is due to preferential amplification of one allele in low concentration samples. Due to the potential for allelic dropout, it is important to quantify starting material prior to capillary electrophoresis based testing. Results of sensitivity testing using dpFRET showed no allelic drop out for a heterozygote sample. Ten-fold diluted concentrations of starting material were tested and preliminary results demonstrated no marked change in final fluorescent signal for diluted samples. This result is expected due primarily to the

amplification protocol. Current forensic protocols incorporating capillary electrophoresis based genotyping typically utilize 30-40 cycles for PCR amplification as opposed to dpFRET protocols that utilizes 50-80 cycles of amplification. The increased number of cycles for dpFRET potentially alleviates sampling error that can be seen with fewer cycle amplification approaches. In other words, additional amplification opportunity (cycles) is provided to produce an equivalent signal for both alleles. Barring an explanation, elimination of the need for pre-quantification and a solution to allelic dropout has enormous importance for forensic analysis.

Both allelic drop out and pre-quantification are also important contributing aspects to mixed sample testing. Results for both laboratory generated and real world mixes demonstrated dpFRET's potential to detect samples containing more than one genotype. Success with producing equivalent results to size based testing for the percent donor contribution to cell fractions for bone marrow transplant provides evidence that real world application is possible with more development. It was noticed that results were somewhat variable for correlation of peak height intensity with true allelic content. A potential explanation for these results is sampling error due to the amplification approach. The protocol that was used for amplification of both laboratory and clinical samples was 40 cycles of double-stranded amplification followed by introduction of a small portion of this reaction into another reaction consisting of 40 cycles of single stranded amplification containing only one primer. Due to the need to test multiple probes per samples, this method was used in an effort to minimize the amount of sample used for testing. Following the 80 cycles of amplification, allele specific probes are supplemented in each

reaction. This minimizes the amount of sample required but also provides potential for introduction of sampling error that could result in peak height variability. Protocols were also tested based on closed tube 80 cycle amplification protocols (data not shown) that demonstrated less variability and better correlation with signal intensity. Unfortunately, this would require multiple aliquots of sample for testing of multiple alleles at each locus. Although limited success for dpFRET mixed sample testing was demonstrated, there is strong potential for obtaining better correlation provided the continuation of additional protocol development with particular attention to sampling methodology.

Optimization of STR Analysis

Current strategies for dpFRET STR analysis are based on standard melt curve analysis of each potential allele. Although proven successful, this approach requires multiple reactions per locus, additional time for analysis, and acquisition of relatively large data sets. These limitations prompted exploration of alternative methods to either reduce the number of reactions required per locus and/or further simplify the melt curve analysis required to differentiate the presence/absence of an allele.

A reduction in the number of reactions required to genotype an individual at a locus would necessitate the ability to genotype with a reduced number of probes. This approach would require moving from a match/mismatch based analysis to a more classical melt based analysis similar to genotyping SNP mutations. Early results demonstrated variation in the mismatch peak melt curve that appeared potentially correlative with the mismatched allele present in the sample.

An example of this phenomenon is depicted in Figure 24. Four individuals all heterozygous for a 9 repeat allele showed the expected match peak when tested with a 9 repeat probe. For their mismatched alleles, two of the individuals contained an 8 repeat and two individuals contained an 11 repeat allele. Results showed a unique melt pattern for the mismatch peak differentially based on the mismatched allele (either 8 or 11). In other words, it appeared that a higher level of discrimination was possible beyond a basic presence/absence type analysis. Similar results were generated with other repeat probes and it was determined that higher repeat number probes resulted in better resolution of mismatched melt peaks. For example, testing with an 11 repeat probe (Figure 25) demonstrated potential to differentiate the full allelic complement of a sample beyond a simple match/mismatch based analysis. The potential to generate an STR genotype for a sample using dpFRET and a minimal number of probes appears likely but will require the use and development of higher resolution equipment and curve fitting analysis.

Classical melt curve based analysis requires time and additional resources to generate a multitude of data points for every temperature point along the curve. Current dpFRET STR analysis produces distinct match/mismatch melt peaks separated by approximately 3 to 4 degrees. To reduce the time and complexity of analysis a minimal number of fluorescent data points can be taken at three temperature points; (1) prior to probe/template denaturation (2) a point midway between melting of a match and mismatched hybrid complex, and (3) following complete denaturation. By comparing the slope ratios between these points (1 to 2 and 2 to 3), a more rapid quantitative method for STR genotyping is possible that requires only three temperature measurements. This method of analysis is depicted in Figure 26 and is capable of

genotyping both homozygotes and heterozygotes. With careful design, this same analysis can potentially be applied for all probes at multiple loci further simplifying analysis. This would necessitate careful control of probe melting temperatures based on reporter and anchor flank sequence design. Further development of this approach could improve the speed and reduce the complexity of dpFRET STR testing as compared to current classical melt curve analysis.

Conclusion

The benefits to using dpFRET for discovering and screening changes in DNA are numerous. It is less costly than many other approaches for SNP and STR genotyping due to the use of an intercalating dye and a probe with a single fluorophore. Probe design is extremely flexible and initial results suggest it to be somewhat sequence independent. Equipment requirements are minimal needing nothing more than is required for basic real-time PCR. Analysis is more objective than other approaches and is amendable to automation. Application of this new approach has the potential to alleviate many of the shortcomings of current forensic approaches and could potentially be applied to any field in need of examining changes in DNA.

APPENDIX A: FIGURES

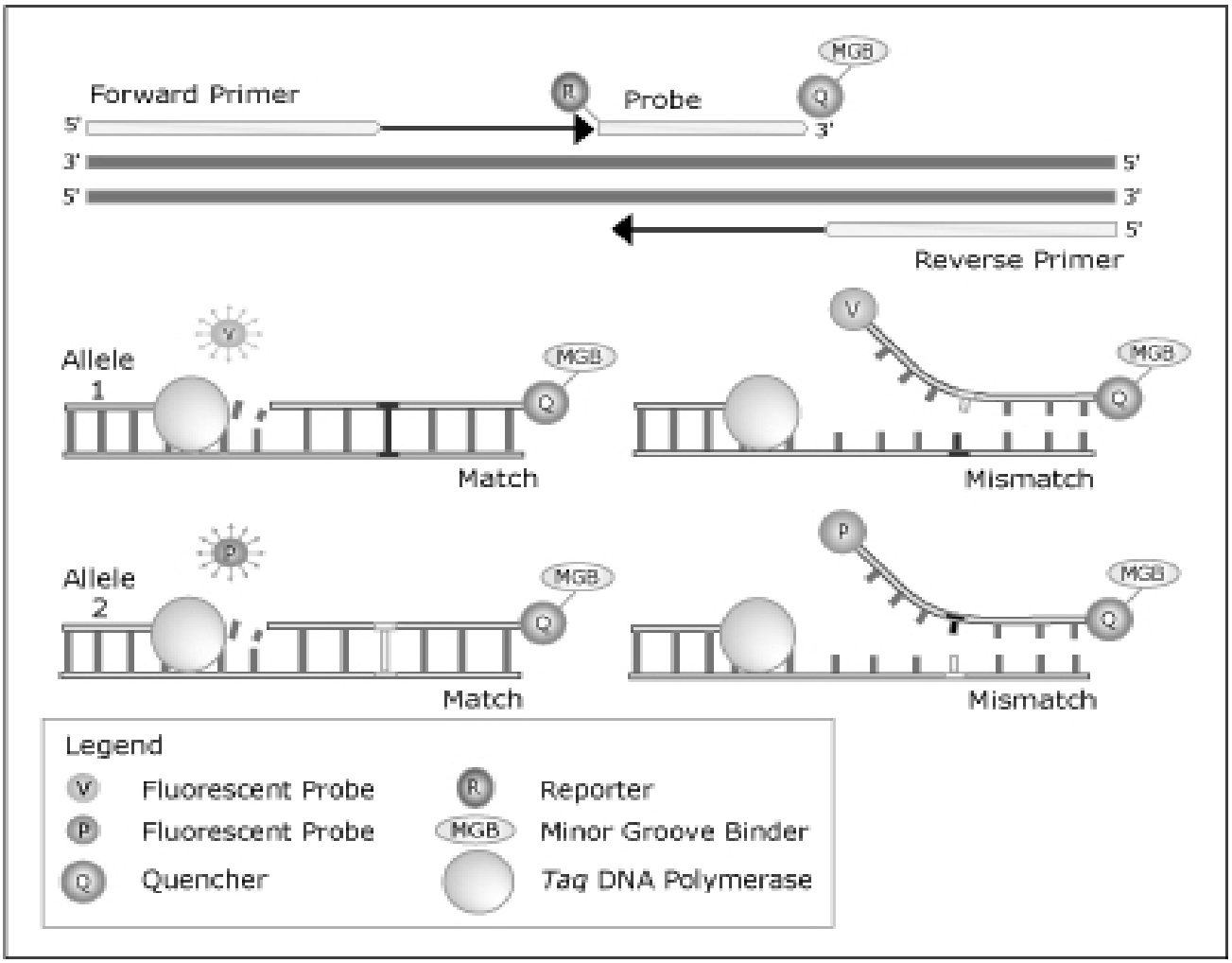
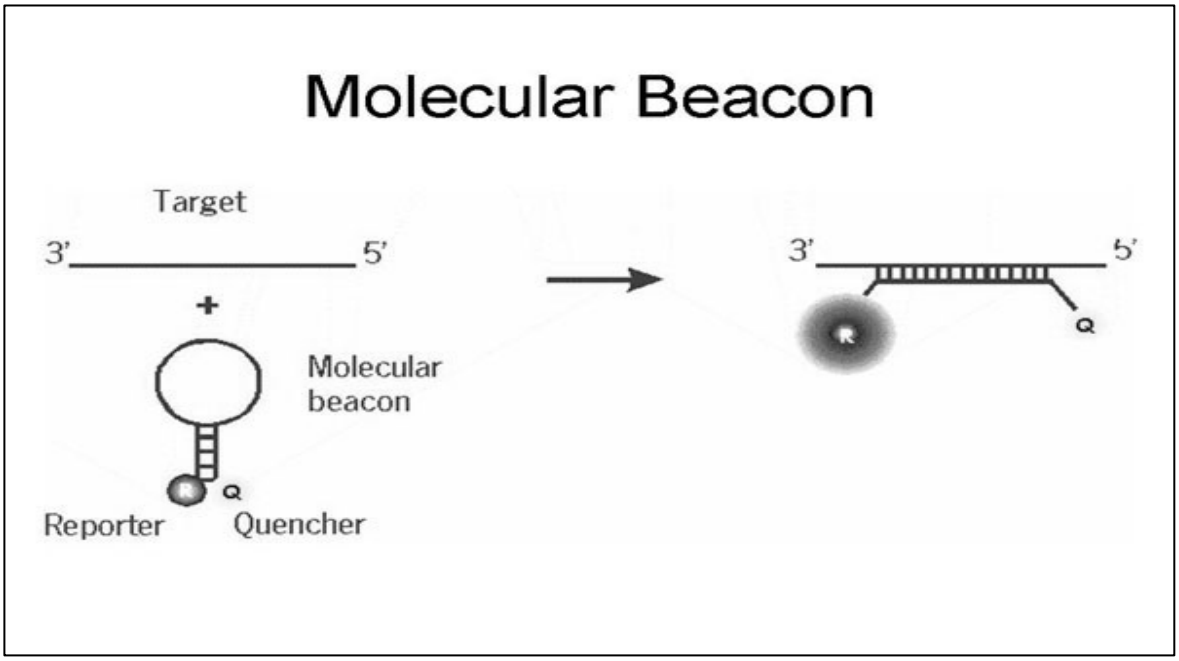


Figure 1: Diagram of Hydrolysis Probe (Taqman) SNP Detection
(Source: http://www.servicexs.com/plaatjes/TaqMan_AD_SNP_assay.jpg)



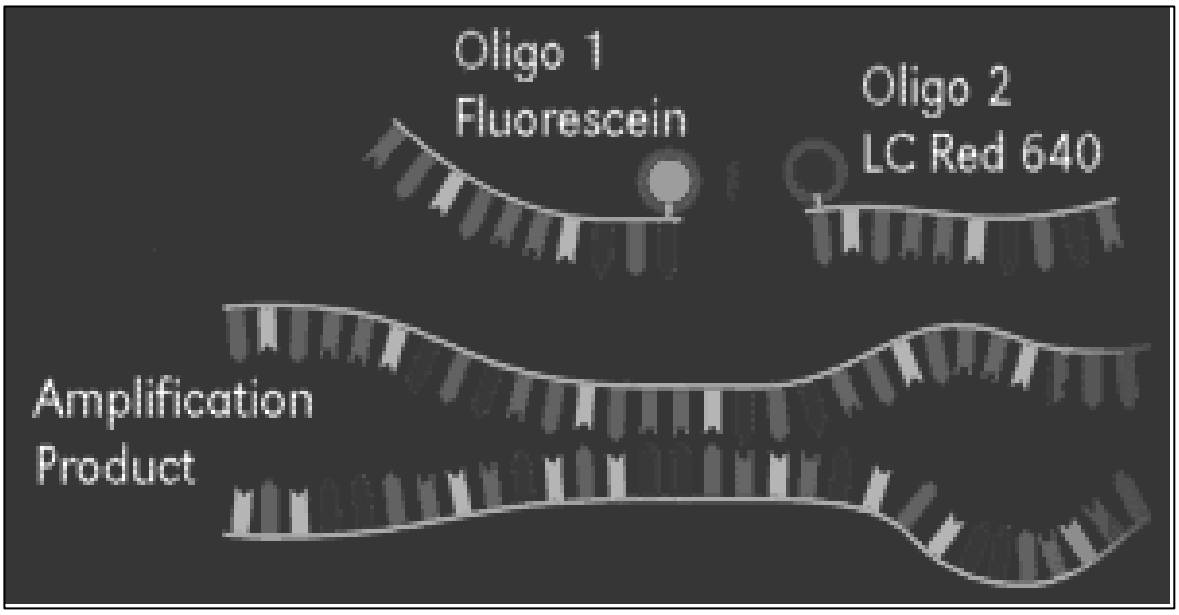


Figure 2: Diagrams of Common Hybridization Probes including Molecular Beacons (top panel) and Dual Labeled Hybridization Probes (bottom panel)
[Sources: (Top panel): http://documents.plant.wur.nl/pri/biointeractions/images/slide9.jpg (Bottom panel): http://www.gene-quantification.de/hyb08_01.gif]

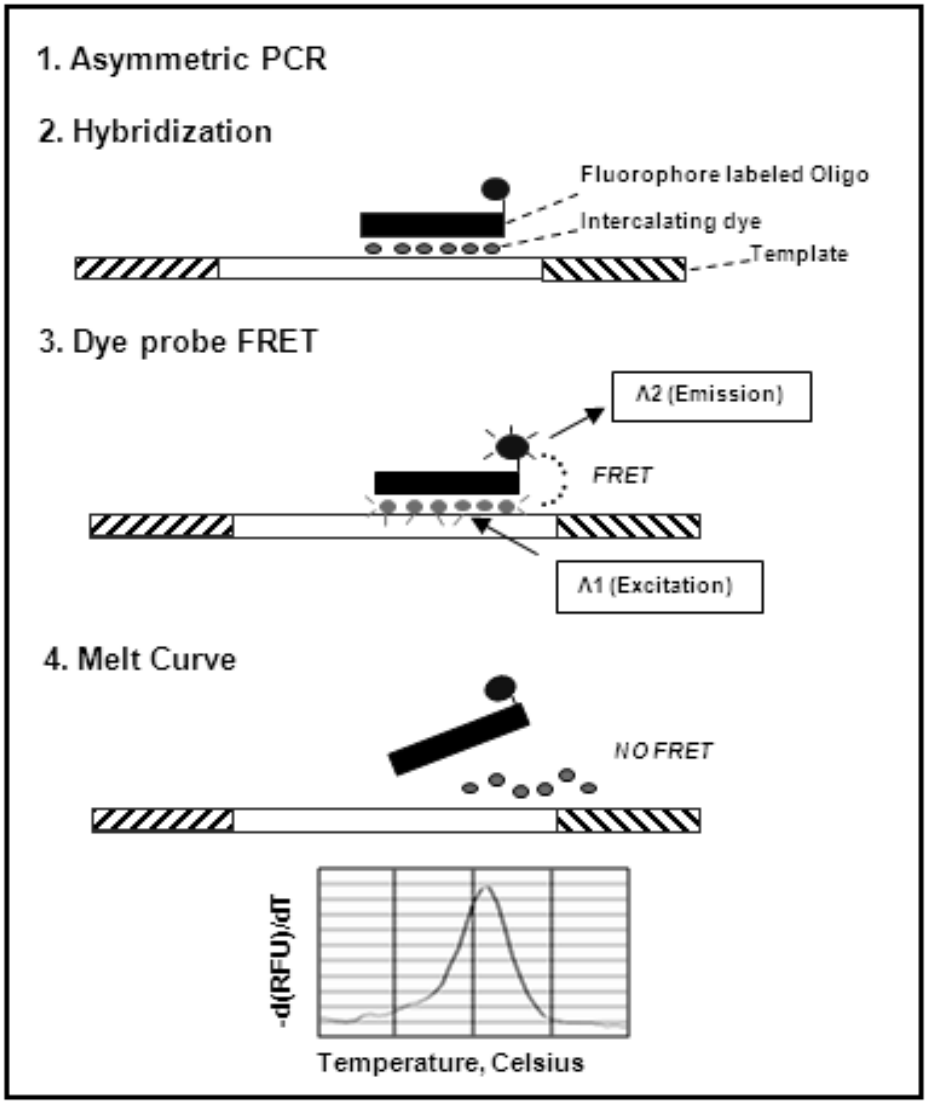


Figure 3: dpFRET SNP Experimental Approach

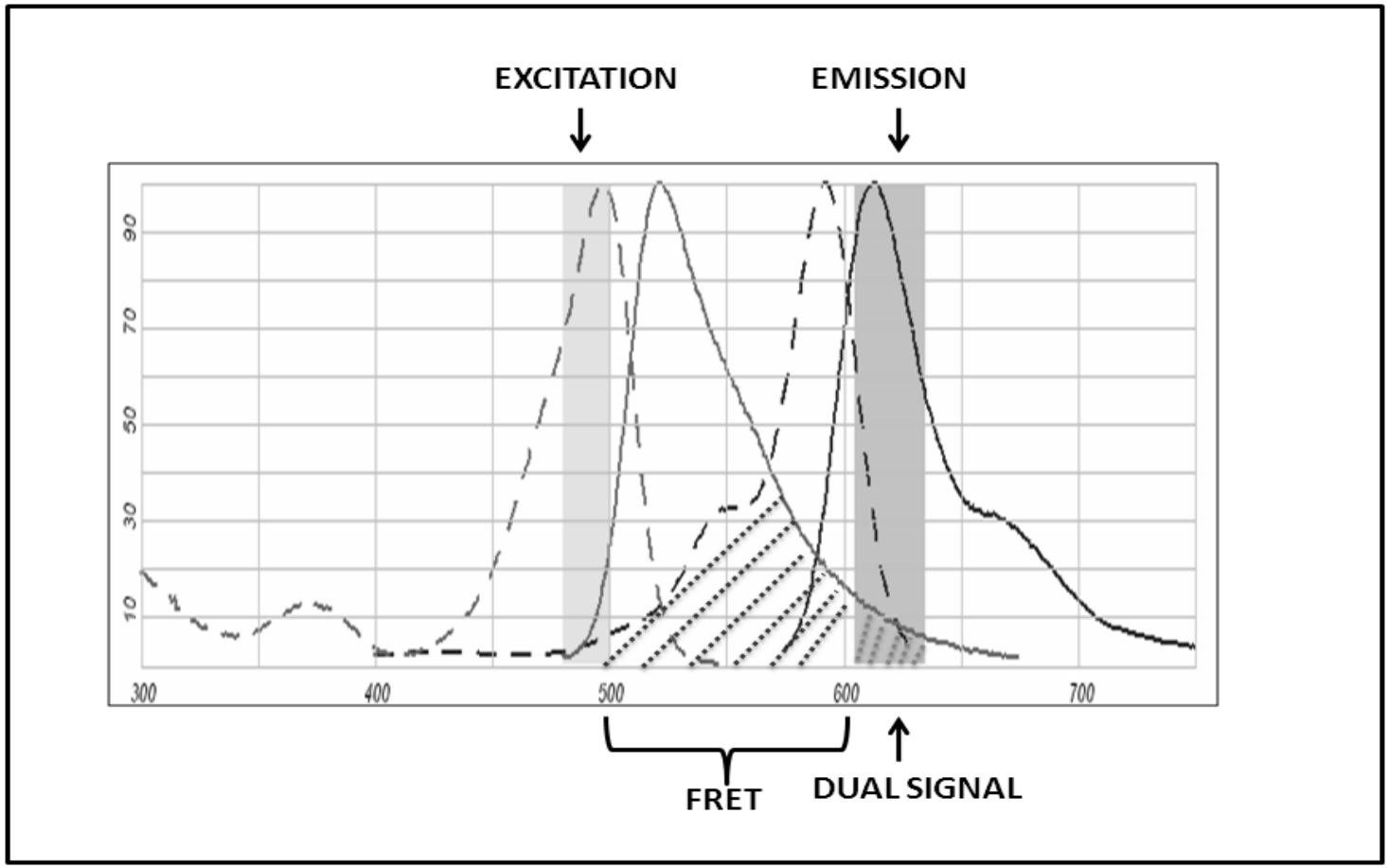


Figure 4: FRET Excitation and Emission for SYBR Green I (Left) and Texas Red (Right)
FRET region of the spectrum (dark hashed), region of emission overlap (light hashed), and excitation (light solid)/emission (dark solid) band widths are shown

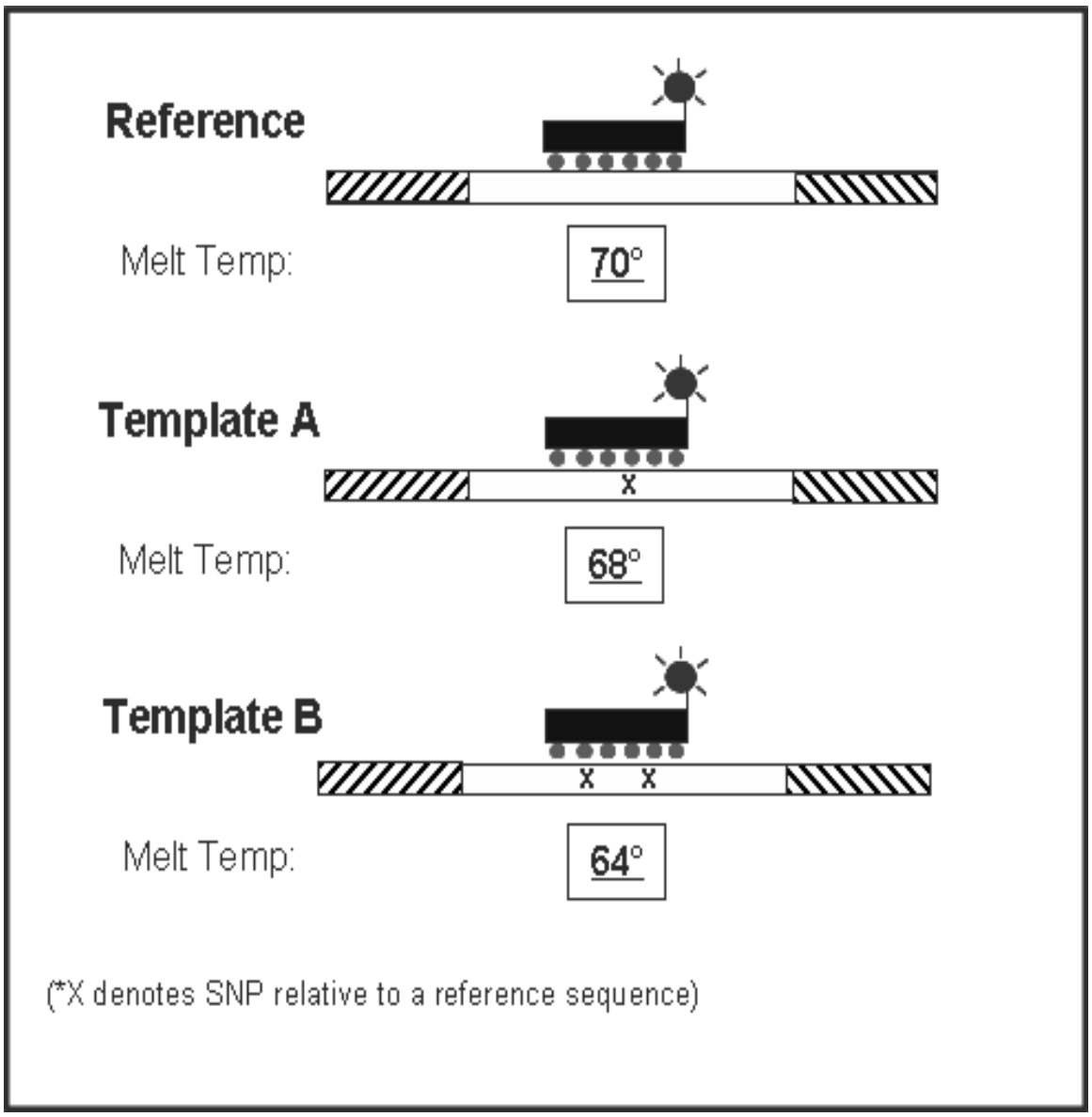


Figure 5: dpFRET Experimental Approach for Genotyping Single and Multiple SNPs

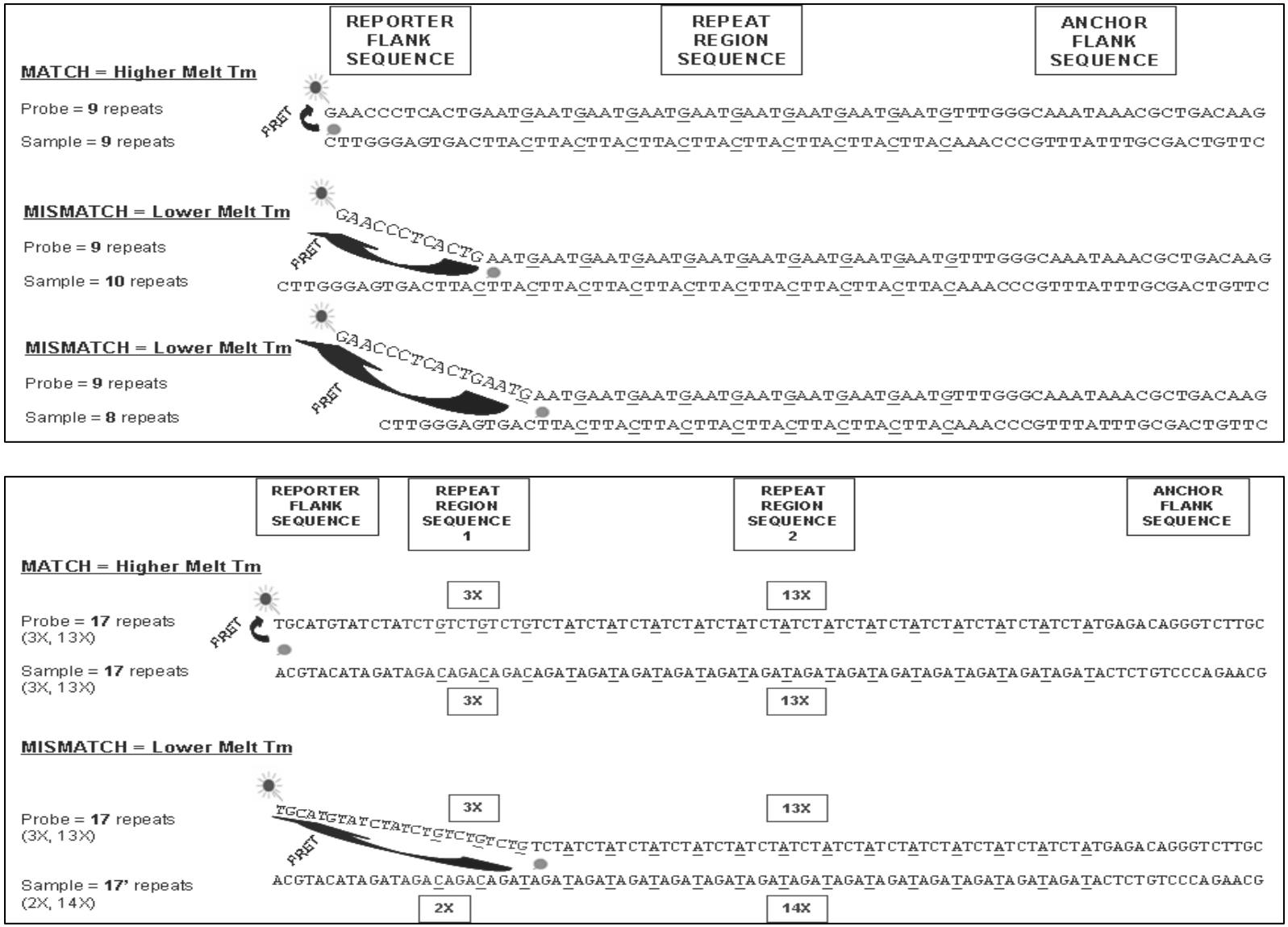


Figure 6: dpFRET Experimental Approach for Genotyping Simple (top panel) and Complex (bottom panel) STRs

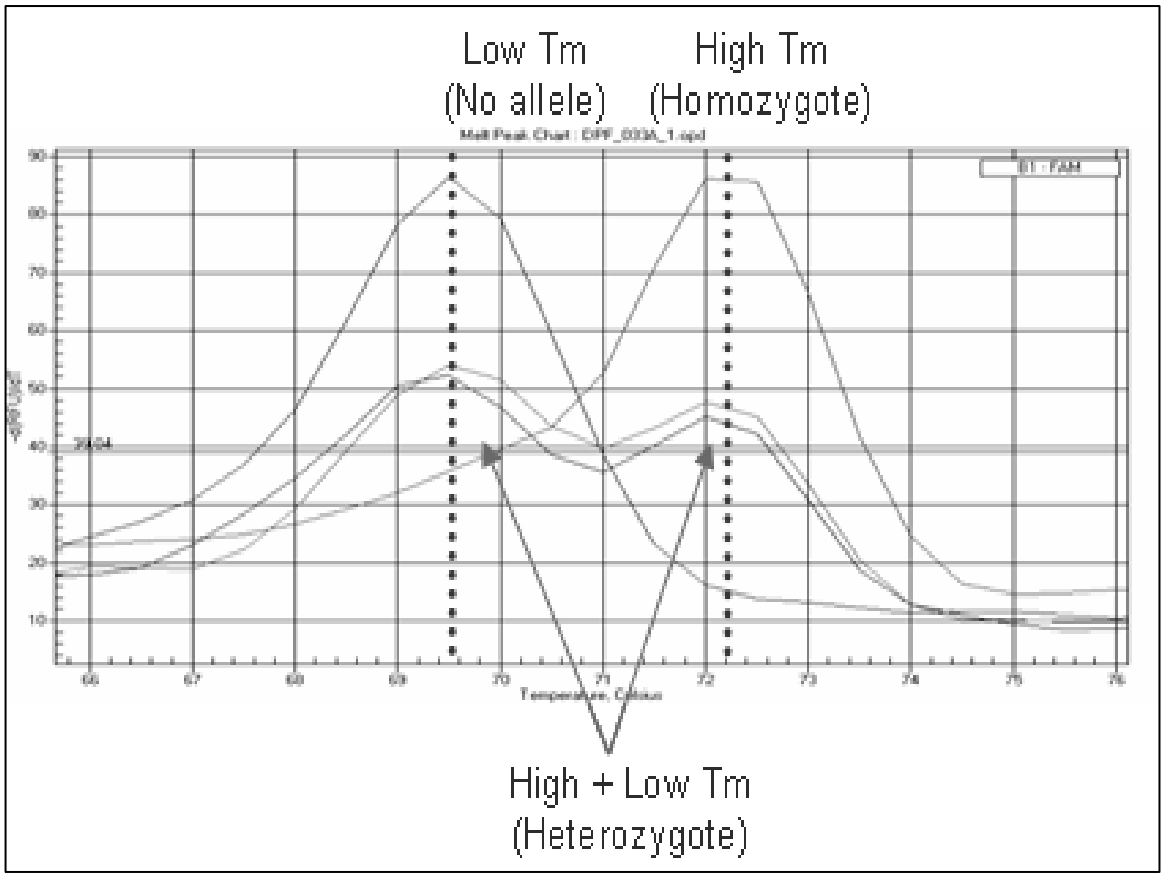


Figure 7: Example of Experimental Results Generated by dpFRET STR Analysis of Homozygous and Heterozygous Samples

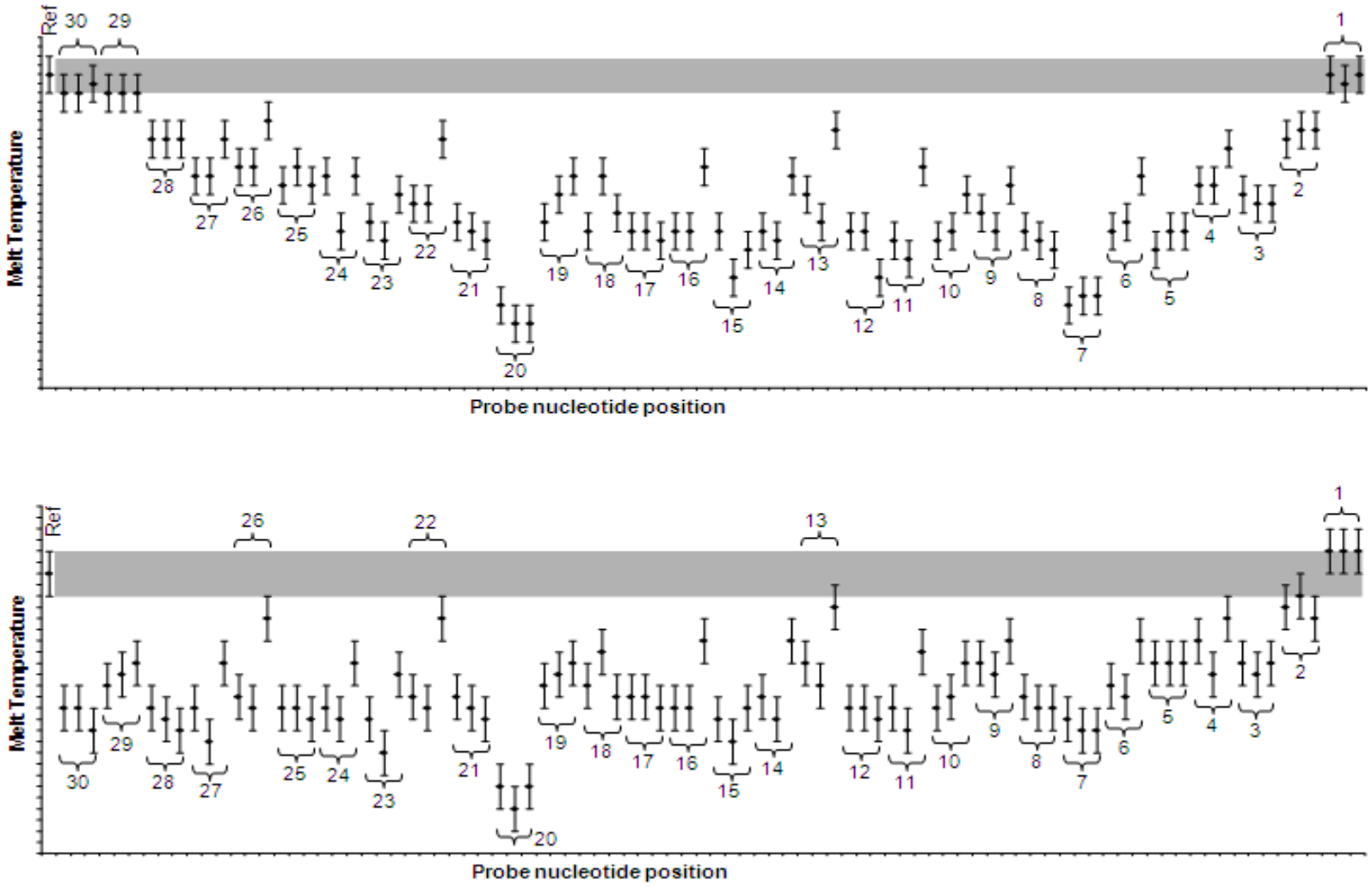


Figure 8: Results for synthetic resolution testing of 30 bp labeled (top panel) and unlabeled (bottom panel) probes
Reference sequence temperature range (grey bar) is highlighted (error bars = ± 0.4 degrees).

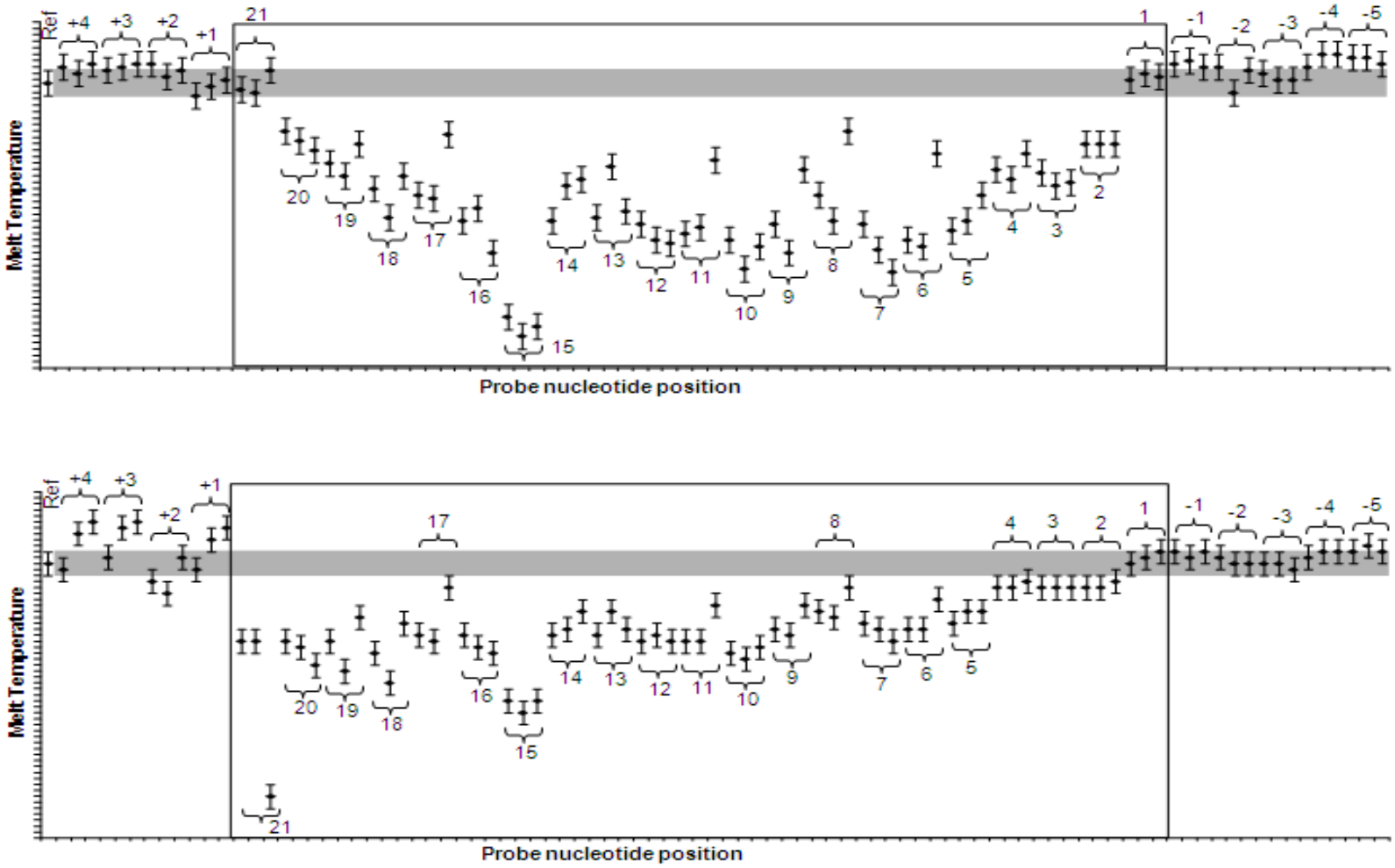


Figure 9: Results for Synthetic Resolution Testing of 21 bp labeled (top panel) and unlabeled (bottom panel) probes
Reference sequence temperature range (grey bar) is highlighted (error bars = ± 0.4 degrees). The portion of templates encompassed by the probe design is located within the box with effects on flanking sequence mutations shown outside the box.

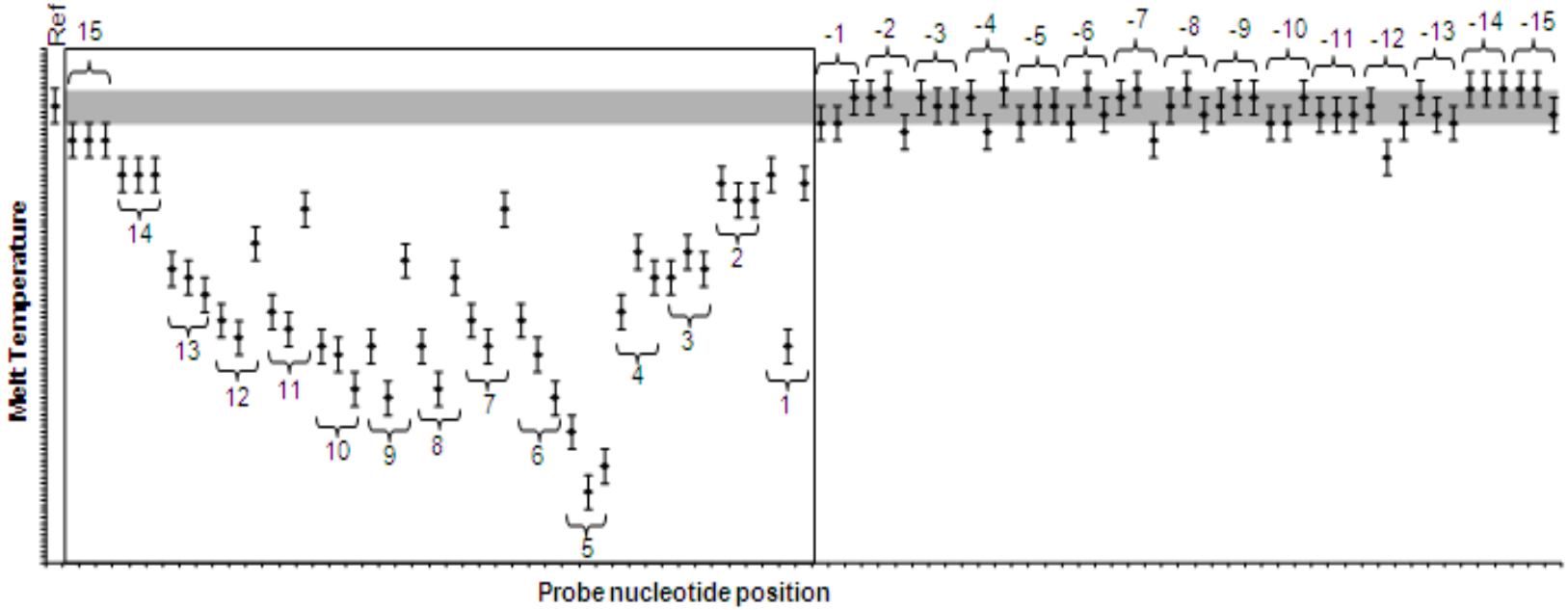


Figure 10: Results for Synthetic Resolution Testing of a 15 bp Labeled Probe

Reference sequence temperature range (grey bar) is highlighted (error bars = ± 0.4 degrees). The portion of templates encompassed by the probe design is located within the box with effects on flanking sequence mutations shown outside the box.

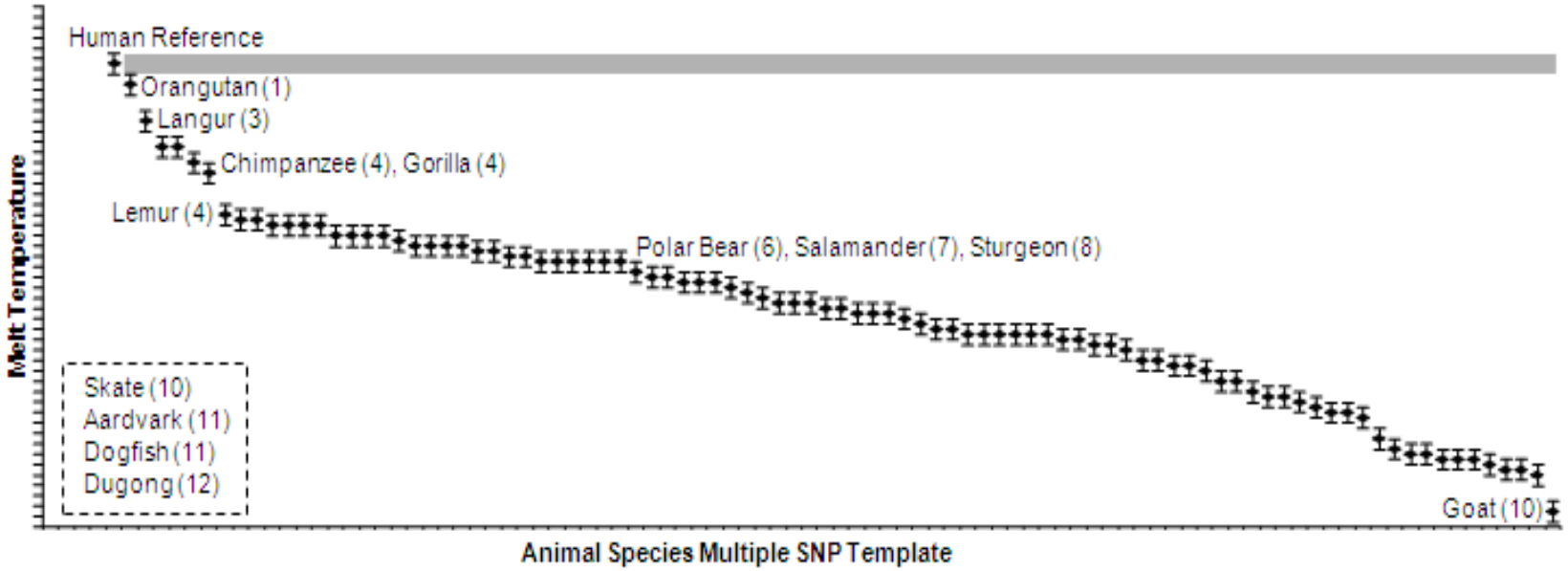


Figure 11: Synthetic SNP Resolution Testing of Multiple Animal Species Templates

Select species are depicted to illustrate divergence with number of SNPs relative to the human reference sequence shown in parenthesis. Four templates did not generate a melt curve (Skate, Aardvark, Dogfish and Dugong). Human reference sequence temperature range (grey bar) is highlighted (Error Bars = ± 1.0 degree).

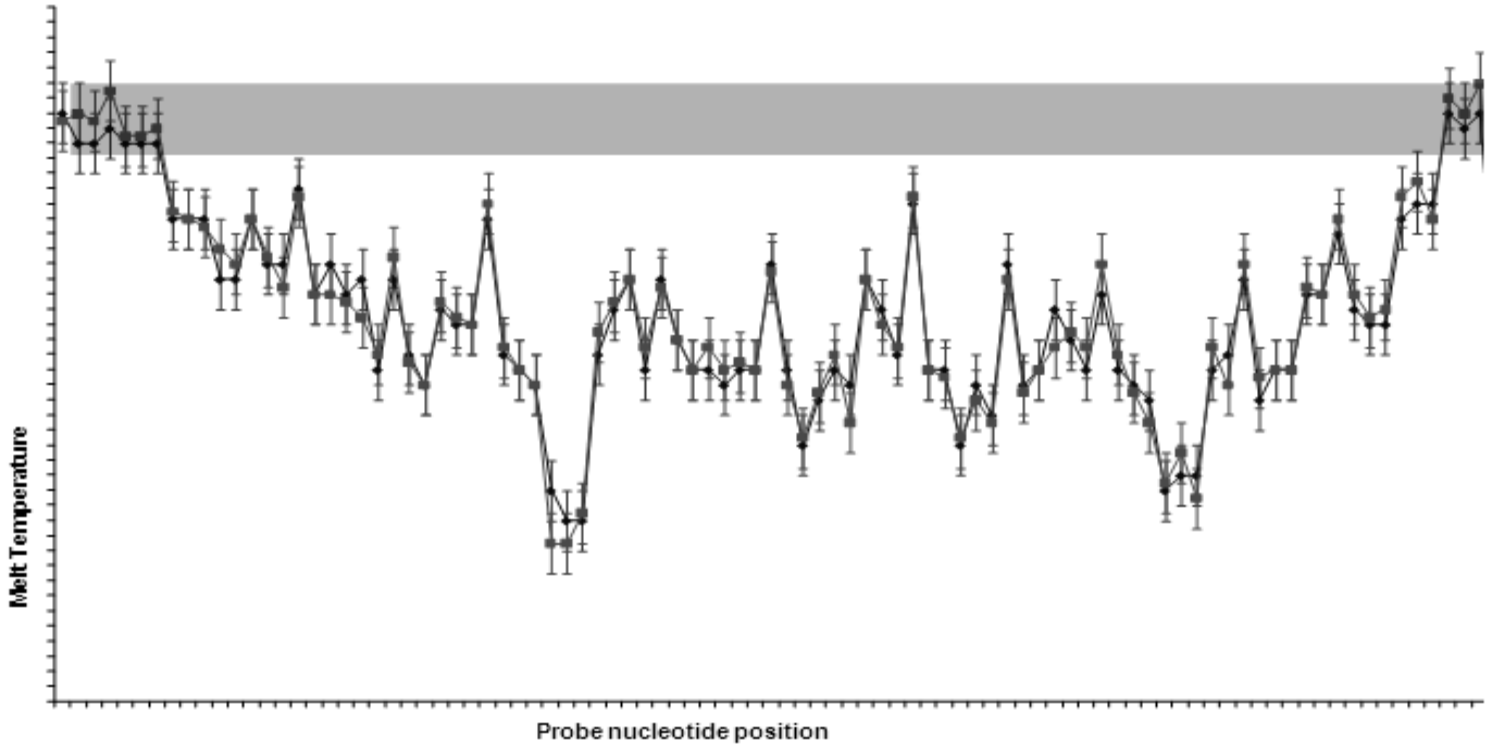


Figure 12: Commercial versus TdT Labeled Probe
Tested against the variable human sequence synthetic library (Error Bars = ± 0.4 Degrees). Reference sequence temperature range (grey bar) is highlighted.

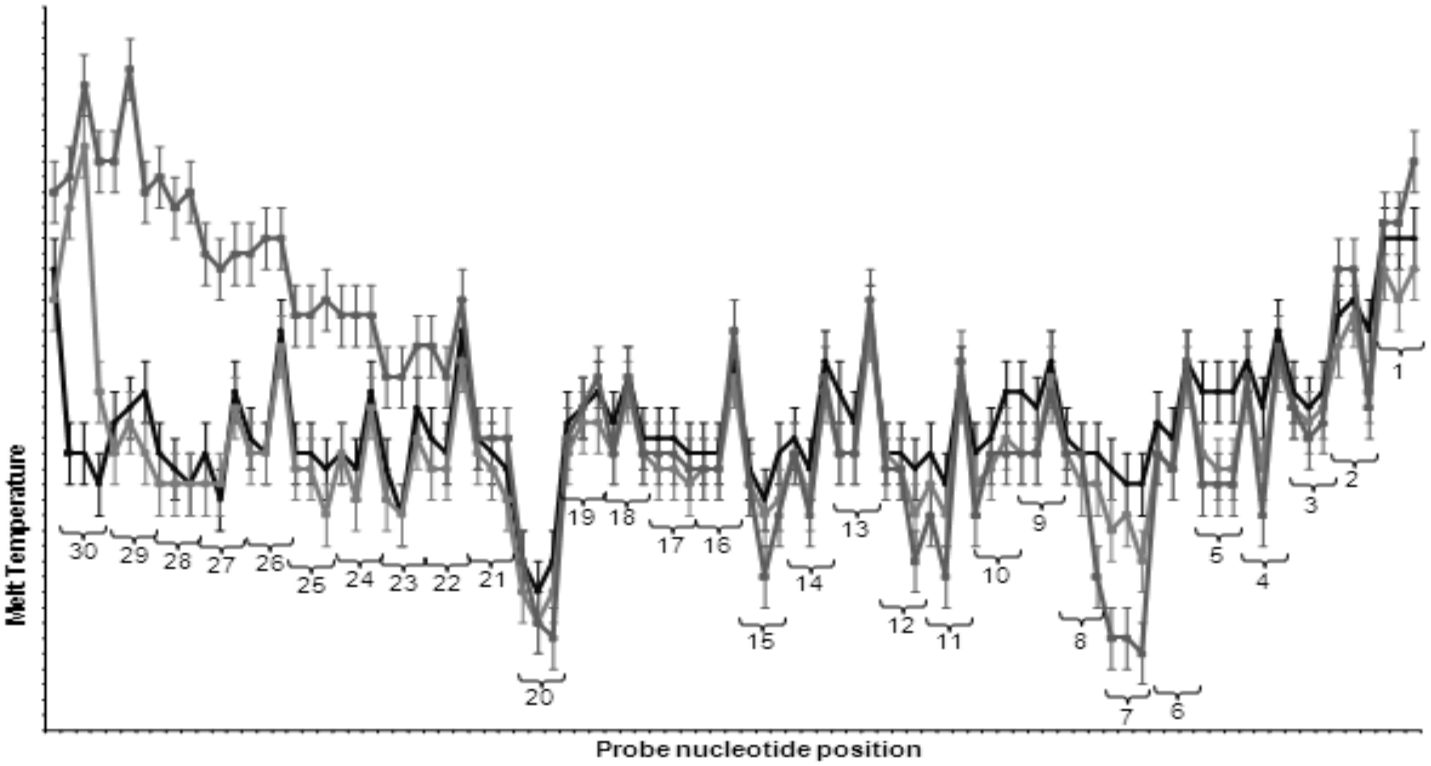
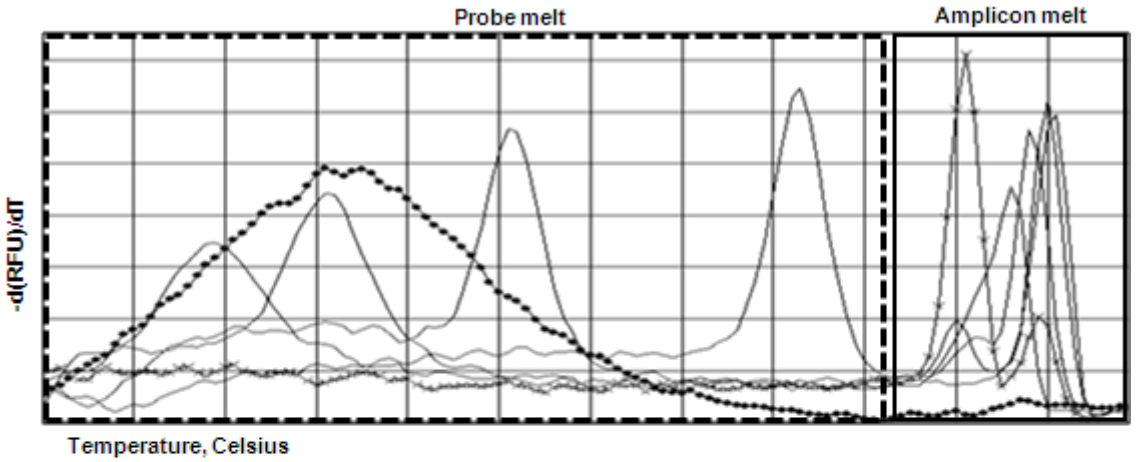
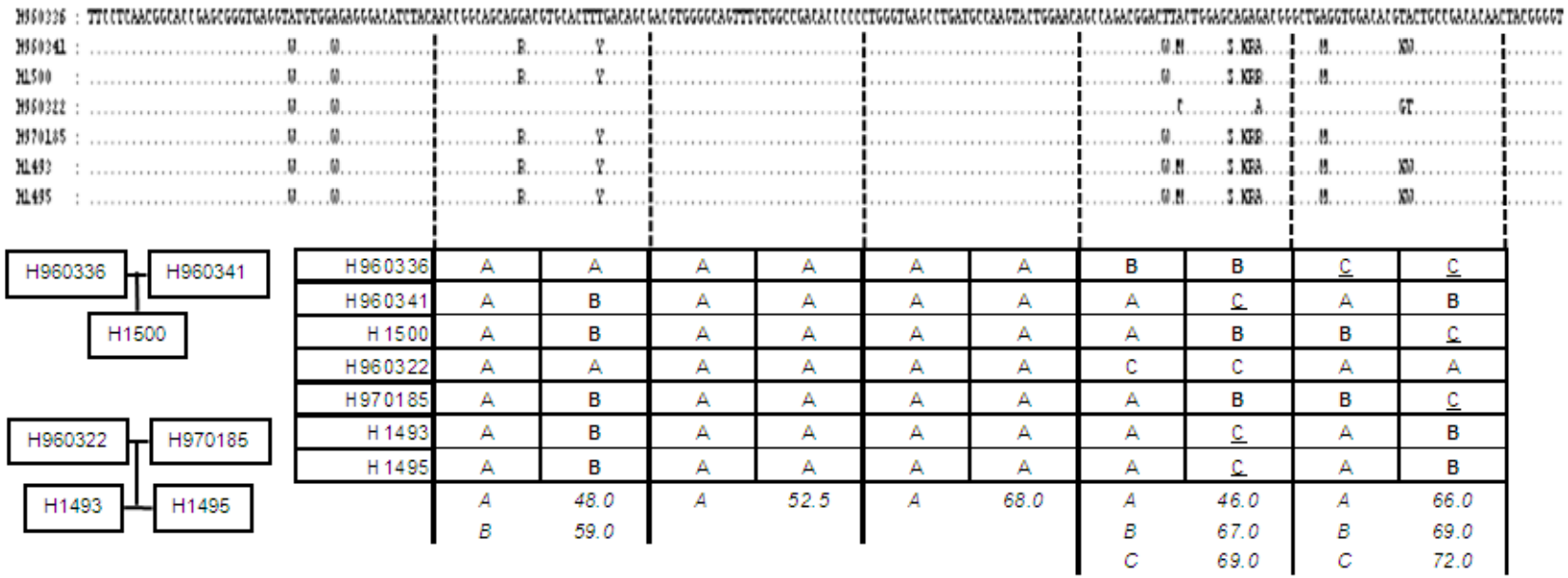


Figure 13: Inosine Probes Tested Against the Synthetic Human Sequence Library
Probe treatments consisted of unmodified, inosine at probe position 30, and inosines at probe positions 28, 29, and 30.



ID	Probe	Amplicon
Human	71.5	85.5
Ferret	55.5	83.0
Penguin	45.5	85.5
Flamingo	39.5	84.0
Python	-	81.0
Negative	47.0	-

Figure 14: dpFRET *Cytochrome B* SNP Real World Species Testing
Multiple species (smooth), NTC (diamond) and python (cross) are shown.



H960336	A	A	A	A	A	A	B	B	C	C
H960341	A	B	A	A	A	A	A	C	A	B
H1500	A	B	A	A	A	A	A	B	B	C
H960322	A	A	A	A	A	A	C	C	A	A
H970185	A	B	A	A	A	A	A	B	B	C
H1493	A	B	A	A	A	A	A	C	A	B
H1495	A	B	A	A	A	A	A	C	A	B
	A	*48.0*	*A*	*52.5*	*A*	*68.0*	*A*	*46.0*	*A*	*66.0*
	B	*59.0*					*B*	*67.0*	*B*	*69.0*
							C	*69.0*	*C*	*72.0*

Figure 15: dpFRET *Mhc DRB* SNP Penguin Paternity Testing

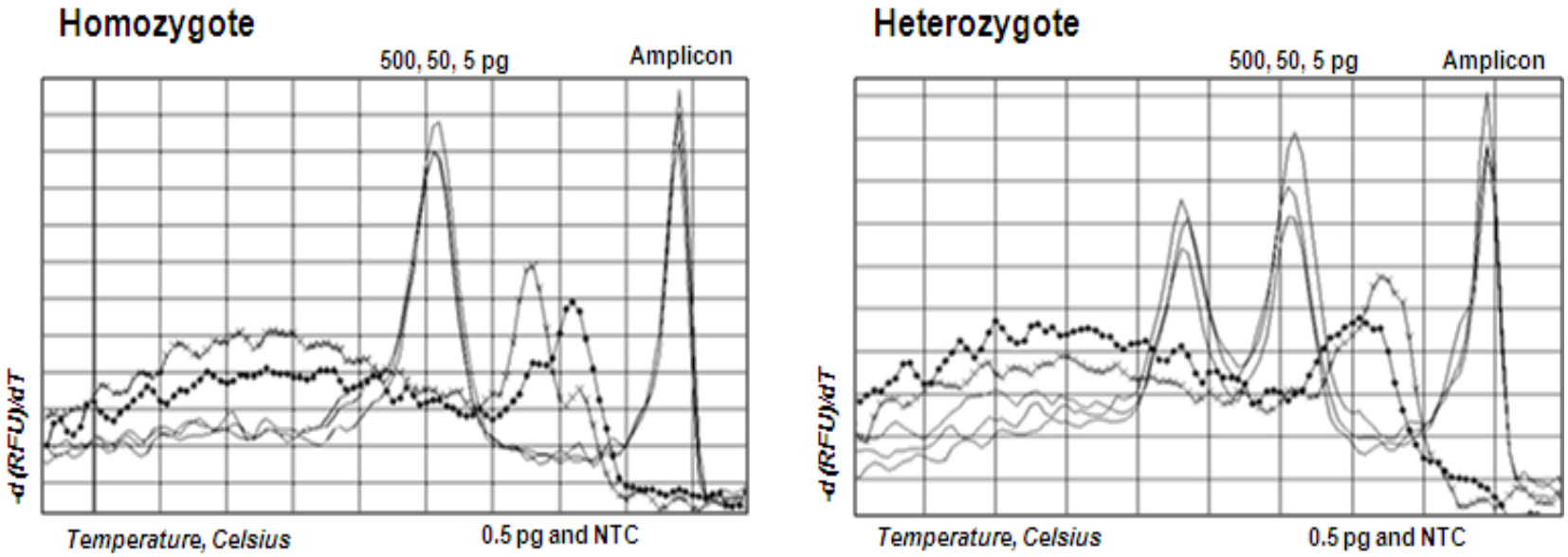


Figure 16: dpFRET SNP Sensitivity and Lack of Allelic Dropout

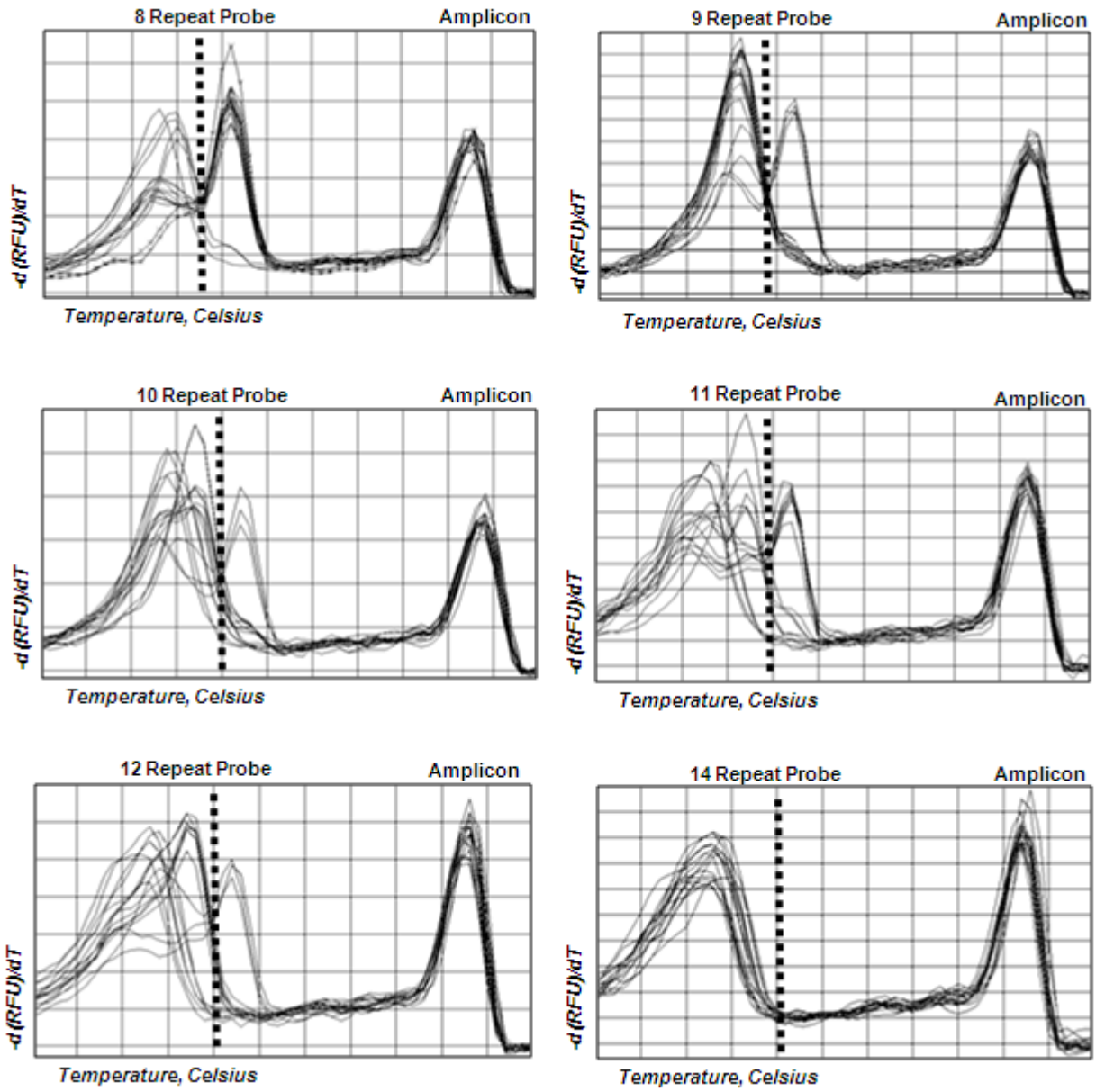


Figure 17: dpFRET Simple STR Testing (TPOX)

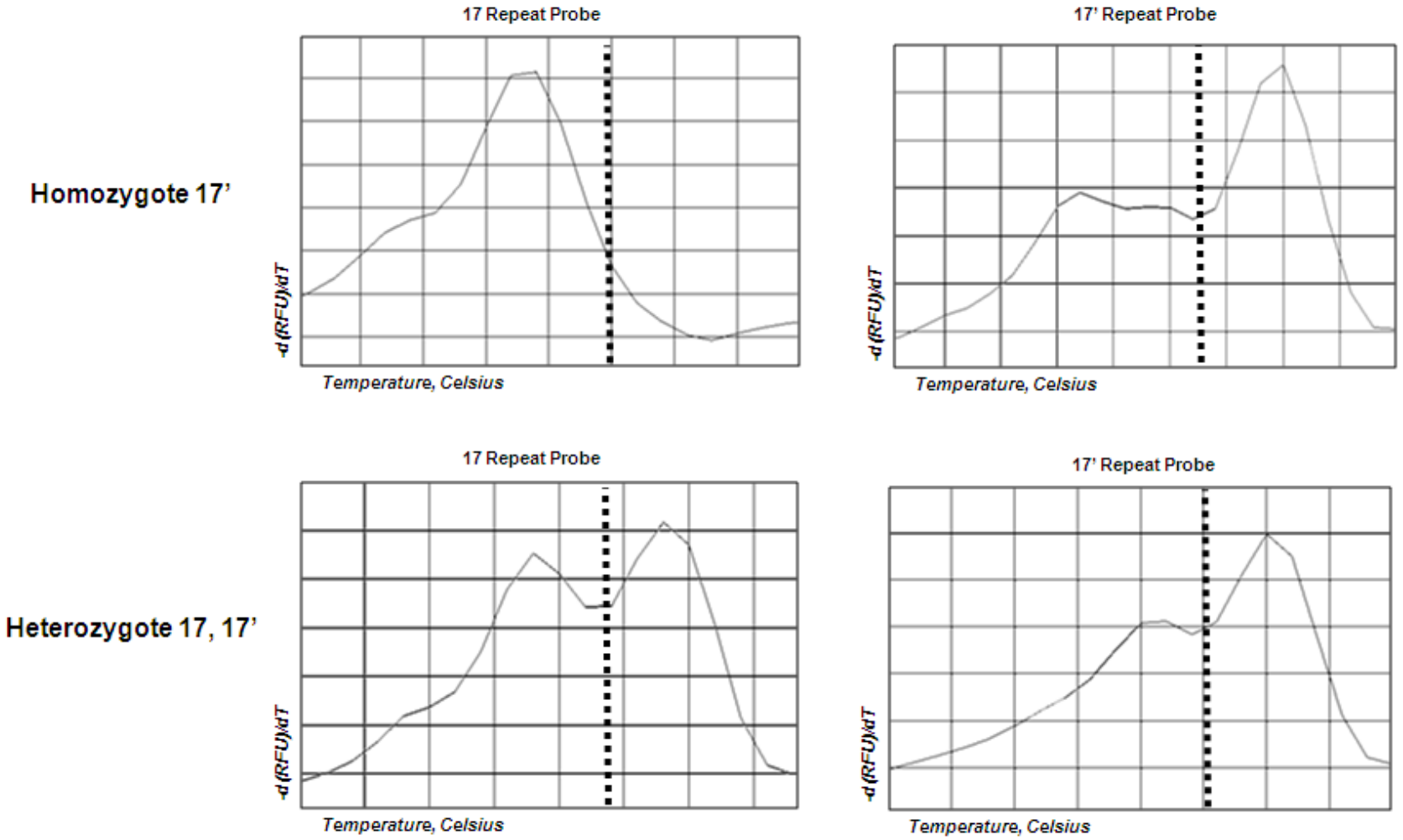


Figure 18: dpFRET Complex STR Testing (D3S1358)

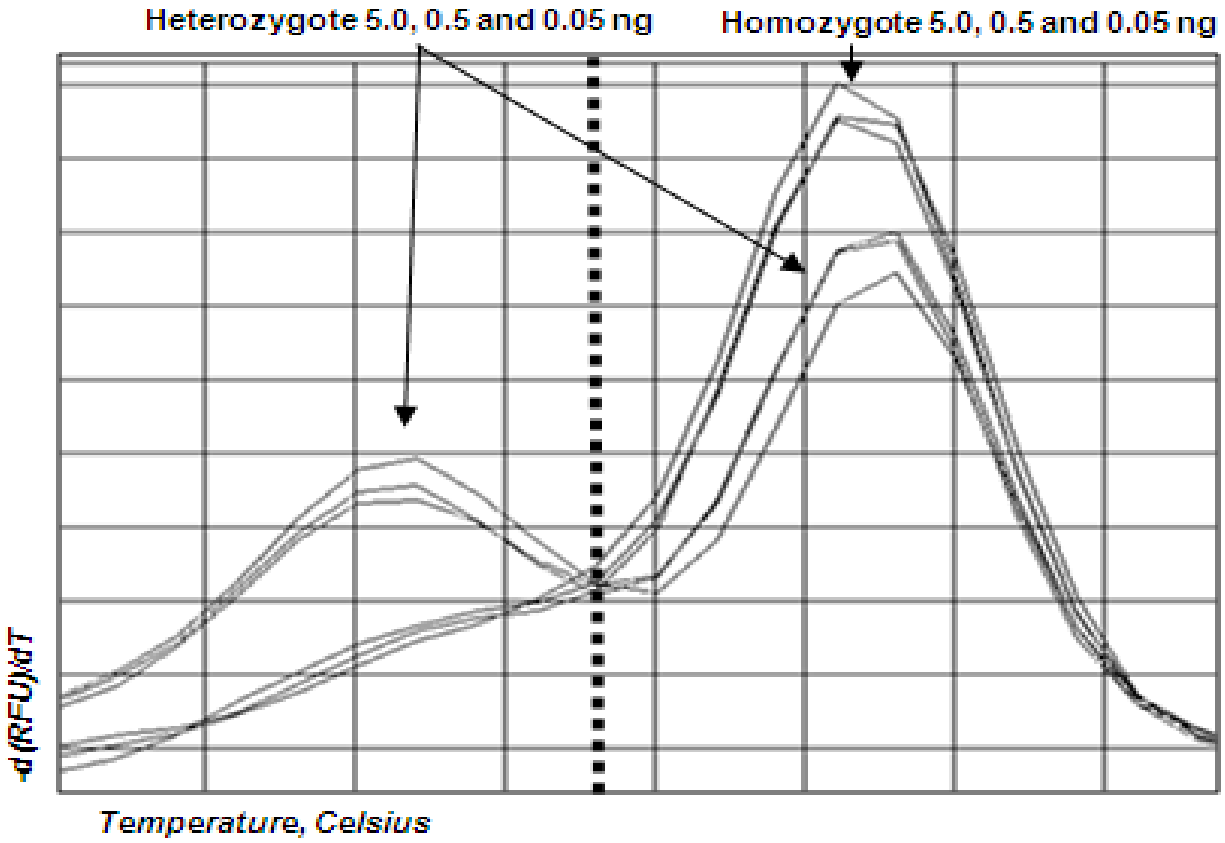


Figure 19: dpFRET STR Sensitivity and Lack of Allelic Dropout

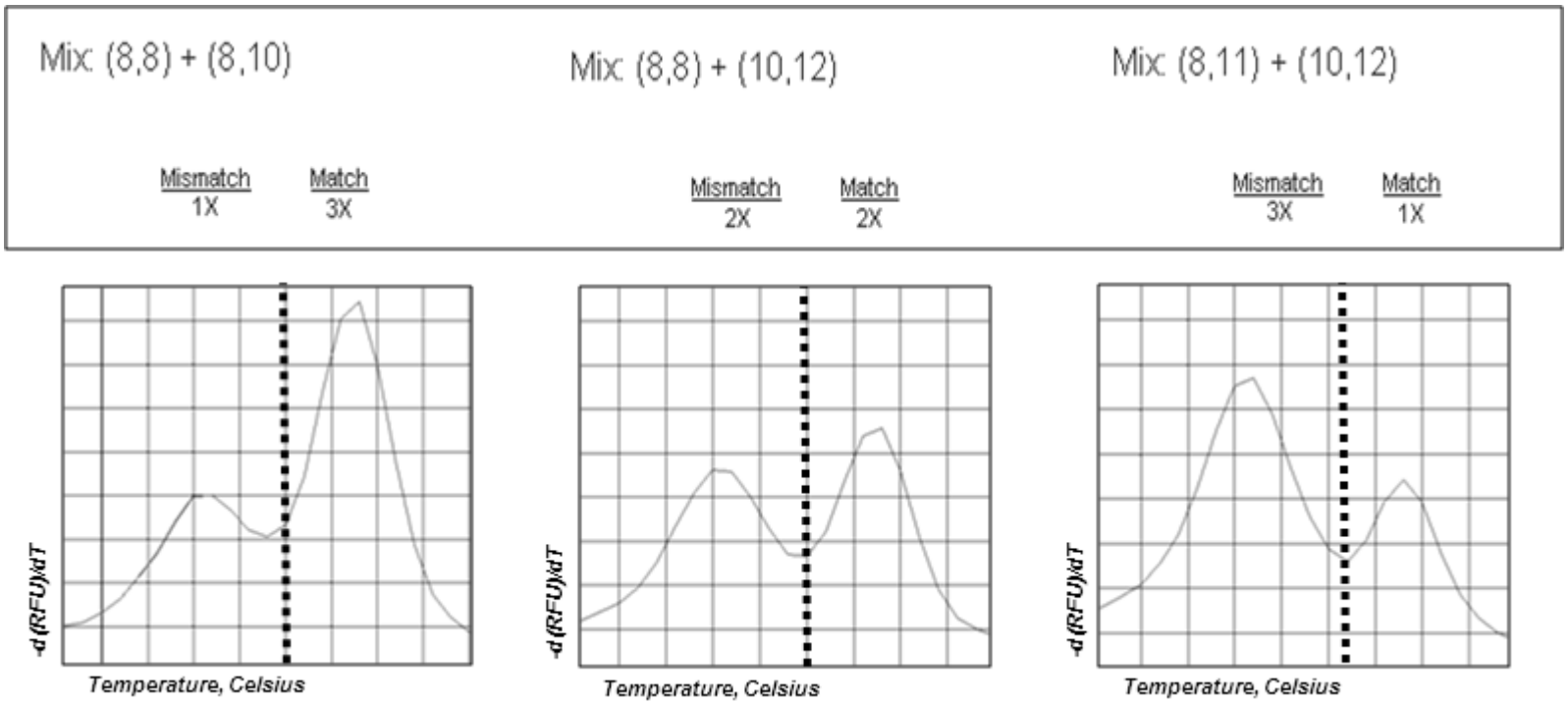
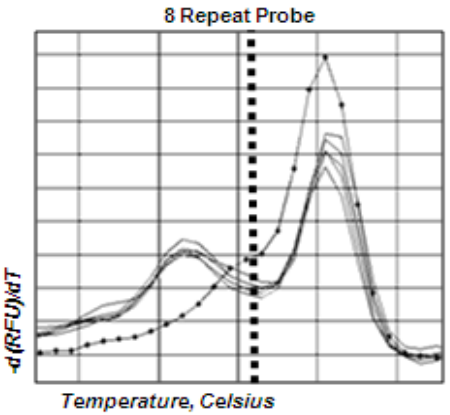
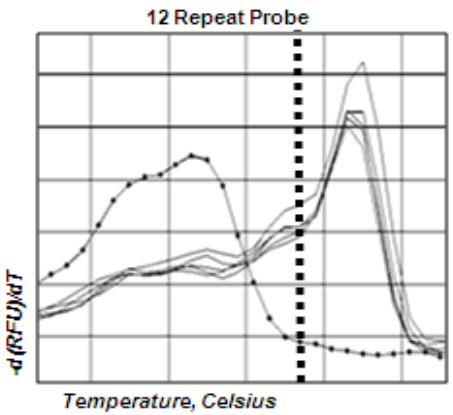


Figure 20: dpFRET STR Artificial Mixed Sample Results

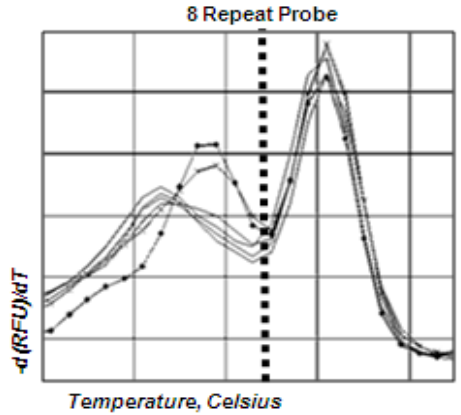
Case 1: donor (8,12) and recipient (8,8)

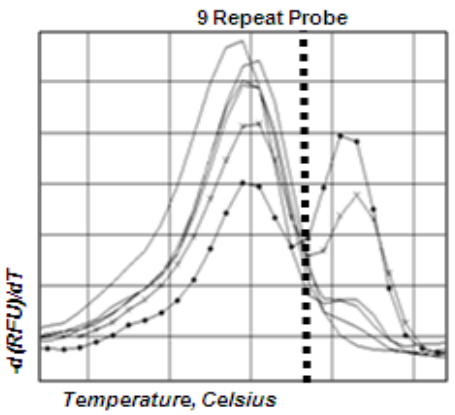


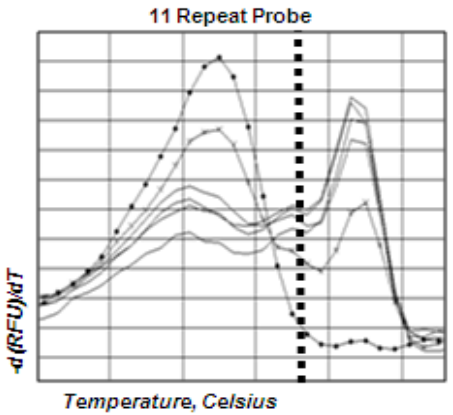


CE Results	
Case 1	**% donor**
Peripheral blood	90
Monocytes	95
Granulocytes	95
Bone Marrow	95
Case 2	**% donor**
Peripheral blood	80
Monocytes	90
Granulocytes	**50**
Bone Marrow	90

Case 2: donor (8,11) and recipient (8,9)







Donor and cell fractions – smooth
Recipient – diamonds
Case 2 granulocytes - crosses

Figure 21: dpFRET STR Real World Mixed Sample Results (Bone Marrow Transplant)

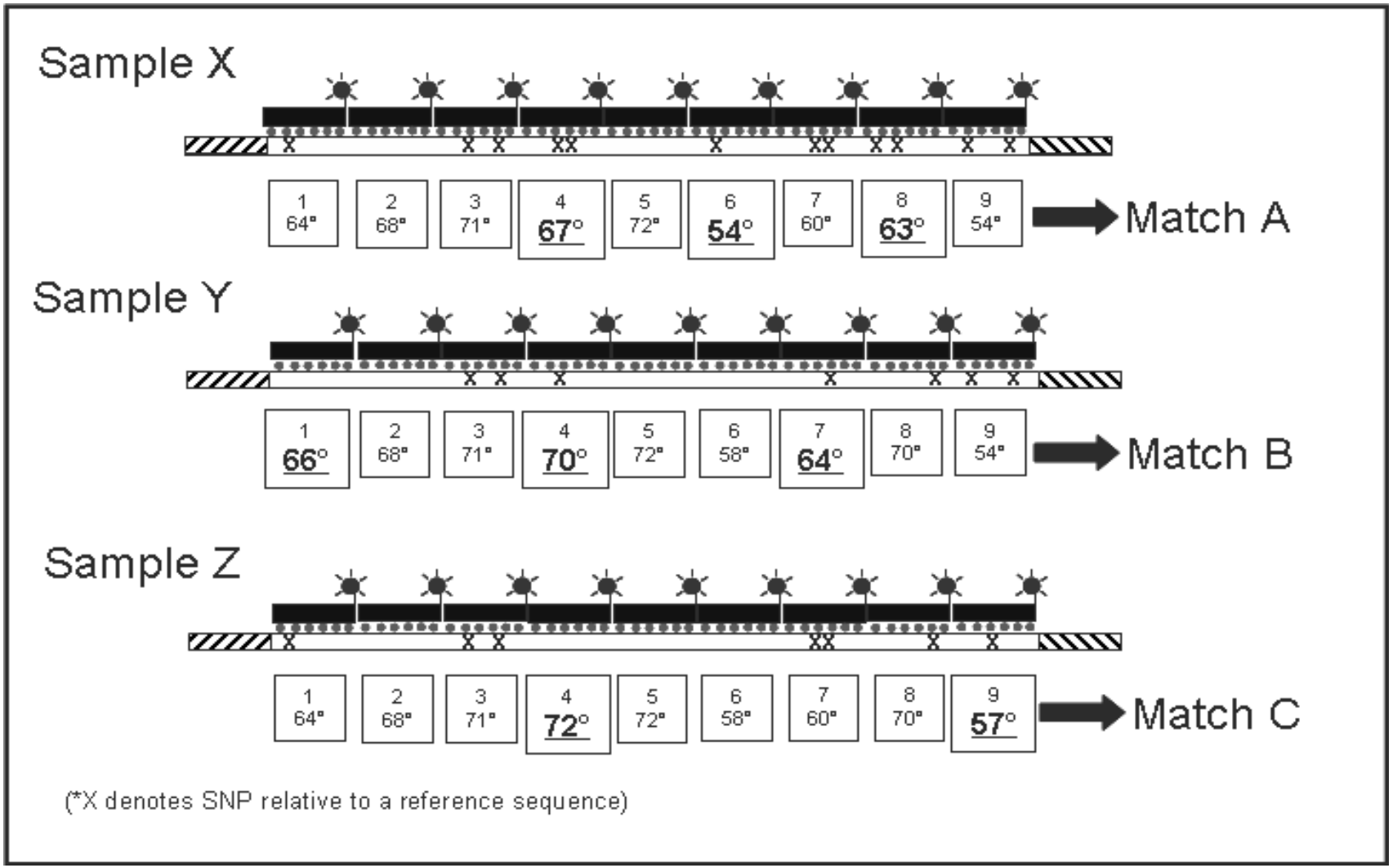


Figure 22: Experimental Approach for "Relative Sequencing"

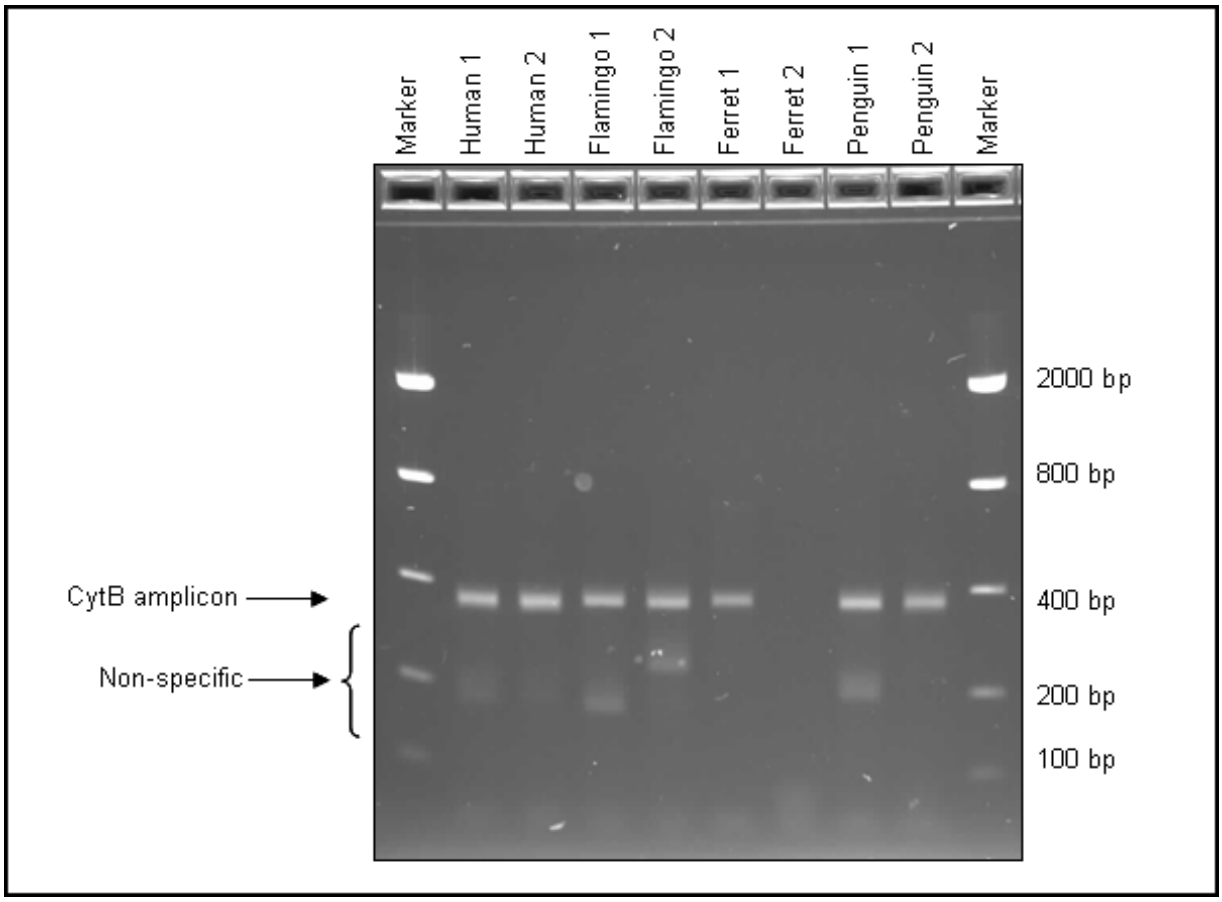


Figure 23: dpFRET 80 cycle *Cytochrome B* Endpoint Products
Specific (350 bp) and Non-specific Amplicons Detected by Gel Electrophoresis

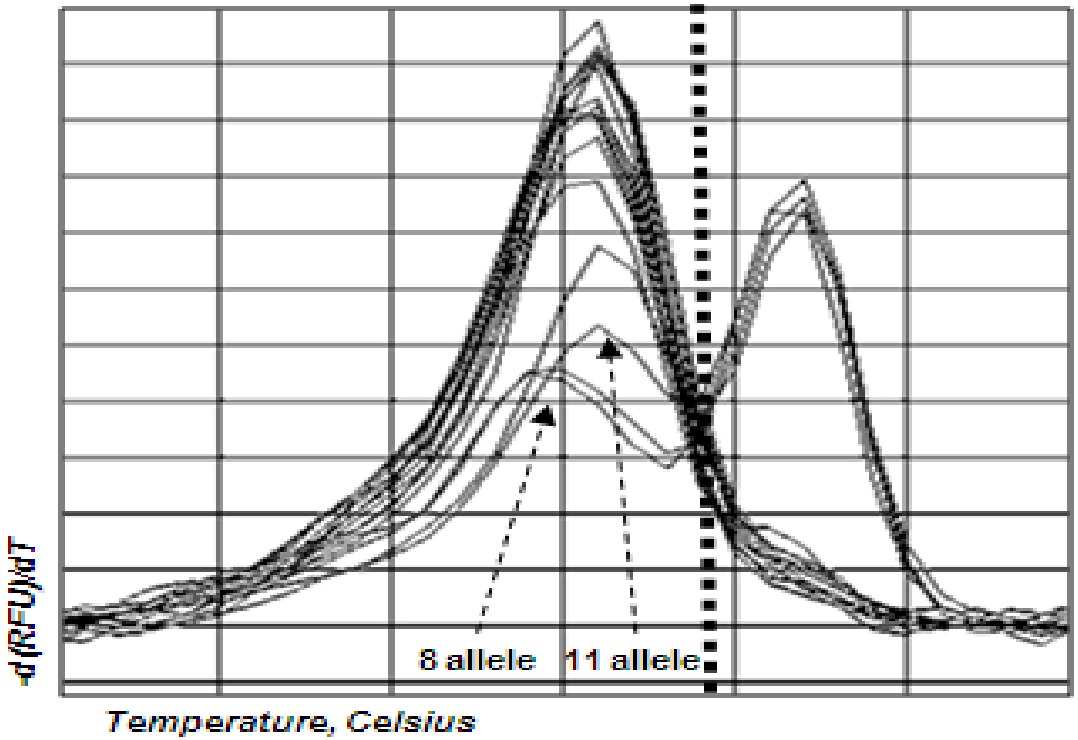


Figure 24: dpFRET STR 9 Repeat Probe Differential Mismatch Peak Profiles for 8 and 11 Alleles

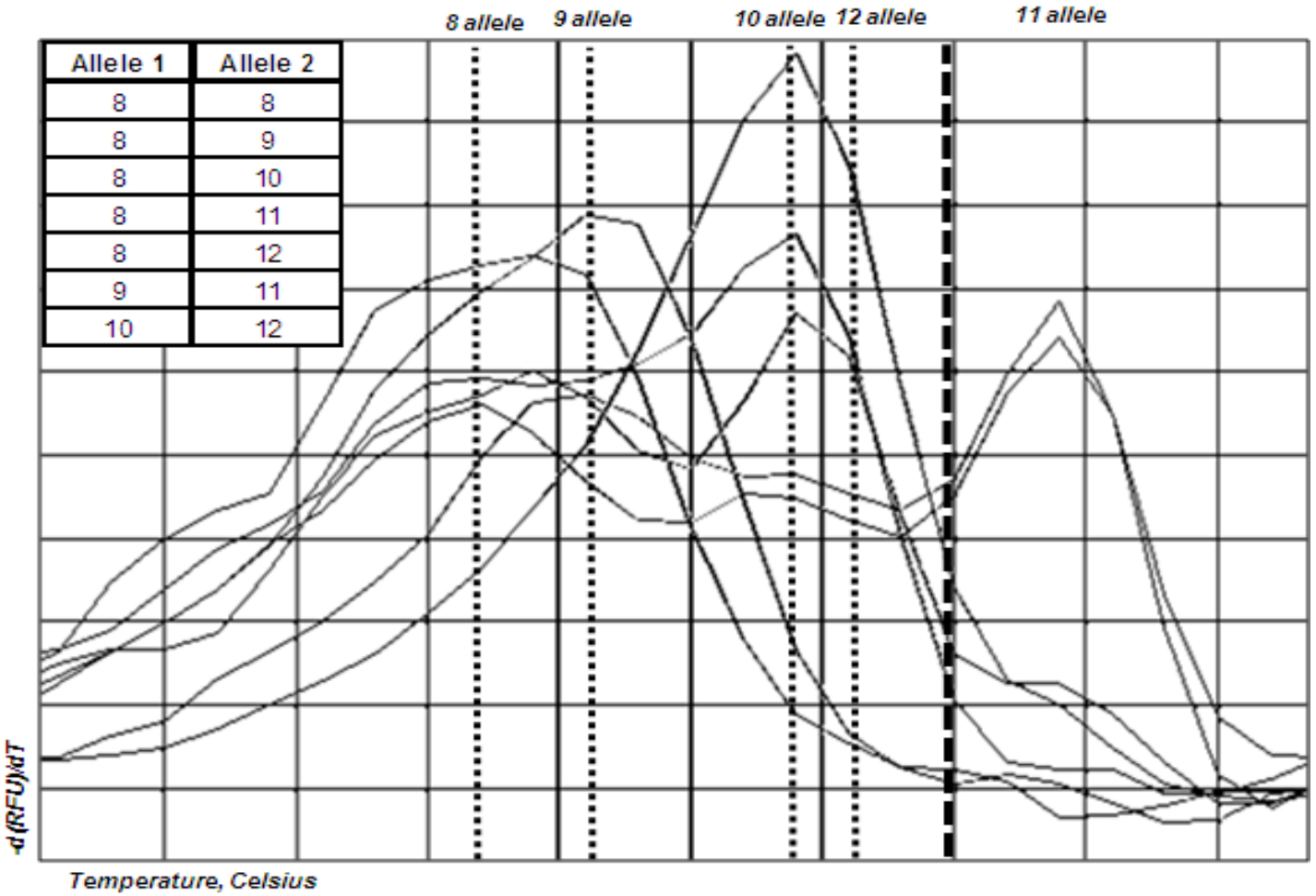


Figure 25: dpFRET STR 11 Repeat Probe Differential Mismatch Peak Profiles
Dotted lines indicate approximate melt temperatures generated with a single probe for different alleles.

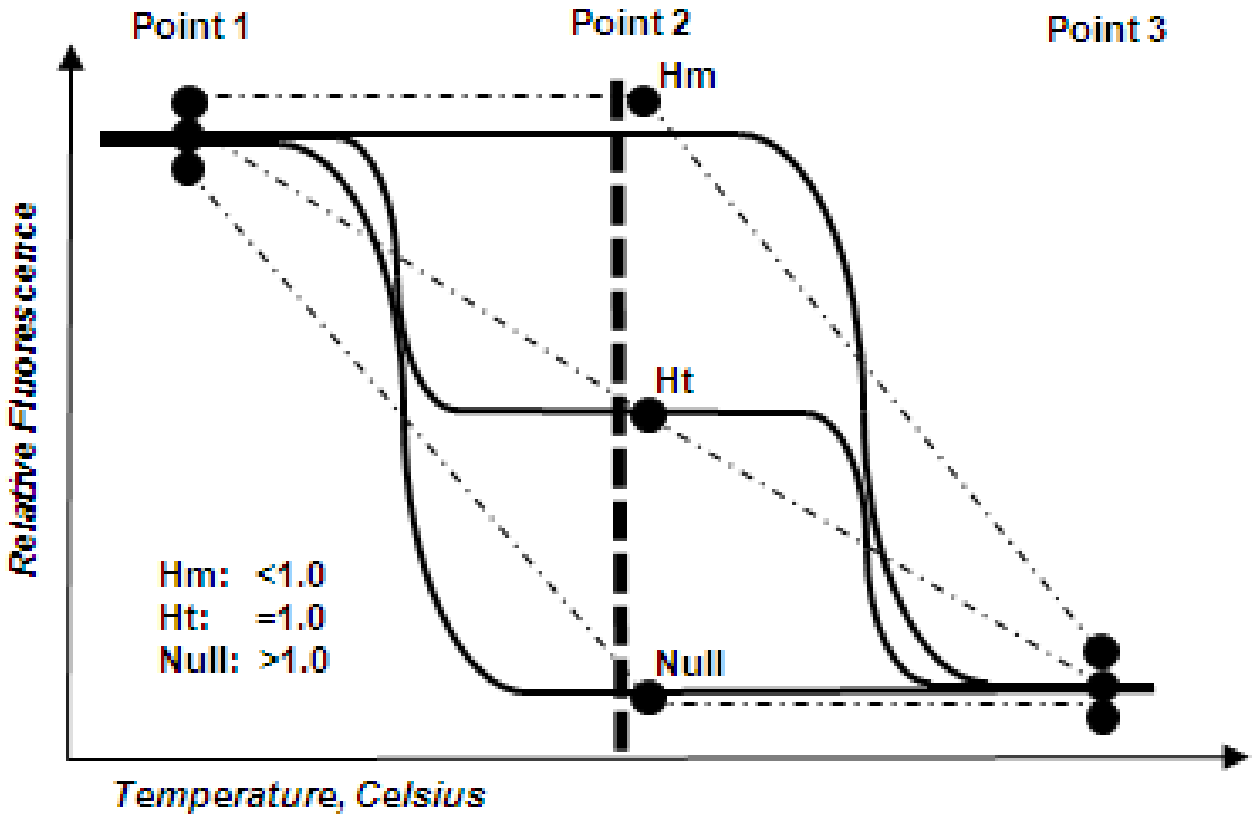


Figure 26: dpFRET STR Slope Ratio Analysis

APPENDIX B: TABLES

Table 1: Species Used for the Variable Animal Species Sequence Template Library

Human	Catfish	FinWhale	Heron	Minnow	Rhino
Aardvark	Cattle	FlyFox	Hippo	MnkSeal	RvrDolphin
AfElephant	Cheetah	Fox	Horse	Mongoose	Salamander
Alpaca	Chicken	Frog	HumWhale	Mouse	Salmon
Armadillo	Chimp	GdFurSeal	Hyrax	Muntjac	Sheep
AsBlkBear	Coelacanth	NtFurSeal	Junglefowl	NileCroc	Skate
AsElephant	Colobus	Goat	Kestrel	Orangutan	Sloth
AtWalrus	Coyote	Goby	Kiwi	Penguin	SptSeal
AuSeaLion	Deer	Gorilla	Langur	Pig	Squirrel
Baboon	Desman	GrayWolf	Lemur	PolarBear	Stingray
BalWhale	Dog	Grebe	Leopard	Porpoise	Sturgeon
Bat	Dogfish	GrnLizard	LfMonkey	Rabbit	TftDeer
BrnBear	Donkey	GrnMonkey	Loach	Rat	TwnVole
Buffalo	Dugong	GuinPig	Loon	Reindeer	Vole
CaspSeal	Eel	Hamster	LprdSeal	RghtWhale	WhtShark
Cat	Finch	Hedgehog	Mammoth	Rhea	Yak

Table 2: dpFRET *Cytochrome B* SNP Real World Species Testing Melt Values

ID	Amplicon	Probe
Human 1	85.0	71.5
Human 2	85.5	71.5
Penguin 1	85.5	46.0
Penguin 2	85.5	45.5
Flamingo 1	84.0	39.5
Flamingo 2	84.0	39.5
Python 1	80.5	-
Python 2	81.0	-
Ferret 2	83.0	55.5
Negative	-	47.0

APPENDIX C: VARIABLE ANIMAL SPECIES SEQUENCE SYNTHETIC TEMPLATE LIBRARY

Species	Sequence
Human	CAA CCG CCT TTT CAT CAA TCG CCC ACA TCA CTC GAG ACG TAA ATT ATG GC
Aardvark	CAA CCG CAT TCT CAT CTG TAA CCC ATA TTT GCC GAG ATG TAA ACT ACG GC
AfElephant	TAA CTG CAT TTT CAT CTA TAT CCC ATA TTT GCC GAG ATG TGA ACT ACG GC
Alpaca	CAA CAG CCT TCT CTT CAG TCG CAC ACA TCT GCC GAG ACG TAA ATT ACG GC
Armadillo	TAA CAG CCT TCT CAT CTG TAA CTC ACA TCT GCC GAG ACG TAA ACT ATG GC
AsBlkBear	CTA CAG CCT TTT CAT CAG TCG CCC ATA TTT GCC GAG ACG TCC ATT ACG GA
AsElephant	TAA CTG CAT TTT CAT CTA TAT CCC ATA TCT GCC GAG ACG TCA ACT ACG GC
AtWalrus	CCA CAG CTT TCT CAT CAA TCA CAC ATA TCT GCC GAG ATG TCA ACT ATG GT
AuSeaLion	CCA CAG CCT TTT CAT CGG TCA CCC ACA TTT GCC GAG ACG TGA ACT ACG GC
Baboon	CCT CTG CCT TCT CTT CAA TCG CAC ACA TCA CCC GAG ACG TAA ACT ATG GC
BalWhale	CAA CCG CTT TCT CAT CAG TCA CAC ACA TTT GCC GAG ACG TAA ACT ACG GC
Bat	CTA CCG CAT TCA ACT CTG TCA CCC ATA TCT GTC GAG ACG TCA ACT ATG GA
BrnBear	CCA CAG CTT TTT CAT CAG TCA CCC ACA TTT GCC GAG ACG TTC ACT ACG GA
Buffalo	CAA CAG CAT TCT CCT CCG TCG CCC ACA TCT GCC GGG ACG TGA ACT ATG GA
CaspSeal	CCA CAG CCT TCT CAT CAG TAA CCC ACA TCT GCC GGG ACG TAA ACT ACG GC
Cat	TAA CCG CCT TTT CAT CAG TTA CCC ACA TCT GTC GCG ACG TTA ATT ATG GC
Catfish	CAA CTG CCT TTT CAT CCG TCG CCC ACA TCT GCC GAG ATG TAA ACT ACG GG
Cattle	CAA CAG CAT TCT CCT CTG TTA CCC ATA TCT GCC GAG ACG TGA ACT ACG GC
Cheetah	TAA CCG CCT TTT CAT CAG TTA CTC ACA TCT GCC GCG ACG TCA ACT ACG GC
Chicken	CCC TAG CCT TCT CCT CCG TAG CCC ACA CTT GCC GGA ACG TAC AAT ACG GC
Chimp	CAA CCG CCT TCT CAT CGA TCG CCC ACA TTA CCC GAG ACG TAA ACT ATG GT
Coelacanth	CAA CAG CAT TCT CAT CAG TAG CCC ACA TCT GCC GAG ATG TAA ACT ATG GA
Colobus	CCT CTG CTT TCT CCT CAG TTG CAC ATA TCA CCC GGG ACG TAA ACT ATG GC
Coyote	CCA CAG CTT TTT CAT CAG TCA CCC ACA TCT GTC GAG ACG TTA ACT ACG GC
Deer	TAA CAG CAT TCT CCT CTG TCA CCC ATA TCT GTC GAG ATG TCA ATT ATG GT
Desman	TAA CAG CCT TCT CAT CAG TAA CCC ATA TTT GCC GAG ATG TAA ACT ACG GA
Dog	CCA CAG CTT TTT CAT CAG TCA CCC ACA TCT GCC GAG ACG TTA ACT ACG GC
Dogfish	CCA CGG CCT TCT CCT CAG TAG TTC ATA TTT GTC GTG ACG TCA ATT ATG GT
Donkey	CAA CTG CCT TCT CAT CCG TCA CCC ATA TCT GCC GAG ACG TTA ACT ACG GA
Dugong	TAA CCG CAT TCT CCT CAG TAA CCC ATA TTT GCC GGG ATG TAA ACT ACG GC
Eel	CGA CCG CTT TCT CCT CAG TTG TCC ATA TCT GCC GAG ATG TAA ACT ATG GC
Finch	CCC TAG CCT TCT CCT CAG TCG CCC ACA TAT GCC GAG ACG TAC AAT TTG GC
FinWhale	CAA CCG CCT TCT CAT CAG TCA CAC ACA TCT GCC GAG ACG TGA ATT ACG GC
FlyFox	CAA CCG CCT TCC AAT CCG TAA CCC ACA TCT GCC GAG ACG TAA ACT ACG GC
Fox	CTA CTG CTT TCT CAT CTG TCA CTC ACA TCT GCC GAG ACG TTA ACT ATG GC
Frog	CCC TTG CAT TCT CAT CTA TTG CCC ACA TCT GTC GAG ATG TTA ATA ACG GC
GdFurSeal	CTA CAG CCT TTT CAT CAG TCA CCC ACA TTT GCC GAG ACG TGA ACT ACG GC
NtFurSeal	CCA CAG CCT TCT CAT CAG TCG CCC ATA TTT GCC GAG ACG TGA ACT ACG GC
Goat	TAA CAG CAT TTT CCT CTG TAA CTC ACA TTT GTC GAG ATG TAA ATT ATG GC
Goby	CCA CAG CTT TTT CTT CTG TAG CCC ATA TCT GCC GGG ATG TTA ACT TTG GT
Gorilla	CAA CCG CCT TCT CAT CAA TTG CCC ACA TCA CCC GAG ATG TAA ACT ATG GC
GrayWolf	CCA CAG CTT TTT CAT CAG TCA CCC ACA TCT GCC GAG ACG TTA ACT ACG GC
Grebe	CCC TAG CCT TCT CAT CCG TCG CCC ACA CAT GTC GAA ACG TAC AGT ACG GC
GrnLizard	CCT CCG CAT TCT CAT CTG TCA CCC ACA TTC ACC GAG ATG TTC AAT ATG GC
GrnMonkey	CTT CTG CCT TCT CTT CAA TCG CAC ACA TCA CCC GAG ACG TAA ACC ACG GC
GuinPig	CCA CGG CAT TCT CGT CTG TCG CCC ACA TTT GCC GAG ACG TAA ACT ATG GC
Hamster	CTA CAG CAT TCT CAT CAG TCA CCC ACA TTT GTC GAG ATG TTA ATT ACG GC
Hedgehog	TTA CAG CAT TTT CAT CCA TTA CTC ACA TTT GCC GAG ATG TAA ACT ACG GT

Heron	CAT TAG CCT TCT CAT CCG TCG CCC ACA CAT GCC GAA ACG TAC AGT ACG GC
Hippo	TCA CCG CAT TCT CAT CGG TAA CCC ACA TCT GCC GTG ATG TAA ACT ACG GG
Horse	CAA CTG CCT TCT CAT CCG TCA CTC ACA TCT GCC GAG ACG TTA ACT ACG GA
HumWhale	CAA CCG CCT TCT CAT CAG TCA CAC ACA TCT GTC GAG ACG TAA ATT ATG GC
Hyrax	TAA CCG CAT TCA CAT CAG TAA CCC ACA TTT GTC GAG ACG TAA ACC ATG GA
Junglefowl	CCC TAG CCT TCT CCT CCG TAG CCC ACA CTT GCC GGA ACG TAC AAT ACG GC
Kestrel	CAC TGG CCT TCT CAT CTG TTG CCC ACA CAT GCC GAA ACG TGC AGT ACG GA
Kiwi	CCC TAG CCT TTT CAT CCA TCG CCC ATA TCT GTC GAA ACG TCC AAT ATG GA
Langur	CCT CAG CCT TCT CCT CAA TCG CCC ATA TCA CTC GAG ACG TAA ACT ACG GC
Lemur	CAA CAG CAT TTT CAT CCA TTG CCC ACA TCT CAC GAG ACG TAA ACT ACG GC
Leopard	TAA CTG CTT TCT CAT CTG TCA CCC ATA TTT GCC GCG ACG TAA ACT ATG GT
LfMonkey	CCT CTG CCT TCT CCT CAA TTG CAC ATA TTA CCC GAG ATG TAA ATT ATG GC
Loach	CTA CTG CCT TTT CAT CCG TAG CCC ACA TCT GCC GAG ATG TTA ACT ATG GA
Loon	CCC TAG CCT TCT CAT CCG TTG CCC ACA CAT GCC GAA ACG TAC AGT ACG GT
LprdSeal	CTA CAG CCT TTT CAT CAG TCA CAC ACA TCT GCC GAG ACG TAA ACT ACG GT
Mammoth	TAA CTG CAT TTT CAT CTA TAT CCC ATA TCT GCC GAG ATG TCA ACT ACG GT
Minnow	CCA CTG CAT TTT CAT CAG TAG CCC ACA TCT GCC GAG ATG TTA ATT ATG GC
MnkSeal	CCA CAG CCT TTT CAT CAA TCA CAC ACA TCT GCC GAG ACG TAA ATT ACG GC
Mongoose	CAA CTG CCT TTT CAT CAG TAA CCC ACA TTT GCC GCG ACG TCA ACT ACG GC
Mouse	TAA CAG CCT TTT CAT CAG TAA CAC ACA TTT GTC GAG ACG TAA ATT ACG GG
Muntjac	TAA CAG CAT TCT CCT CGG TTA CCC ATA TCT GCC GAG ACG TCA ACT ATG GC
NileCroc	CCC TAG CTT TTA TAT CTG TCG CTT ATA CTT CAC GAG AAG TTT GAT ACG GC
Orangutan	CCA CTG CCT TTT CAT CAA TCG CCC ACA TCA CTC GAG ATG TAA ACT ACG GC
Penguin	CCC TAG CCT TCT CCT CCA TCG CCC ACA CAT GCC GAA ATG TAC AGT ACG GC
Pig	CAA CAG CTT TCT CAT CAG TTA CAC ACA TTT GTC GAG ACG TAA ATT ACG GA
PolarBear	CCA CAG CTT TTT CAT CAG TCA CCC ACA TTT GCC GAG ACG TTC ACT ACG GG
Porpoise	CAA CCG CTT TTT CAT CAG TCG CAC ATA TCT GTC GAG ACG TTA ATT ATG GC
Rabbit	CAA CAG CAT TCT CAT CAG TAA CCC ATA TTT GCC GAG ATG TTA ACT ATG GC
Rat	TAA CAG CAT TTT CAT CAG TCA CCC ACA TCT GCC GAG ACG TAA ACT ACG GC
Reindeer	TAA CAG CAT TCT CCT CTG TTA CTC ACA TCT GTC GAG ACG TCA ATT ATG GC
RghtWhale	CAA CCG CCT TCT CAT CAA TCA CAC ACA TCT GTC GAG ACG TAA ACT ACG GT
Rhea	CAT TAG CCT TCT CAT CCG TAG CCC ACA CCT GCC GCA ACG TCC AAT ATG GT
Rhino	TAA CTG CCT TCT CAT CTG TCG CCC ATA TCT GTC GAG ACG TGA ATT ACG GC
RvrDolphin	CAA CCG CCT TCT CAT CCA TCA CAC ACA TTT GCC GAG ACG TCA ACT ACG GC
Salamander	CTT CCG CAT TTT CAT CAG TCG TAC ATA TCT GCC GAG ACG TAA ACT ATG GA
Salmon	CAA CAG CTT TTT CCT CTG TCT GCC ACA TCT GCC GAG ATG TTA GTT ACG GC
Sheep	CAA CAG CAT TCT CCT CTG TAA CCC ACA TTT GCC GAG ACG TAA ACT ATG GC
Skate	CCT CCG CTT TCT CCT CAG TTG TTC ACA TCT GCC GAG ATG TGA ATT ATG GA
Sloth	CCA CCG CCT TCT CAT CCG TAA CCC ACA TCT GCC GAG ACG TAA ACT ACG GC
SptSeal	CCA CAG CCT TCT CAT CAG TAA CCC ACA TCT GCC GAG ACG TAA ACT ACG GC
Squirrel	TAA CAG CTT TTT CTT CCG TTA CTC ACA TCT GCC GAG ACG TAA ATT ATG GC
Stingray	CAA CCG CAT TCT CCT CAG TAG CAC ATA TCT GCC GAG ACG TAA ACT ACG GC
Sturgeon	CAA CAG CCT TCT CTT CTG TCG CCC ACA TCT GCC GAG ATG TAA ATT ACG GA
TftDeer	TAA CAG CAT TTT CCT CTG TAA CCC ACA TTT GCC GAG ACG TCA ACT ATG GG
TwnVole	CAA CAG CAT TCT CAT CAG TAG CCC ATA TCT GCC GAG ACG TCA ACT ACG GC
Vole	CAA CAG CAT TCT CAT CAG TAG CCC ACA TTT GTC GAG ACG TAA ACT ATG GC
WhtShark	CTA TAG CCT TCT CCT CAG TAA CCC ACA TCT GCC GTG ACG TCA ATT ACG GC
Yak	CAA CAG CAT TCT CCT CCG TTG CCC ATA TCT GCC GAG ACG TGA ACT ACG GC

APPENDIX D: VARIABLE HUMAN SEQUENCE SYNTHETIC TEMPLATE LIBRARY

Label	Sequence
CytB_Reference	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_1A	CAACCGCCTTATCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_1C	CAACCGCCTTCTCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_1G	CAACCGCCTTGTCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_2A	CAACCGCCTTTACATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_2C	CAACCGCCTTTCCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_2G	CAACCGCCTTTGCATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_3T	CAACCGCCTTTTTATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_3A	CAACCGCCTTTTAAATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_3G	CAACCGCCTTTTGATCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_4T	CAACCGCCTTTTCTTCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_4C	CAACCGCCTTTTCCTCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_4G	CAACCGCCTTTTCGTCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_5A	CAACCGCCTTTTCAACAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_5C	CAACCGCCTTTTCACCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_5G	CAACCGCCTTTTCAGCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_6T	CAACCGCCTTTTCATTAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_6A	CAACCGCCTTTTCATAAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_6G	CAACCGCCTTTTCATGAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_7T	CAACCGCCTTTTCATCTATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_7C	CAACCGCCTTTTCATCCATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_7G	CAACCGCCTTTTCATCGATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_8T	CAACCGCCTTTTCATCATTCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_8C	CAACCGCCTTTTCATCACTCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_8G	CAACCGCCTTTTCATCAGTCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_9A	CAACCGCCTTTTCATCAAACGCCCACATCACTCGAGACGTAAATTATGGC
CytB_9C	CAACCGCCTTTTCATCAACCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_9G	CAACCGCCTTTTCATCAAGCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_10T	CAACCGCCTTTTCATCAATTGCCCACATCACTCGAGACGTAAATTATGGC
CytB_10A	CAACCGCCTTTTCATCAATAGCCCACATCACTCGAGACGTAAATTATGGC
CytB_10G	CAACCGCCTTTTCATCAATGGCCCACATCACTCGAGACGTAAATTATGGC
CytB_11C	CAACCGCCTTTTCATCAATCCCCCACATCACTCGAGACGTAAATTATGGC
CytB_11T	CAACCGCCTTTTCATCAATCTCCCACATCACTCGAGACGTAAATTATGGC
CytB_11A	CAACCGCCTTTTCATCAATCACCCACATCACTCGAGACGTAAATTATGGC
CytB_12T	CAACCGCCTTTTCATCAATCGTCCACATCACTCGAGACGTAAATTATGGC
CytB_12A	CAACCGCCTTTTCATCAATCGACCACATCACTCGAGACGTAAATTATGGC
CytB_12G	CAACCGCCTTTTCATCAATCGGCCACATCACTCGAGACGTAAATTATGGC
CytB_13T	CAACCGCCTTTTCATCAATCGCTCACATCACTCGAGACGTAAATTATGGC
CytB_13A	CAACCGCCTTTTCATCAATCGCACACATCACTCGAGACGTAAATTATGGC
CytB_13G	CAACCGCCTTTTCATCAATCGCGCACATCACTCGAGACGTAAATTATGGC
CytB_14T	CAACCGCCTTTTCATCAATCGCCTACATCACTCGAGACGTAAATTATGGC
CytB_14A	CAACCGCCTTTTCATCAATCGCCAACATCACTCGAGACGTAAATTATGGC
CytB_14G	CAACCGCCTTTTCATCAATCGCCGACATCACTCGAGACGTAAATTATGGC
CytB_15T	CAACCGCCTTTTCATCAATCGCCCTCATCACTCGAGACGTAAATTATGGC
CytB_15C	CAACCGCCTTTTCATCAATCGCCCCCATCACTCGAGACGTAAATTATGGC
CytB_15G	CAACCGCCTTTTCATCAATCGCCCGCATCACTCGAGACGTAAATTATGGC
CytB_16T	CAACCGCCTTTTCATCAATCGCCCATATCACTCGAGACGTAAATTATGGC
CytB_16A	CAACCGCCTTTTCATCAATCGCCCAAATCACTCGAGACGTAAATTATGGC

CytB_16G	CAACCGCCTTTTCATCAATCGCCCAGATCACTCGAGACGTAAATTATGGC
CytB_17T	CAACCGCCTTTTCATCAATCGCCCACTTCACTCGAGACGTAAATTATGGC
CytB_17C	CAACCGCCTTTTCATCAATCGCCCACCTCACTCGAGACGTAAATTATGGC
CytB_17G	CAACCGCCTTTTCATCAATCGCCCACGTCACTCGAGACGTAAATTATGGC
CytB_18A	CAACCGCCTTTTCATCAATCGCCCACAACACTCGAGACGTAAATTATGGC
CytB_18C	CAACCGCCTTTTCATCAATCGCCCACACCACTCGAGACGTAAATTATGGC
CytB_18G	CAACCGCCTTTTCATCAATCGCCCACAGCACTCGAGACGTAAATTATGGC
CytB_19T	CAACCGCCTTTTCATCAATCGCCCACATTACTCGAGACGTAAATTATGGC
CytB_19A	CAACCGCCTTTTCATCAATCGCCCACATAACTCGAGACGTAAATTATGGC
CytB_19G	CAACCGCCTTTTCATCAATCGCCCACATGACTCGAGACGTAAATTATGGC
CytB_20T	CAACCGCCTTTTCATCAATCGCCCACATCTCTCGAGACGTAAATTATGGC
CytB_20C	CAACCGCCTTTTCATCAATCGCCCACATCCCTCGAGACGTAAATTATGGC
CytB_20G	CAACCGCCTTTTCATCAATCGCCCACATCGCTCGAGACGTAAATTATGGC
CytB_21T	CAACCGCCTTTTCATCAATCGCCCACATCATTCGAGACGTAAATTATGGC
CytB_21A	CAACCGCCTTTTCATCAATCGCCCACATCAATCGAGACGTAAATTATGGC
CytB_21G	CAACCGCCTTTTCATCAATCGCCCACATCAGTCGAGACGTAAATTATGGC
CytB_22A	CAACCGCCTTTTCATCAATCGCCCACATCACACGAGACGTAAATTATGGC
CytB_22C	CAACCGCCTTTTCATCAATCGCCCACATCACCCGAGACGTAAATTATGGC
CytB_22G	CAACCGCCTTTTCATCAATCGCCCACATCACGCGAGACGTAAATTATGGC
CytB_23T	CAACCGCCTTTTCATCAATCGCCCACATCACTTGAGACGTAAATTATGGC
CytB_23A	CAACCGCCTTTTCATCAATCGCCCACATCACTAGAGACGTAAATTATGGC
CytB_23G	CAACCGCCTTTTCATCAATCGCCCACATCACTGGAGACGTAAATTATGGC
CytB_24T	CAACCGCCTTTTCATCAATCGCCCACATCACTCTAGACGTAAATTATGGC
CytB_24A	CAACCGCCTTTTCATCAATCGCCCACATCACTCAAGACGTAAATTATGGC
CytB_24C	CAACCGCCTTTTCATCAATCGCCCACATCACTCCAGACGTAAATTATGGC
CytB_25T	CAACCGCCTTTTCATCAATCGCCCACATCACTCGTGACGTAAATTATGGC
CytB_25C	CAACCGCCTTTTCATCAATCGCCCACATCACTCGCGACGTAAATTATGGC
CytB_25G	CAACCGCCTTTTCATCAATCGCCCACATCACTCGGGACGTAAATTATGGC
CytB_26T	CAACCGCCTTTTCATCAATCGCCCACATCACTCGATACGTAAATTATGGC
CytB_26A	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAAACGTAAATTATGGC
CytB_26C	CAACCGCCTTTTCATCAATCGCCCACATCACTCGACACGTAAATTATGGC
CytB_27T	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGTCGTAAATTATGGC
CytB_27C	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGCCGTAAATTATGGC
CytB_27G	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGGCGTAAATTATGGC
CytB_28T	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGATGTAAATTATGGC
CytB_28A	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGAAGTAAATTATGGC
CytB_28G	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGAGGTAAATTATGGC
CytB_29T	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACTTAAATTATGGC
CytB_29A	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACATAAATTATGGC
CytB_29C	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACCTAAATTATGGC
CytB_30A	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACGAAAATTATGGC
CytB_30C	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACGCAAATTATGGC
CytB_30G	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGACGGAAATTATGGC
CytB_27T28T29T30A	CAACCGCCTTTTCATCAATCGCCCACATCACTCGAGTTTAAAATTATGGC
CytB_1A2A3A4T	CAACCGCCTTTAAATTCAATCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_5A6A7T8T	CAACCGCCTTTTCAAATTTCGCCCACATCACTCGAGACGTAAATTATGGC
CytB_22A23A24A25T	CAACCGCCTTTTCATCAATCGCCCACATCACAAATGACGTAAATTATGGC
REP_ATTTTA	CAACCGCCTTATTTTAATTTTAATTTTAATTTTAATTTTAAAATTATGGC

APPENDIX E: REFERENCES

Andréasson H, Asp A, Alderborn A, Gyllensten U, Allen M. Mitochondrial sequence analysis for forensic identification using pyrosequencing technology. Biotechniques. 2002 Jan;32(1):124-6, 128, 130-3.

Balitzki-Korte B, Anslinger K, Bartsch C, Rolf B. Species identification by means of pyrosequencing the mitochondrial 12S rRNA gene. Int J Legal Med. 2005 Sep;119(5):291-4.

Bellis C, Ashton KJ, Freney L, Blair B, Griffiths LR. A molecular genetic approach for forensic animal species identification. Forensic Sci Int. 2003 Jul 8;134(2-3):99-108.

Bernard PS, Wittwer CT. Homogeneous amplification and variant detection by fluorescent hybridization probes. Clin Chem. 2000 Feb;46(2):147-8.

Blackett RS, Keim P. Big game species identification by deoxyribonucleic acid (DNA) probes. J Forensic Sci. 1992 Mar;37(2):590-6.

Bommarito S, Peyret N, SantaLucia J Jr. Thermodynamic parameters for DNA sequences with dangling ends. Nucleic Acids Res. 2000 May 1;28(9):1929-34.

Branicki W, Kupiec T, Pawlowski R. Validation of cytochrome b sequence analysis as a method of species identification. J Forensic Sci. 2003 Jan;48(1):83-7.

Butler JM, Li J, Shaler TA, Monforte JA, Becker CH. Reliable genotyping of short tandem repeat loci without an allelic ladder using time-of-flight mass spectrometry. Int J Legal Med. 1999;112(1):45-9.

Case-Green SC, Southern EM. Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides. Nucleic Acids Res. 1994 Jan 25;22(2):131-6.

Chardon P, Renard C, Vaiman M. The major histocompatibility complex in swine. Immunol Rev. 1999 Feb;167:179-92.

De Schutter H, Spaepen M, Mc Bride WH, Nuyts S. The clinical relevance of microsatellite alterations in head and neck squamous cell carcinoma: a critical review. Eur J Hum Genet. 2007 Jul;15(7):734-41.

Doktycz MJ, Morris MD, Dormady SJ, Beattie KL, Jacobson KB. Optical melting of 128 octamer DNA duplexes. Effects of base pair location and nearest neighbors on thermal stability. J Biol Chem. 1995 Apr 14;270(15):8439-45.

Doxiadis GG, Otting N, de Groot NG, Bontrop RE. Differential evolutionary MHC class II strategies in humans and rhesus macaques: relevance for biomedical studies. Immunol Rev. 2001 Oct;183:76-85.

Ellegren H. Microsatellites: simple sequences with complex evolution. Nat Rev Genet. 2004 Jun;5(6):435-45.

Gill P, Werrett DJ, Budowle B, Guerrieri R. An assessment of whether SNPs will replace STRs in national DNA databases--joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). Sci Justice. 2004 Jan-Mar;44(1):51-3.

Gudnason H, Dufva M, Bang DD, Wolff A. Comparison of multiple DNA dyes for real-time PCR: effects of dye concentration and sequence composition on DNA amplification and melting temperature. Nucleic Acids Res. 2007;35(19):e127.

Guglich EA, Wilson PJ, White BN. Forensic application of repetitive DNA markers to the species identification of animal tissues. J Forensic Sci. 1994 Mar;39(2):353-61.

Guha S, Kashyap VK. Development of novel heminested PCR assays based on mitochondrial 16s rRNA gene for identification of seven pecora species. BMC Genet. 2005 Aug 11;6:42.

Gut IG. Automation in genotyping of single nucleotide polymorphisms. Hum Mutat. 2001 Jun;17(6):475-92.

Halachmi S, Cohen M, Szargel R, Cohen N. Molecular detection of bladder cancer by fluorescence microsatellite analysis and an automated genetic analyzing system. ScientificWorldJournal. 2007 Sep 17;7:1553-7.

Hayes FN, Lilly EH, Ratliff RL, Smith DA, Williams DL. Thermal transitions in mixtures of polydeoxyribodinucleotides. Biopolymers. 1970;9(9):1105-17.

Herrmann MG, Durtschi JD, Bromley LK, Wittwer CT, Voelkerding KV. Amplicon DNA melting analysis for mutation scanning and genotyping: cross-platform comparison of instruments and dyes. Clin Chem. 2006 Mar;52(3):494-503.

Howell WM, Jobs M, Brookes AJ. iFRET: an improved fluorescence system for DNA-melting analysis. Genome Res. 2002 Sep;12(9):1401-7.

Howell WM. Detection of DNA hybridization using induced fluorescence resonance energy transfer. Methods Mol Biol. 2006;335:33-41.

Igloi GL. Substrate properties of fluorescent ribonucleotides in the terminal transferase-catalyzed labeling of DNA sequencing primers. Biotechniques. 1996 Dec;21(6):1084-92.

Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature. 1985 Jul 4-10;316(6023):76-9.

Kemp JT, Davis RW, White RL, Wang SX, Webb CD. A novel method for STR-based DNA profiling using microarrays. J Forensic Sci. 2005 Sep;50(5):1109-13.

Kwok PY. Methods for genotyping single nucleotide polymorphisms. Annu Rev Genomics Hum Genet. 2001;2:235-58.

Lane MJ, Paner T, Kashin I, Faldasz BD, Li B, Gallo FJ, Benight AS. The thermodynamic advantage of DNA oligonucleotide 'stacking hybridization' reactions: energetics of a DNA nick. Nucleic Acids Res. 1997 Feb 1;25(3):611-7.

Lareu M, Puente J, Sobrino B, Quintáns B, Brión M, Carracedo A. The use of the LightCycler for the detection of Y chromosome SNPs. Forensic Sci Int. 2001 May 15;118(2-3):163-8.

Leijon M, Gräslund A. Effects of sequence and length on imino proton exchange and base pair opening kinetics in DNA oligonucleotide duplexes. Nucleic Acids Res. 1992 Oct 25;20(20):5339-43.

Lion T. Summary: reports on quantitative analysis of chimerism after allogeneic stem cell transplantation by PCR amplification of microsatellite markers and capillary electrophoresis with fluorescence detection. Leukemia. 2003 Jan;17(1):252-4.

Lopez-Andreo M, Lugo L, Garrido-Pertierra A, Prieto MI, Puyet A. Identification and quantitation of species in complex DNA mixtures by real-time polymerase chain reaction. Anal Biochem. 2005 Apr 1;339(1):73-82.

Murray BW, McClymont RA, Strobeck C. Forensic identification of ungulate species using restriction digests of PCR-amplified mitochondrial DNA. J Forensic Sci. 1995 Nov;40(6): 943-51.

Nadir E, Margalit H, Gallily T, Ben-Sasson SA. Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. Proc Natl Acad Sci U S A. 1996 Jun 25;93(13):6470-5.

Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M. Species identification by means of the cytochrome b gene. Int J Legal Med. 2000;114(1-2):23-8.

Pourmand N, Caramuta S, Villablanca A, Mori S, Karhanek M, Wang SX, Davis RW. Branch migration displacement assay with automated heuristic analysis for discrete DNA length measurement using DNA microarrays. Proc Natl Acad Sci U S A. 2007 Apr 10;104(15): 6146-51.

Radtkey R, Feng L, Muralhidar M, Duhon M, Canter D, DiPierro D, Fallon S, Tu E, McElfresh K, Nerenberg M, Sosnowski R. Rapid, high fidelity analysis of simple sequence repeats on an electronically active DNA microchip. Nucleic Acids Res. 2000 Apr 1;28(7):E17.

Reed GH, Wittwer CT. Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clin Chem. 2004 Oct;50(10):1748-54.

SantaLucia J Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U S A. 1998 Feb 17;95(4):1460-5.

Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R. Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philos Trans R Soc Lond B Biol Sci. 2005 Oct 29;360(1462):1805-11.

Spink CH, Garbett N, Chaires JB. Enthalpies of DNA melting in the presence of osmolytes. Biophys Chem. 2007 Mar;126(1-3):176-85.

Takatsu K, Yokomaku T, Kurata S, Kanagawa T. A FRET-based analysis of SNPs without fluorescent probes. Nucleic Acids Res. 2004 Nov 8;32(19):e156.

Tully G, Sullivan KM, Nixon P, Stones RE, Gill P. Rapid detection of mitochondrial sequence polymorphisms using multiplex solid-phase fluorescent minisequencing. Genomics. 1996 May 15;34(1):107-13.

Wambaugh, J. The Blooding. Copyright 1989. New York: Bantam Books.

Watkins NE Jr, SantaLucia J Jr. Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes. Nucleic Acids Res. 2005 Nov 1;33(19):6258-67.

Wetmur JG. DNA probes: applications of the principles of nucleic acid hybridization. Crit Rev Biochem Mol Biol. 1991;26(3-4):227-59.

Zhou L, Myers AN, Vandersteen JG, Wang L, Wittwer CT. Closed-tube genotyping with unlabeled oligonucleotide probes and a saturating DNA dye. Clin Chem. 2004 Aug;50(8): 1328-35.

Zhou L, Wang L, Palais R, Pryor R, Wittwer CT. High-resolution DNA melting analysis for simultaneous mutation scanning and genotyping in solution. Clin Chem. 2005 Oct;51(10): 1770-7.