January 2016

A Novel Forensic DNA Profiling Method Based On Molecular Beacons Without Dna Purification

Benjamin C. Shaw
Eastern Kentucky University

A NOVEL FORENSIC DNA PROFILING METHOD BASED ON MOLECULAR BEACONS WITHOUT DNA PURIFICATION

By

Benjamin Charles Shaw

Thesis Approved:

Dr. Jamie Fredericks
Chair, Advisory Committee

Dr. Judith Jenkins
Member, Advisory Committee

Dr. David Cunningham
Member, Advisory Committee

Dr. Jerry J. Pogatshnik
Dean, Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at Eastern Kentucky University, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made. Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature 

Date 13 April 2016

A NOVEL FORENSIC DNA PROFILING METHOD BASED ON MOLECULAR BEACONS WITHOUT DNA PURIFICATION

By

Benjamin Charles Shaw

Bachelor of Science

Eastern Kentucky University

Richmond, Kentucky

2015

Submitted to the Faculty of the Graduate School of
Eastern Kentucky University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 2016

DEDICATION

This thesis is dedicated to my father and beautiful wife and daughter, who are always pushing me to achieve more.

This is also dedicated to the countless professors on whom I have leaned over the course of this work. Your counsel has been invaluable to my success.

ABSTRACT

Analysis of polymorphisms in nucleic acid sequences provides the basis for identification of individuals and their genetic deficiencies. Currently, the accepted method of analysis for profiling is (STR) profiling. This is a lengthy process, typically taking up to 3 days. The time necessary to generate an STR profile, along with the ever-increasing reliance on DNA to solve crimes, has led to a large DNA sample backlog, with violent crime turnaround taking an average of 103 days. The time and resource investment required for STR analysis is significant, and not all samples generate useful profiles.

The current methods for use of technologies require an isolated template sample. This isolation typically requires hours of extractions and incubations, followed by still more time for analysis. The considerable length of time necessary for this process makes it inherently expensive, while also increasing the backlog. A universal protocol allowing amplification from various, frequently used samples would allow extremely rapid sampling and results. Further, these templates are faster and easier to amplify than standard STRs, which reduces the risk of resources and time on a sample which may not amplify.

Common forensic samples include blood, hair, saliva, and buccal swabs. Using a single, universal protocol to prepare these samples for analysis without extensive isolation allows the simultaneous preparation of multiple samples. Accordingly, this work explores the development of a preparatory method for multiple forensic samples coupled with the optimization of polymerase chain reaction conditions to facilitate the real-time monitoring of the interaction of molecular beacons (MBs) with the template. These MBs can then be used to identify the presence or absence of specific nucleotide polymorphisms. This increase in throughput has extensive application in forensic and medical applications.

ACKNOWLEDGEMENTS

This Masters project has been a great undertaking—one which I would not have been able to accomplish alone. Over the past two years, a number of people have provided invaluable support and assistance. I would like to acknowledge all who made this achievement possible, particularly the following individuals.

My project advisor, Dr. Jamie Fredericks, has allowed me to develop as an independent scientist capable of further research. I haven't always been the easiest graduate student to train, but his patience with me has been exceptional and his guidance has been essential to my success. His willingness to come to the lab even on his days off during vacation for me to ask questions and correct my rookie mistakes has allowed me to finish this project well before I was told was possible.

Dr. Judy Jenkins and Dr. David Cunningham have also provided a great deal of support throughout this project. Both have brought their areas of expertise, well outside of forensic biochemistry, and found ways to make them integral. They were always available for drop-in questions, even those taking much longer than expected.

Finally, I would like to thank my friends and family. Morgan, Seth, Leland, and Tyler: I'm sure there have been weeks or months when I've left you wondering if I was still alive even; this is the product of those times when I fell off the grid. Thank you for shuffling your own plans to make time for me when I did decide to come out of the lab, infrequent as it was. To my father, Craig, for his constant push for me to achieve more: thanks for telling me there's always more to do. To my wife and daughter, Heather and Hayley: I know I wasn't always home for dinner, bed time stories, or play dates in the park, yet you stood by me anyway. Your unwavering support has given me the drive for all my successes in life. I may not be around to say it enough, but I love you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABBREVIATIONS

BHQ1	Black Hole Quencher 1
BLAST	Basic Local Alignment Search Tool
bp	Base Pair
BSA	Bovine Serum Albumin
C_T	Cycle Threshold
DAP	Donor-Acceptor Pair
dATP	Deoxyadenosine Triphosphate
dCTP	Deoxycytosine Triphosphate
dGTP	Deoxyguanosine Triphosphate
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
DPCR	Direct PCR
dsDNA	Double-stranded DNA
dTTP	Deoxythymidine Triphasphate
DVI	Disaster Victim Identification
FAM	6-carboxyfluorescein
FRET	Förster Resonance Energy Transfer
gDNA	Genomic DNA
HEX	6-carboxy-2',4,4',5',7,7'-hexachlorofluorescein
InDel	Insertion-Deletion
kb	Kilobase Pair (1000 bp)
LCN	Low Copy Number
LOD	Limit of Detection
LOQ	Limit of Quantification
MB	Molecular Beacon

MIQE	Minimum Information for qPCR Experiments
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
NFW	Nuclease Free Water
nt	Nucleotide
NTC	No Template Control
PCR	Polymerase Chain Reaction
PIC	Pseudoisocyanine Chloride
qPCR	Quantitative PCR
RFLP	Restriction Fragment Length Polymorphism
SNP	Single Nucleotide Polymorphism
SNR	Signal-to-Noise Ratio
ssDNA	Single-stranded DNA
STR	Short Tandem Repeat
Taq	*Thermus aquaticus*
TBE	Tris-Borate-EDTA Buffer
TE	Tris-EDTA Buffer
T_M	Dissociation Temperature
VNTR	Variable Nucleotide Tandem Repeat
$^{wt}/_{v}$	Weight-by-volume
ΔG	Change in Gibbs' Free Energy

GLOSSARY OF TERMS

1X TBE: Buffer system commonly used in biochemistry for gel electrophoresis. Contains 0.0889 M tris base, 0.0889 M boric acid (as borate), 0.002 M EDTA, adjusted to pH 8.3 with HCl.

1X TE: Buffer system commonly used in PCR techniques. Contains 0.010 M Tris base, 0.001 M EDTA, adjusted to pH 8.0 with HCl.

Allele: A variation of a gene.

Amplicon: Also known as PCR fragment size.

Bathochromic shift: Shift in optical spectrum to a longer wavelength than observed under standard conditions.

Biallelic: A gene or polymorphism for which there are only two known variations.

Chromophore: A moiety capable of absorbing light.

Cofactor: An ion or molecule essential for the function of an enzyme.

Electrophoresis: Separation technique based on size and charge differences used in biochemistry. Capable of separating DNA by size using applied voltage through a porous gel, typically composed of agarose.

Erythrocyte: Also known as a red blood cell. These cells, upon maturation and before release into the bloodstream, purge their nucleus and any DNA.

***Ex vivo*:** Outside of a living organism.

Exonuclease: Enzyme capable of cleaving nucleotides from an end of an oligonucleotide.

Fluorophore: A moiety capable of emitting light after excitation and internal conversion of energy.

Förster distance: Distance at which FRET is 50% efficient for a donor-acceptor pair.

Hemoglobin: Protein in erythrocytes capable of transporting O_2. Uses iron ions as a cofactor.

Heterozygote: An individual who has inherited two different alleles of the same gene.

Homozygote: An individual who has inherited two of the same alleles of the same gene.

Hybridization probe: Any labeled oligonucleotide capable of hydrogen-bonding to a target sequence to give signal. In the absence of a target sequence, no signal should be generated. These probes are highly specific to individual sequences.

Hydrolysis probes: A type of hybridization probe which works by Taq polymerase's intrinsic ability to cleave a fluorophore and quencher from a hybridized oligonucleotide, thereby generating signal. These are linear probes with no secondary structure.

***In silico*:** Software-based

***In vitro*:** Inside living, cultured cells, but outside an organism.

***In vivo*:** Inside a living organism's cells.

InDels: Type of genetic polymorphism. Includes either an inserted or deleted sequence of a short number of nucleotides.

Intercalating dye: Fluorophore which binds to the minor or major grooves in DNA, forming a J-aggregate and increasing fluorescence quantum yield.

Leukocyte: Also known as white blood cells. These cells contain a nucleus and DNA.

Locus (pl. loci)**:** A point or region in an organism's genome.

Low Copy Number sample: Sample with an extremely low concentration of DNA. Preferential amplification of one allele in a heterozygous sample is possible, making interpretation of results complex.

Micro-satellite region: Very short repeating motif capable of use in DNA profiling. See also STR, SNP, InDel, and Y chromosome STR.

Mini-satellite region: Short repeating motif originally used in DNA profiling, but longer than a micro-satellite. See also VNTR.

mtDNA: Mitochondrial DNA. Inherited maternally, and capable of identifying maternal lineage.

Multiplex PCR: Amplification of multiple loci in one reaction mixture.

Oligonucleotide: Short, synthetic polymer of nucleotides.

Polymerase: Enzyme which forms a polymer.

Polymorphism: Variation in a genetic sequence. See also SNP and InDel.

qPCR: See Real-Time PCR.

Quantum yield: Decimal value of the ratio of photons emitted to photons absorbed.

Real-Time PCR: Technique used to amplify DNA while simultaneously monitoring its rate of amplification. Can be adapted to suit many methods and chemistries.

Restriction Fragment Length Polymorphism: Technique which employs the use of bacterial enzymes capable of cleaving at specific DNA sequences. Sometimes used to discern alleles.

Reticulocyte: Mature erythrocyte.

RFLP: See Restriction Fragment Length Polymorphism.

Scorpion probes: Hybridization probe which incorporates a PCR primer joined to a MB. Highly specific, and reduces required amplicon size. More complex to design for specificity.

SNP: Single Nucleotide Polymorphism. Point mutation, or variation of only one nucleotide in a sequence without changing the number of nucleotides.

Southern blotting: Biochemical technique for the analysis of DNA combining electrophoresis and hybridization probes to identify specific nucleic acid sequences.

Spectrofluorometric thermocycler: An instrument capable of adjusting to specified temperatures for specified times, while also collecting fluorescence measurements.

STR: Short Tandem Repeat. Micro-satellite. Currently used as the primary method of DNA profiling. Requires sample collection, purification, amplification, and capillary electrophoresis before analysis.

TaqMan probes: See hydrolysis probes.

Tris: 2-Amino-2-hydroxymethyl-propane-1,3-diol. A common base used in biochemistry buffers.

VNTR: Variable Number Tandem Repeat. Mini-satellite. First method used for DNA profiling. Superseded by STR profiling.

Watson-Crick base pairing: Nucleic acid base pairing for optimal stability. In DNA, adenine pairs with thymine, guanine pairs with cytosine. Deviations are possible but lead to lower stability and are not thermodynamically favored.

CHAPTER 1: INTRODUCTION

1.1: Purpose of the project

Forensic laboratories are under enormous stress as the need for identification of victims and perpetrators rises every year. These laboratories are also responsible for the identification of victims of terrorist attacks and natural disasters. Terrorist attacks are growing at an alarming rate, with ever-increasing devastation as seen recently in Paris, Egypt, and Syria. Natural disasters such as tsunamis and earthquakes can result in large numbers of casualties, particularly in less developed areas such as Haiti and Nepal.

Typically, non-degraded samples can produce full DNA profiles for an individual with relative ease due to current methods. However, in both of these scenarios—terrorist attacks and natural disasters—there is great potential for extreme degradation of human remains. Terrorist attacks leave behind charred remains, damaging the soft tissue normally used for profiling. Disaster victims' remains are often left for days to weeks at a time before recovery, allowing decay. The success of profiling a sample from degraded remains is highly uncertain, and failure results in wasted time and resources.

In the case of DNA profiling, the integrity of the sample is crucial. Ideally, analysts should be able to assess this integrity prior to amplification. Pre-screening a sample would allow analysts and scientists to make better predictions about a degraded sample. Colorimetry (Fredericks et al., 2015) and Fourier Transform Infrared (FTIR) spectroscopy (Fredericks et al., 2012) can be used to determine the probability of successful amplification, and multiplex polymerase chain reaction (PCR) may help to determine sample viability (von Wurmb-Schwark et al., 2004).

Colorimetry and FTIR are limited in their analysis, however, as they do not give direct quantity or quality of DNA available; their uses are derived from correlations between optical properties of samples and the probabilities of successful amplification depending on the length of an amplified fragment, or amplicon. A sample which is deemed suitable for amplification through these methods may still fail to yield a usable profile. Multiplex PCR, while a more reliable standard against which to judge the suitability of a sample, is limited to simply informing the investigator that multiplex PCR is possible. This technique gives no information as to possible preliminary profiles.

Molecular beacons (MBs) are short, single-stranded nucleic acid probes which form a strong secondary structure in the shape of a hairpin. These probes consist of a fluorogenic moiety at the 5' terminus and a nonfluorescent chromophore at the 3' terminus to control signal generation. A MB is composed of a probe region and a stem region. The probe region is a reverse complement to the target DNA and capable of hydrogen bonding with the target. The stem region is a reverse complement to itself, flanking both ends of the probe region and providing the secondary structure necessary. The two-part nature of a MB probe gives rise to its greatest advantage over standard qPCR dyes and linear probes: specificity. Intercalating dyes used for qPCR are completely nonspecific; they bind to double-stranded DNA (dsDNA) indiscriminately. Linear probes do not form secondary structures, so specificity is limited by the length of the probe.

The strong thermodynamic competition between the probe and stem sequences allows for the discrimination of a single nucleotide polymorphism (SNP) by a MB probe (Marras et al., 2003, Marras et al., 1999, Mhlanga and Malmberg, 2001, Root et al., 2004). These polymorphisms are the smallest possible in a sequence—only one base is changed, and none deleted or inserted. The stem region

provides a thermodynamic drive to hold the dyes in close proximity and quench signal in the absence of a target. However, in the presence of an exact target, the MB will undergo conformational changes allowing fluorescence, and thus signal. These MBs are typically used in a real-time PCR (qPCR) format without a lengthy melt analysis. This technology is roughly 1-2 days faster than current STR technology, and hours faster than HyBeacon probes, from sample collection to results (Butler, 2005, French et al., 2006).

There are no satellite regions—tandemly repeating, noncoding DNA—to analyze, and as such the amplicon sizes necessary are extremely small. These fragment sizes can be limited to the length of the forward and reverse primers, plus the length of the MB used in the assay, resulting in a typical amplicon length of less than 150 bp, while satellite analysis can require upwards of 350 bp (Butler, 2005). Amplicon size is inversely proportional to the probability of successful amplification from a degraded sample, and as such it is advantageous to use as small an amplicon as possible in analysis.

Given the MB's ability to discriminate SNPs, there is potential for the development of a forensic panel based on this technology. This would not only allow an investigator to determine the probability of a successful profile based on currently accepted techniques such as STRs, but also give a preliminary profile. This preliminary profile could be used as a benchmark for determining the likelihood of a match to a reference sample from a disaster victim, terror attack victim, or perpetrator.

1.2: Overview of the project

Profiling individuals through DNA can take many routes: VNTR, STR, SNP, Y chromosome STR, and InDels. Variable Number Tandem Repeat (VNTR)

regions are extremely long in comparison to other methods. Since all DNA profiling is the product of statistical analysis of genotypes, and a longer region is less likely to be successfully amplified, VNTR profiling is inherently less reliable than more current methods. This method was made obsolete upon the development of the STR method. Single Nucleotide Polymorphism (SNP) profiling requires the smallest amplicon length, but also has the lowest power of discrimination; there are only three possible genotypes per locus. Insertion-Deletion (InDel) polymorphisms are similar to SNPs in that there are only 3 possible genotypes; however, they are more compatible with nucleic acid probes such as MBs given the additional differences in the sequence. The STR method relies on much shorter repeating motifs than VNTRs, but still requires a 400 bp or longer amplicon. A variation of this method, Y-chromosome STR analysis, is specific to men only.

Currently, the accepted method of DNA profiling for legal proceedings such as trials is STR analysis of 13 core loci, along with the amelogenin gene on the sex chromosome for gender discrimination. This method is time- and resource-consuming, and as such a DNA testing backlog has occurred. As of 2011, more than 100,000 samples were awaiting testing with 83,000 of those over 30 days old. According to the National Institute of Justice, an additional 241,000 samples were added for analysis in 2011 (Nelson et al., 2013).

Current STR methods require sample extraction and purification, amplification, and analysis; this is a 2-3 day process, with multiple opportunities for contamination. This project proposes details the development of a single-step method for DNA sample analysis and profiling. The proposed method uses molecular beacons (MBs) in conjunction with established direct PCR methods. The MB targets are single nucleotide polymorphisms (SNPs). The literature review establishes the mechanism by which a MB is capable of distinguishing even a single

nucleotide difference. This is a proof-in-concept project, establishing the capability of MBs to discriminate alleles in a complex analytical matrix, leading to further work using insertion-deletion (InDel) alleles for the final discriminatory panel. InDels will have increased discriminatory ability compared to SNPs, as there is a greater thermodynamic penalty due to increased base mismatching.

1.3: Definition of a molecular beacon

A MB is defined as a synthetic oligonucleotide hybridization probe. It has a self-complementary stem sequence flanking a target-complementary probe sequence. The probe sequence is typically at least twice as many nucleotides (nt) long as the stem sequence. A fluorescent moiety is covalently linked to the 5' terminus of the oligonucleotide, and an appropriate nonfluorescent quenching moiety is covalently linked to the 3' terminus. These are located such that the stem sequence holds them in close proximity, forming a hairpin-like structure under standard conditions. Their proximity facilitates quenching. Upon the introduction of a complementary target in the form of single-stranded DNA (ssDNA), the MB will spontaneously unfold and anneal, changing conformation and separating the fluorescent and quenching moieties. This separation results in signal generation.

1.4: Forensic samples examined

This project focuses on the application MBs to forensic biology, and this is a proof-in-concept. As such, only simple, freshly collected, unpreserved samples were studied. These samples include fresh hair, saliva, buccal (cheek) swabs, and blood. Future studies should examine samples with anticoagulants or other matrix additives in blood samples, along with other common forensic samples such as

semen and vaginal swabs. Mixtures of sample types or individuals should also be studied at a later date.

1.5: Aim and scope of the project

The aim of the project is to develop a sample preparation and MB assay methodology that is sufficiently robust to known PCR inhibitors and optical impediments for non-purified samples. It is hypothesized that MB assays can be adapted to a direct PCR (DPCR) format. Further, it is hypothesized that the length of time required to perform these assays can be considerably shortened. Multiple experiment types were necessary to test these hypotheses. All experiments in this project were PCR-based, primarily serving either to verify previous work or optimize the newly developed assay. Novel work for this project included the use of MB assays in a DPCR format, capable of successfully genotyping individuals. Further, this was done in a 2-step PCR cycling format; a thorough literature review did not find any other work showing a 2-step MB assay.

The objectives:

1. Verify primer functionality
2. Develop a thermal profile of the MBs
3. Verify a working MB assay
4. Screen for genotypes necessary for the project
5. Test various PCR master mixes for compatibility
6. Develop a working MB-DPCR assay
7. Optimize the MB-DPCR assay

CHAPTER 2: LITERATURE REVIEW

CHAPTER 2: LITERATURE REVIEW

2.1: Introduction

The following is a primer on the current state of the field of single nucleotide polymorphism (SNP) genotyping and the theoretical bases of the project as a whole. Contents include in-depth descriptions of the mechanics of PCR, an introduction to real-time PCR, the mechanics of molecular beacons (MBs) and their design, theory of fluorescence and quenching, challenges of direct PCR, and the basis of DNA profiling. Where appropriate, more exhaustive reviews of a particular topic are suggested.

2.2: Mechanics of PCR

The current state of forensic biology is largely a result of the development of polymerase chain reaction (PCR). This technique was first described in a 1986 patent application by K. B. Mullis as a way to synthetically replicate DNA in solution (Mullis et al., 1987). This replication is necessary for forensic biology as "many forensic samples would be impossible to analyze…[because] obtaining a cleaner, more concentrated sample is normally not possible," and current forensic methods require high quality DNA (Butler, 2005).

As PCR is a chain reaction, each nascent strand of DNA functions as a template in template in the following step to synthesize more strands of DNA (Mullis et al., 1987, Reynolds et al., 1991, Saiki et al., 1988, Butler, 2005). This is an enzymatic process involving repeated heating and cooling of a reaction with a complex analytical matrix. There are three phases of a typical PCR cycling: a 95°C

denaturation phase to break the hydrogen bonds between strands of dsDNA, a variable (but typically 55°C) annealing phase to allow primers and enzymes to form new hydrogen bonds with individual strands of dsDNA, and a 72°C extension phase during which the enzyme is best able to polymerize a nascent strand. A diagram of the thermocycle is shown as

Figure 2.1. Figure 2.2: Diagram of PCR mechanism As PCR continues, products follow geometric growth. Each cycle doubles the amount of template present in the previous cycle, provided it is working at 100% efficiency. For instance, after three cycles, the original number of copies is multiplied by 2^3, or 8; after 20 cycles, the original copy number is multiplied by 2^{20}, or roughly 10^6 times as many copies. Figure 2.2 shows a diagram of the PCR mechanism.

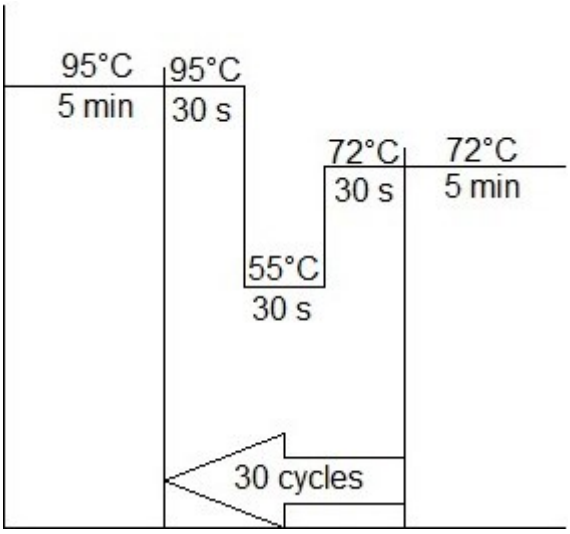

Figure 2.1: PCR thermocycling
This figure shows an example setup for a PCR reaction. The initial 95°C hold is to fully denature all dsDNA. The 30 second, 95°C hold is used to again denature the dsDNA at the beginning of each cycle. The 30 second, 55°C hold allows the primers and polymerase to anneal to the existing strands. The 30 second, 72°C hold allows the polymerase to synthesize a nascent strand complementary to the existing strand, at the site of primer annealing.

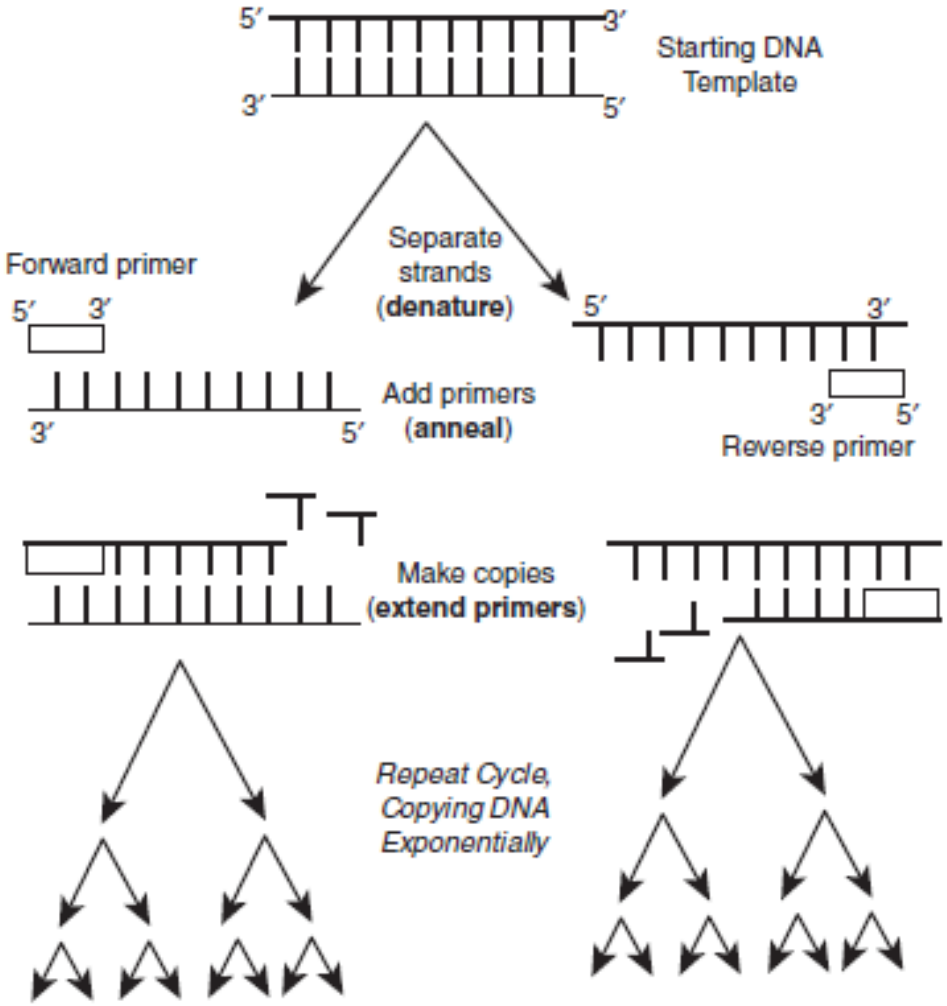

Figure 2.2: Diagram of PCR mechanism
Butler, J. M. (2005) Forensic DNA typing : biology, technology, and genetics of STR markers, 2nd ed. ed., Elsevier Academic Press, Burlington, MA.
This figure shows the molecular-level process of each phase of the PCR. A denaturation phase separates the individual strands, a primer anneals to the single strands, and the enzyme polymerizes a nascent strand. This occurs for each strand of the dsDNA during each cycle, giving rise to its geometric growth. This is represented as the 2^n growth, where n is the cycle number; this is an ideal, theoretical limitation.

Table 2.1 shows a list of common components in a PCR reaction. A Tris-HCl solution buffered to pH 8.3 at 25°C is commonly employed as the bulk of the solution. Divalent cations are required as cofactors (Berg, 2012). Magnesium is the most commonly used cation for this purpose; Mg^{2+} will coordinate with the 3' hydroxyl group of the primer or nascent strand, while also coordinating with the dNTP (Antony and Subramaniam, 2002). Since the terminal phosphate group will still be negatively charged, another Mg^{2+} binds to the dNTP temporarily to allow the dNTP to coordinate with the primer or nascent strand (Berg, 2012). Since all that is required is a divalent cation, manganese (Mn^{2+}) can also be used, but has a markedly increased error rate due to the size difference (Wilson and Keefe, 2001, Zhou et al.). Potassium chloride is used to coordinate the primer and template strands. The four dNTPs—deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxyguanosine triphosphate (dGTP), and deoxycytidine triphosphate (dCTP)—are used as the building blocks of the nascent strand. A DNA polymerase is necessary to form the nascent strand. A thermally stable polymerase enzyme is crucial to PCR. Since enzymes are proteins, these polymerases are susceptible to denaturation due to heat. Thermal denaturation would almost certainly occur after one cycle with human DNA polymerase, thus necessitating aliquots of expensive enzymes with each iteration (Saiki et al., 1988). A thermally stable polymerase would not require iterative aliquots, and would be able to withstand extended incubations near boiling. *Thermus aquaticus* (Taq) was discovered to have such an enzyme in 1976 (Chien et al., 1976).

Bovine serum albumin (BSA) serves a few purposes. Bovine serum albumin is used to stabilize the enzyme and prevent adsorption to the reaction vessel. The addition of BSA to a PCR reaction helps to prevent any copurified

contaminants or inhibitors from interfering with the polymerase enzyme (Comey et al., 1994). Primers are used to provide an initiation point for polymerization.

Table 2.1: Common PCR components and concentrations
Adapted from Carracedo, A. (2005) Forensic DNA typing protocols, Humana Press, Totowa, N.J.

Reagent	Optimal Concentration
Tris-HCl, pH 8.3 (25°C)	10-50 mM
Magnesium chloride	1.2-2.5 mM
Potassium chloride	50 mM
Deoxynucleotide triphosphates (dNTPs)	200 µM each dATP, dTTP, dCTP, dGTP
DNA Polymerase, thermally stable	0.5-5 U
Bovine serum albumin	100 µg/mL
Primers	0.1-1.0 µM
Template DNA	1-10 ng purified genomic DNA

2.3: Introduction to qPCR: terminology and mechanism

Most current qPCR methods are based on fluorescence technology. These fluorescent methods require the use of a spectrofluorometric thermocycler. This instrument is capable of heating and cooling reaction containers from room temperature to almost 100°C (although some have a wider range) while simultaneously using a light source to excite fluorophores and a detector to measure the emitted light at a certain wavelength. This instrument is the pillar of qPCR.

A fluorophore must also be used in conjunction with a spectrofluorometric thermocycler. Depending on the technology used, this fluorophore can exist freely (i.e., unbound in solution), or attached to a small oligonucleotide (hybridization

probe) (Li, 2002, Livak et al., 1995, Nazarenko et al., 1997, Solinas et al., 2001, Tyagi and Kramer, 1996, Whitcombe et al., 1999, Morrison et al., 1998). Free fluorophores, such as SYBR Green I, are intercalating dyes which bind to either the major or minor grooves of DNA, thereby greatly increasing quantum yield (Zipper et al., 2004). Hybridization probes include MBs, Scorpion probes, hydrolysis (TaqMan) probes, and others based on fluorescence quenching and oligonucleotide hybridization (Didenko, 2006).

Both free fluorophore and hybridization probe assays are considered qPCR techniques, and can be used for qualitative and quantitative purposes. Hybridization probes, however, are designed with specificity in mind, and are much more useful in qualitative experiments than free fluorophores such as SYBR Green I. Since SYBR Green I binds to dsDNA indiscriminately, nothing is known about the specific genotype of a given template at the amplification site (Mackay et al., 2002, Morrison et al., 1998).

According to Bustin *et al.*, even qualitative experiments require a minimum amount of quantification in order to deduce qualitative results; that is, there must be some lower or upper limit at which the results are deemed positive or negative (Bustin et al., 2009). No minimum data set had yet been established as of 2008, and misleading data was being published (Böhlenius et al., 2007, Garson et al., 2009). Thus, the Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) was developed. BustinThis publication is centered around standardization of terminology, increasing transparency, and reducing ambiguity in interpretation of results (Bustin et al., 2009). Nomenclature for both qualitative and quantitative real-time PCR experiments is termed qPCR. While qPCR traditionally involves an intercalating dye used for quantification of DNA, any real-time PCR experiment is inherently quantitative, regardless of the mechanism used. Any

analytical technique must have a limit of detection (LOD), which is important in determining the sensitivity of an assay. The analytical specificity must also be included; this demonstrates the number of correctly identified samples in an assay, compared to the total number of samples. A low analytical specificity results in an unviable assay, as it will produce false positives or false negatives, thus limiting its use. Accuracy is also included, specifically for quantification experiments. This is a comparison between measured and actual concentrations. Precision, both short- and long-term, are to be reported as well (Bustin et al., 2009).

The above qualities required for MIQE data are fulfilled in a number of ways. Sequences, thermodynamic data, PCR conditions, vendors, and many other items are required in order to publish. A suggested list of required items for publication is shown in **Error! Reference source not found.**. Use of *in silico* tools such as BLAST (NCBI, Maryland, USA) are virtually required to fulfill the requirements, and their use is good practice for any qPCR experimental design.

The mechanism of qPCR is identical to that of standard PCR, with the exception that the fluorophore or labeled oligonucleotide is added to the reaction, and during one of the phases of cycling (depending on the technique used), fluorescent measurements are obtained. The fluorescent data is plotted as a function of cycle number and interpreted by the researcher. This gives rise to the name "real-time PCR," as it allows the researcher the opportunity to monitor amplification through fluorescence measurements during the experiment as opposed to only after amplification. The fluorescence signal indicates the presence of either dsDNA or even a specific sequence, depending on the technique employed, and is also capable of quantitation.

2.4: Overview of molecular beacons: an introduction to the biochemical and analytical bases of signal generation

Molecular beacons were described in 1996 as a diagnostic tool for analyzing nucleic acid sequences (Tyagi and Kramer, 1996). These MBs consist of four components: a complementary probe sequence, a stem sequence, a reporting fluorophore attached to one side of the stem, and a fluorescent quencher attached to the other side of the stem. A typical MB is single-stranded, and between 15-30 bp total length (Goel et al., 2005, Kostrikis et al., 1998b, Tyagi and Kramer, 1996). An example of a MB is shown as Figure 2.3: Basic diagram of a molecular beacon (Goel et al., 2005). These MBs are used in conjunction with typical polymerase chain reaction components for real-time monitoring. During the reaction, the annealing phase temperature is set at 3-5°C below the calculated melting temperature of the probe sequence. This allows hybridization of the MB to a template strand. Once the MB comes in contact with a template strand at this temperature, the MB will expand and separate the fluorophore from the quencher; Figure 2.4 shows a diagram of a hybridized MB (Wang et al., 2009).

A probe sequence is designed to be a reverse complement to the intended target. This allows for the probe sequence to bind to the target through traditional Watson-Crick base pairing. The probe sequence is typically at least twice as long as the stem sequence in order to give thermodynamic favorability of interaction with a target sequence rather than internal coupling (Tsourkas et al., 2003). The interaction of a MB with a target sequence is the basis for the signal generation. The thermodynamic favorability of the intermolecular versus intramolecular interactions ensures optimal signal generation.

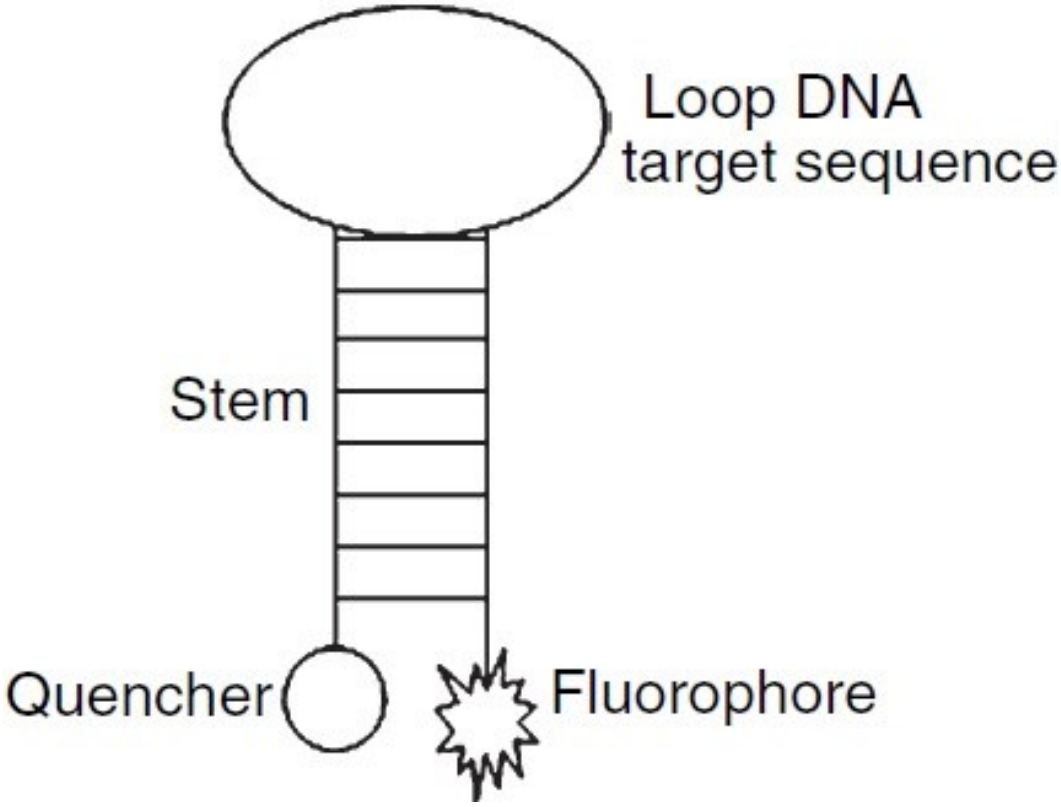

Figure 2.3: Basic diagram of a molecular beacon
Goel, G. et al. (2005) Molecular beacon: a multitask probe, Journal of Applied Microbiology 99, 435-442
The "Loop DNA target sequence" is the probe segment; it is a reverse complement to the target site, which allows hydrogen bonding and duplex formation. The "stem" is a self-complementary sequence flanking the probe segment and holds the fluorophore and quencher dyes in close proximity to facilitate quencing in the absence of target. The probe segment is at least twice as long as the stem sequence, allowing a thermodynamic drive to separate the stem sequence and separate the dyes to generate fluorescent signal.

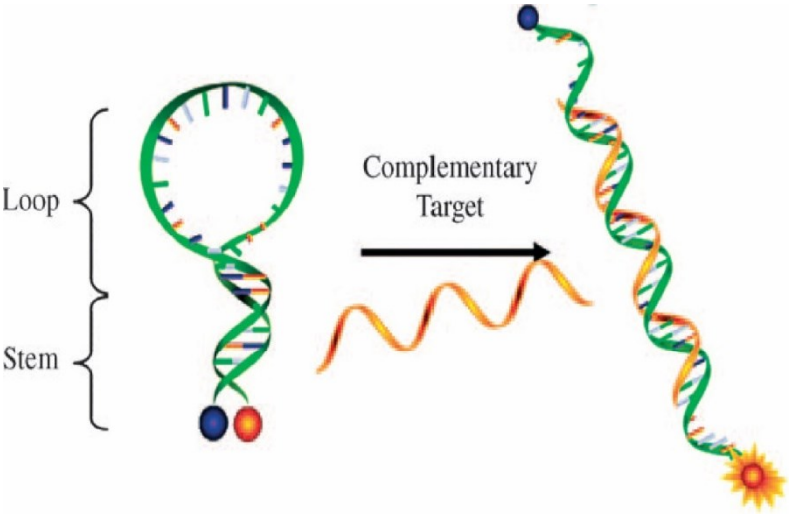

Figure 2.4: Structure of a hybridized molecular beacon
Wang, K. M. et al. (2009) Molecular Engineering of DNA: Molecular Beacons, Angew Chem Int Edit 48, 856-870.
This figure shows the structure of a MB in greater detail, along with its hybridization to a target.

The stem sequence has the opposite task, however. The stem sequence is typically four to seven bases in length flanking either side of the probe sequence, not complementary to the target sequence, but in fact complementary to itself (Kostrikis et al., 1998b, Ortiz et al., 1998, Tyagi et al., 1998, Tyagi and Kramer, 1996). The purpose behind this is to form a secondary hairpin structure. This hairpin structure brings the DAP within the Förster distance, and allows the DAP to resonate energy internally when not interacting with the intended target. This allows for differentiation between alleles.

The quencher-fluorophore pair works through two processes: Förster Resonance Energy Transfer (FRET) and static quenching (Tyagi and Kramer, 1996). In FRET, the donor molecule does not actually emit a photon; rather, it acts as an oscillating dipole which resonates energy to the acceptor, capable of receiving energy in the same resonance frequency (Lakowicz, 2013). The efficiency of this resonance is dependent upon the overlap integral (Clegg, 2009, Lakowicz, 2013).

The overlap integral is a function of the area enclosed underneath the emission spectrum of the donor and the absorption spectrum of the acceptor. Additionally, the efficiency is a distance-dependent, inverse, sixth order relationship between the DAP. The distance at which efficiency for a DAP is 50% is termed the Förster distance. The Förster distance is between 20-60 Å for most molecules (Lakowicz, 2013). Thus, when designing a MB, the choice of a proper donor-acceptor pair (DAP) is essential. Static quenching is less reliant on spectral overlap; the primary factor in static quenching is the orientation and proximity of the dyes. This is discussed in greater detail in 2.9: Signal generation and quenching: Dynamic and Static quenching.

The purpose behind the DAP is signal generation (Goel et al., 2005). During the annealing phase of the PCR, in the absence of the target sequence, the DAP is in close proximity. The average distance between strands in a DNA helix is between 22-26 Å, well within the Förster distance of most DAPs (Lakowicz, 2013, Mandelkern et al., 1981, Wu and Brand, 1994). The DAP exists split between the two strands, and as such are spaced roughly 25 Å apart when closed. This allows the DAP to resonate energy internally, thus quenching the signal by preventing photon emission. If the target sequence is present and interacts with the MB, the MB expands and separates the DAP (Goel et al., 2005). Given that the average distance between DNA bases is 3.3 Å, the DAP is separated by over 80 Å in a 25 bp MB. This distance prevents FRET, and optical signal can be generated from an extended MB.

The signal can be monitored using optical instruments coupled with a PCR thermocycler. A qPCR incorporates an excitation source and a photodetector to measure any photon emission. Monitoring is possible during the annealing phase only; during an extension or denaturation phase of the PCR, the temperature would

dissociate the MB itself, preventing FRET, and this signal cannot be distinguished from a signal attributed to a hybridized MB (Goel et al., 2005). The qPCR emits a specific wavelength of light corresponding to the fluorophore during the annealing phase while simultaneously monitoring any emission attributed to the fluorophorea high-output LED excitation source and photomultiplier tube with monochromator. The fluorescent intensity is plotted against the cycle number during which the data was collected. An example of this data is shown Figure 2.5 (Abravaya et al., 2003). In order to discriminate alleles, the data must show an exponential increase in signal at some point for a positive result, or a linear relationship for a negative result. A linear relationship results from the lack of exponential signal growth due to the absence of the allele interrogated.The exponential increase is due to the mechanism of PCR itself. For every cycle of PCR, the theoretical increase in template strands is 2^n, where n is the cycle number.

The likelihood of an interaction of a MB and template strand is a probability distribution, and with an exponential increase in template strand comes an exponential increase in probability of interaction. The thermodynamic order of the interaction between the MB and target sequence is second-order when concentrations are within an order of magnitude, which accounts for the exponential increase in signal. After a few cycles of this exponential increase, though, the template concentration exceeds the order of magnitude parameter, which results in a pseudo-first-order interactionbetween template DNA and MBs, given the mass excess of template in comparison to the MBs, which are generating the signal (Tsourkas et al., 2003).

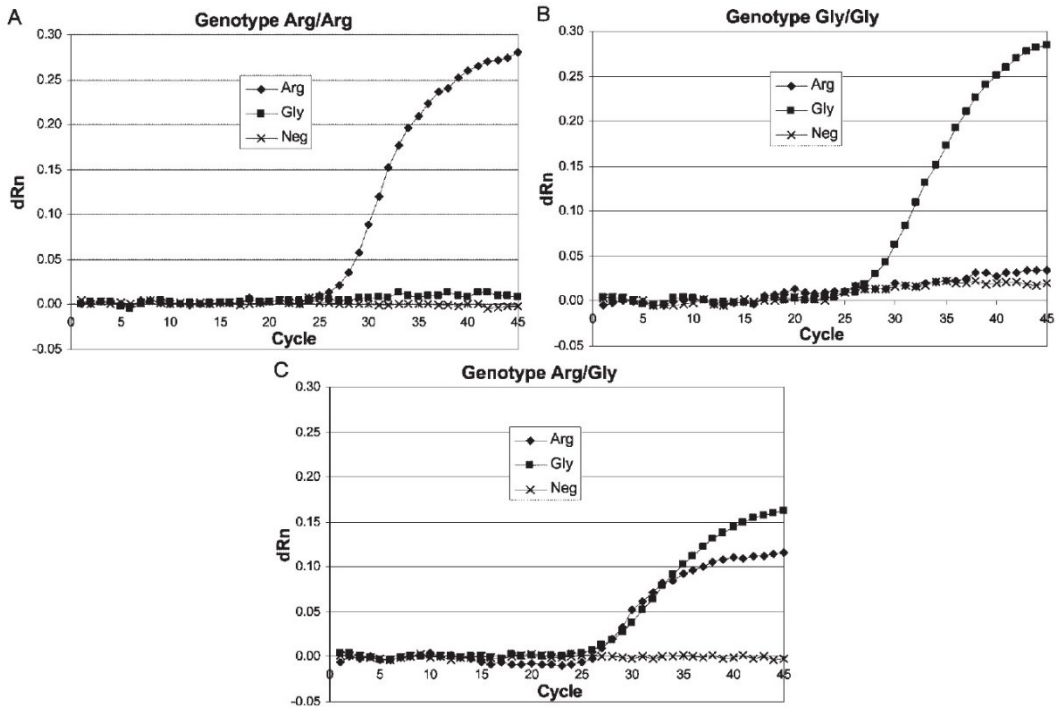

Figure 2.5: Graphical representation of spectral genotyping
Parts A and B show homozygous genotypes, while Part C shows a heterozygous genotype. The y-axis represents fluorescent signal. The marked lines correspond to signal from the respective probes.
Abravaya, K., Huff, J., Marshall, R., Merchant, B., Mullen, C., Schneider, G., and Robinson, J. (2003) Molecular beacons as diagnostic tools: Technology and applications, Clin Chem Lab Med 41*, 468-474.*

This pseudo-first-order interaction results in data that appears linear. The resulting positive signal thus shows an exponential increase in signal for a time, followed by a linear increase in signal due to the change in kinetics. Negative results, on the other hand, show only a linear relationship between relative fluorescent intensity and cycle number, as the signal is only first-order. There is no interaction between a non-complementary template and a MB, meaning the kinetics of the signal depend only on the concentration of the MB which remains constant throughout the PCR.

2.5: Experimental design for MB assays

Experimental design of MB assays requires care in variable manipulation, as with any experimental design. Only one variable should be manipulated per experiment. For example, only annealing temperature should be changed, not annealing temperature and time. Additionally, a known, purified sample should be used as a positive control which is capable of eliciting signal from all MBs in the assay. A no-template control (NTC), containing all components with nuclease-free water (NFW) in place of template DNA, should also be used to check for contamination in the assay.

2.6: Thermodynamic basis for the hybridization of MBs to targets

The MB extension is a thermodynamically driven process. During the annealing phase of a PCR cycle, the template strands of DNA have already been denatured; this melting of template strands produces ssDNA for a brief time. In the presence of complementary DNA at elevated temperatures, with an appropriate salt concentration, the target DNA will begin wrapping itself around the MB probe sequence, causing the MB stem sequence to separate and form a MB-ssDNA hybrid. This tendency is governed by the temperature of the cycle, the complementarity of the template ssDNA, and the salt concentration of the solution. For the purposes of this review, the effects of salt concentration will not be discussed; no work was done concerning the salt concentration, and as such does not impact the scope of this research.

As a ssDNA-MB hybrid forms, the MB stem is less thermodynamically favored. The ΔG of the stem becomes larger in magnitude, while still negative, than that of the ssDNA-MB hybrid, and the stem separates fully, allowing complete

hybridization of the probe sequence. This separation causes extension of the MB, thereby separating the DAP. The separation of the DAP allows fluorescence to occur.

2.7: Thermodynamic basis for allelic discrimination

During ssDNA-MB hybridization, the thermodynamic favorability of hybrid formation outweighs that of stem formation in a perfect complement between the target and probe sequence. However, deviations from complementarity impose thermodynamic penalties, making stem formation preferable. According to Aboul-ela, a base-base mismatch has approximately a 2.1-5.1 $kcal \cdot mol^{-1}$ difference in hybrid formation at 50°C (Aboul-ela et al.). This penalty does not take into account the competing thermodynamic favorability of stem formation, thus causing underestimation of the actual penalty.

In the case of a base-base mismatch, the thermodynamic penalty causes the ΔG of hybridization to stay lower in magnitude than that of stem formation, and the beacon remains in a closed conformation at the annealing temperature. It is important to ensure the use of an optimal annealing temperature for allelic discrimination, as false positives—results which generate signal where none should be detected—are possible.

The most likely cause of a false positive result is the use of a suboptimal annealing temperature. If the annealing temperature is set too far below optimum, the thermodynamic penalty of a base-base mismatch is minimized, thus resulting in hybridization and extension without a perfect complement. If the temperature is sufficiently low, the bases surrounding the mismatch will still orient in such a way to cause hybridization, especially in the case of a single nucleotide polymorphism (SNP). Given that the amount of hydrogen bonding in a ssDNA-MB hybrid is

greater than that of a MB stem, even with a base-base mismatch, the temperature of the reaction must be raised to prevent this nontarget hybridization. This hybridization is generally a small fraction of the total MB concentration, but may still be significant in terms of qualitative allelic discrimination (Tsourkas et al., 2003, Tyagi et al., 1998, Tyagi and Kramer, 2012).

2.8: Kinetics of hybridization

Molecular beacon hybridization reactions are second-order overall, first order with respect to the concentration of both the ssDNA target and the molecular beacon. This is intuitive as the rate of hybrid formation is dependent on the amounts of both target and beacon present. If one of the two is absent, hybridization—and thus exponential signal generation—cannot occur.

Tsourkas *et al*. confirmed this in a 2003 paper detailing the kinetics of hybridization interactions (Tsourkas et al., 2003). They discussed a three-step mechanism for the hybridization of a MB to a target. Initially, the target begins bonding to the probe domain of the MB, while the signal remains quenched. As the number of bases hybridized increases, the stem domain separates. Once the stem separates and signal generation begins, the remaining bases on the probe domain hybridize and the DAP is at maximum separation. The difference between the signal at step two and step three is minimal. As the next section describes, the length of an extended MB would far exceed the distance required for any mechanism of quenching.

At a molecular level, the rate of hybridization is governed by the following equation set:

$$B + T \overset{k_1}{\rightarrow} D$$
$$B + T \underset{k_2}{\leftarrow} D$$

Equation 2.1 Equilibrium rate expression of hybridization
Variables are as follows: B—beacon; T—target; D—duplex (hybrid).

$$\frac{d[D]}{dt} = k_1[B][T] - k_2[D]$$

Equation 2.2: Differential form of rate expression of hybridization
Variables are as follows: B—beacon; T—target; D—duplex (hybrid).

This leads to the differential form expressed in Equation 2.2. The on-rate constant, k_1, can be affected by stem length, probe length, and labelling dyes. There is a moderate, positive correlation between the probe length and on-rate constant. The addition of a single base to the probe sequence can double the rate constant. The addition of a single base to the stem sequence, however, can increase the rate constant by an order of magnitude, thus showing a strong, positive correlation between the stem length and on-rate constant (Tsourkas et al., 2003).

At a macroscopic level, this second order interaction can be visualized at the exponential phase of the amplification (Li et al., 2008). In Figure 2.6, which depicts a typical positive response for an allele in the presence of an appropriate MB,, there are four highlighted regions. The normalized fluorescence is plotted along the Y axis against the cycle number on the X axis. The first region, a mostly horizontal line, is the baseline phase. This region describes a period in the amplification during which the concentration of DNA is many orders of magnitude less than that of the MBs. It provides the background fluorescence measurement. During this time, amplification is occurring but is not represented by the fluorescence data as there is not enough DNA target in solution to provide a meaningful amount of DNA-MB hybrid. The second region shows exponential growth of the fluorescent signal. At this point, there has been enough replication such that the DNA target and MB are in roughly equivalent concentration, allowing

visualization of the second-order kinetics. The DNA target will continue to grow in concentration, and eventually eclipse the concentration of MB by many orders of magnitude. At this point, the data suggests a pseudo-first-order process as the DNA. This is represented by a linear growth in fluorescent signal. This pseudo-first-order process is heavily influenced by increasing competition between complementary, nascent strands and the MBs. Since the MBs are in a constant concentration and the nascent strands are theoretically doubling every cycle, the nascent strands will have a much higher probability of associating to the MB target. This increased competition prevents MB hybridization, thus slowing the growth of signal generation and causing apparent linearity. The linear growth occurs as hybridization becomes more likely every cycle, but not enough MBs are present to show exponential growth. Finally, the plateau phase occurs when the reaction is depleted of primers or dNTPs, as no further growth can occur. As amplification of DNA ceases, so too does the signal growth.

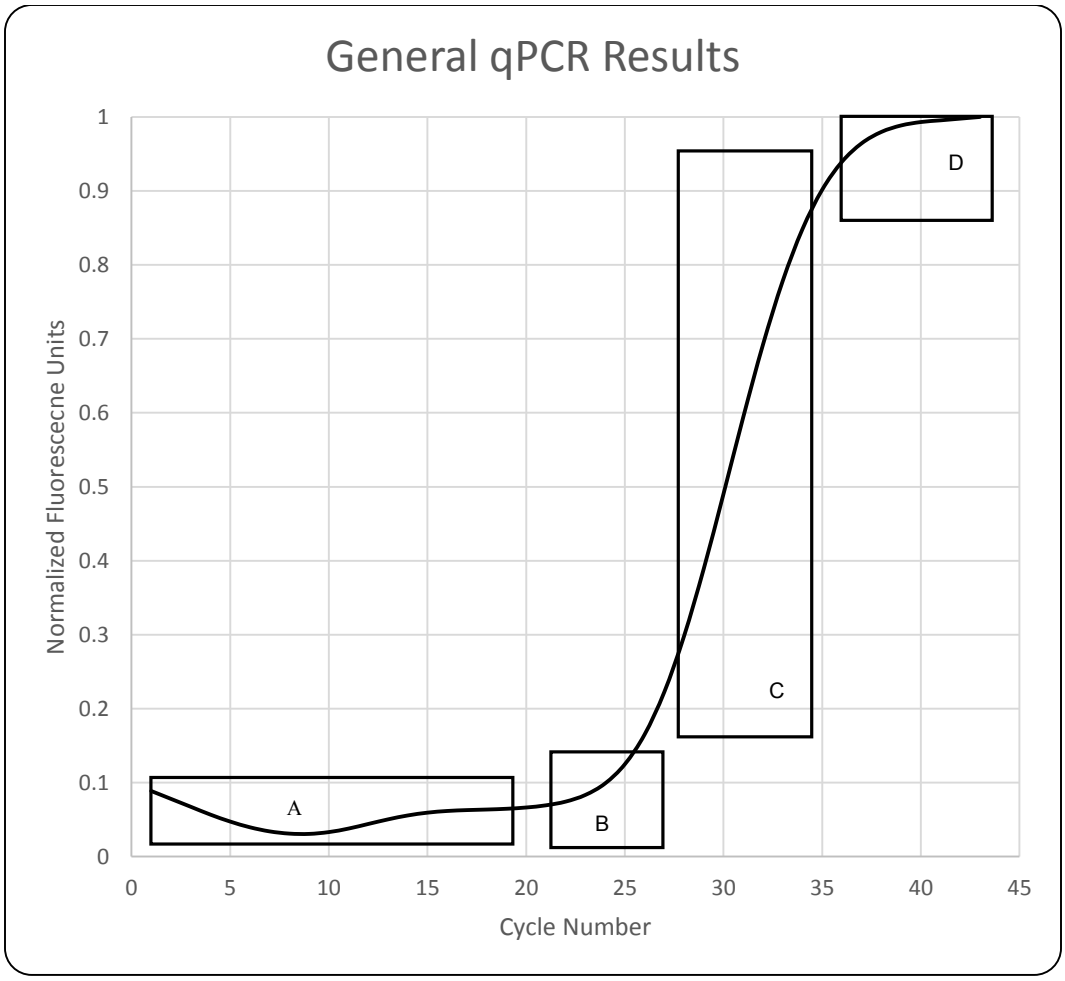

Figure 2.6: Graphical results of a general qPCR reaction
Part A shows the baseline phase. Part B corresponds to the exponential phase, with second-order kinetics shown. Part C is the linear phase, where DNA target concentration has greatly exceeded the MB concentration. Part D is the plateau phase, or the end of amplification.

2.9: Signal generation and quenching: Dynamic and Static quenching

The scope of this project does not include studies into fluorescent data generation, and as such a fully in-depth review of fluorescence and quenching is not provided. The following review provides the foundations for what is necessary to understand the scope of the project. For a more detailed review of fluorescence and quenching, see the most current edition of *Principles of Fluorescence Spectroscopy* by Joseph R. Lakowicz (Lakowicz, 2013).

2.9.1: Static quenching: Contact quenching

Static quenching, or contact quenching, occurs as a result of strong intramolecular dipole-dipole interactions between two labeling moieties (Kasha, 1963, Kasha, 1991, Scholes and Ghiggino, 1994). In the case of MBs, these labeling moieties are the fluorophore and quencher. According to Bernacchi, this formation of a ground-state intramolecular dimer allows strong coupling between transition dipoles, resulting in delocalization of excitation and large changes in spectral properties (Bernacchi and Mély, 2001).

Dyes in solution, such as 6-carboxyfluorescein (FAM), 6-carboxy-2',4,4',5',7,7'-hexachlorofluorescein (HEX), and Black Hole Quencher 1 ((BHQ1), tend to aggregate due to electrostatic, steric, and hydrophobic forces (Arbeloa, 1981a, Arbeloa, 1981b, Johansson et al., 2002, Khairutdinov and Serpone, 1997, West and Pearce, 1965). Structures of FAM, HEX, and BHQ1 are shown in . Given the steric conditions of a MB, however, it is not so much the inherent, spontaneous aggregation of the dyes which causes this interaction as it is the close proximity in which they are held by the stem sequence (Tyagi et al., 1998, Tyagi and Kramer, 1996). Between the stem sequence immobilizing the dyes in close proximity and

the aggregation of the DAP, the DAP forms an H-aggregate heterodimer with significantly diminished fluorescence (Wang et al., 2000).

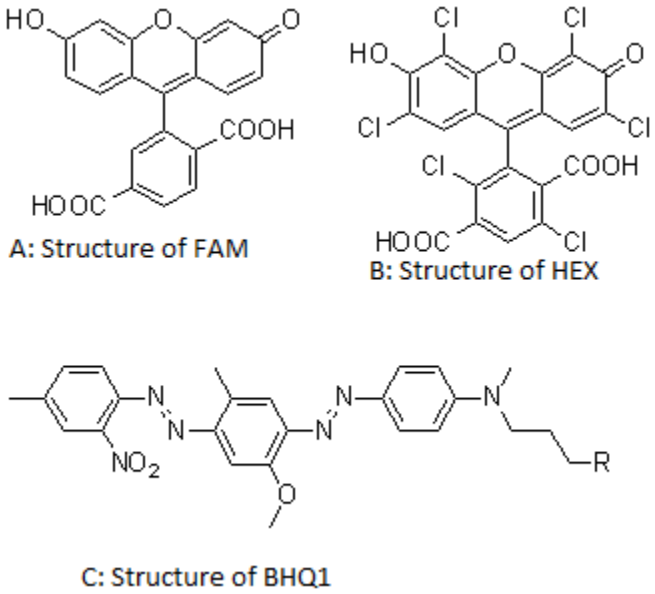

Figure 2.7: Strucure of FAM, HEX, and BHQ1
Part A shows the structure of FAM; Part B shows the structure of HEX; Part C shows the structure of BHQ1.

Dye aggregation follows two primary patterns: J- and H-aggregation8. In J-aggregation, supramolecular self-organization occurs through the stacking of dye molecules. This stacking can take many shapes, but there is a principle axis upon which the self-organization occurs (Jelley, 1936). The stacking occurs roughly in a brickstone manner according to Vacha, but more studies are needed to confirm this as multiple stacking angles are observed (Vacha et al., 2001). The stacking occurs in such a way that the absorption is bathochromically shifted, with a marked increase in fluorescence (Johansson et al., 2002). Much of what is known about J-aggregates is due to the work of E. E. Jelley and G. Scheibe with pseudoisocyanine chloride (PIC) and subsequent research (Jelley, 1936, Jelley, 1937, Würthner et al., 2011). J-aggregates are present in qPCR techniques other than MBs, specifically

involving SYBR Green I. This dye binds to the minor groove in DNA, thereby forming dye aggregates through stacking. It has been determined that these dye aggregates are J-aggregates, as the absorption pattern is bathochromically shifted and fluorescence greatly increases (Wang et al., 2000).

H-aggregates are of specific interest with regards to MBs. These aggregates also form due to supramolecular self-organization through stacking, but do so with a different orientation than J-aggregates, primarily through the angle of stacking (Berlepsch and Böttcher, 2012, Emerson et al., 1967). A diagrammatic representation of this is shown in Figure 2.8. An additional difference between H- and J-aggregates is the number of molecules involved in the supramolecular structure: J-aggregates tend to have a much larger supramolecular structure than do H-aggregates.

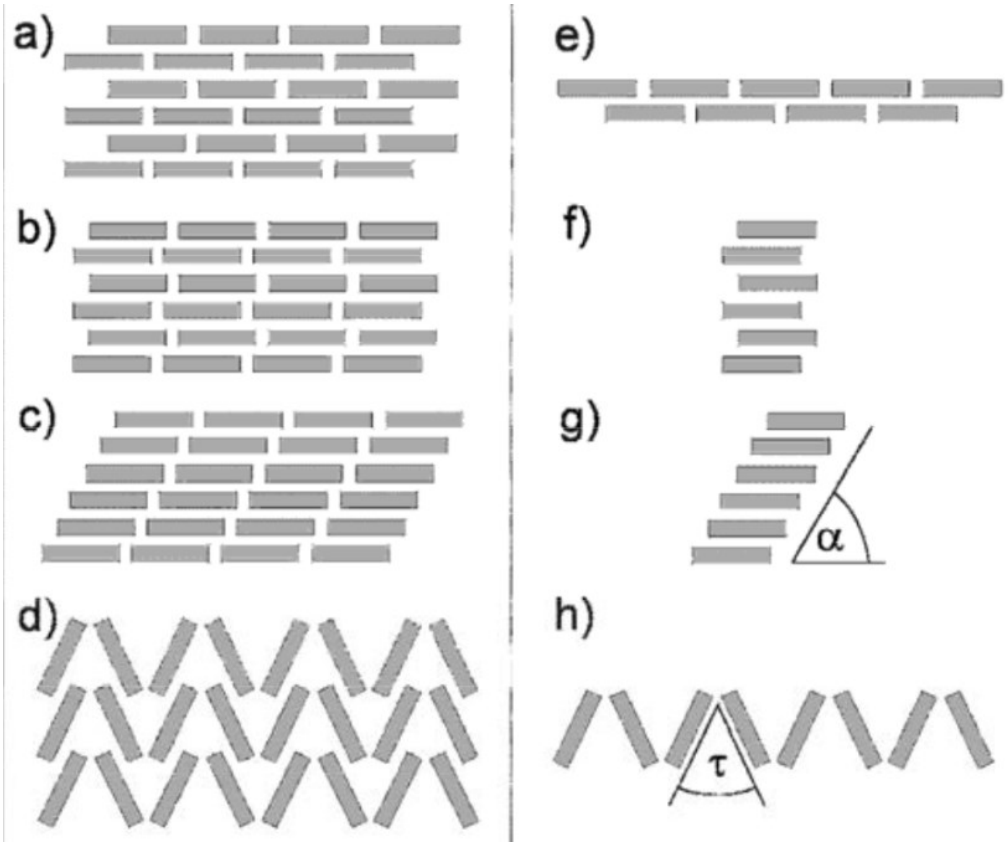

Figure 2.8: 2D and Quasi-1D assemblies of dye molecules for H and J aggregates Conceivable arrangements of dye molecules in two-dimensional (a–d) and quasi one-dimensional assemblies (e–h).
Adapted from Berlepsch, H. v., and Böttcher, C. (2012) The Morphologies of Molecular Cyanine Dye Aggregates as Revealed by Cryogenic Transmission Electron Microscopy, In J-aggregates *(Kobayashi, T., Ed.), World Scientific Publishing Company, Singapore.*

Dye aggregates have markedly different electronic properties arising from their geometric orientation. The geometric orientation of the dyes allows the interaction of dipole moments, producing "nonfluorescent heterodimers with their own distinct absorption spectra," (Johansson et al., 2002). An analogue of H-aggregates occurs in MB DAPs resulting in fluorescence quenching. These heterodimers of fluorophore and nonfluorescent quenching dyes have minimal fluorescence in spite of the sometimes limited spectral overlap between the fluorescence spectrum of the reporter and absorbance spectrum of the quencher (Johansson, 2006). It is known that static quenching dominates the quenching mechanisms in MBs (Antony and Subramaniam, 2002, Carracedo, 2005, Didenko, 2006, Emerson et al., 1967, Johansson, 2006, Johansson et al., 2002, Livak et al., 1995, Marras et al., 2002). A static quenching mechanism is made more feasible through the linkages between the dye and oligonucleotide, limiting the rotation and keeping the orientation of the DAP constant.

2.9.2: Dynamic quenching: FRET

While static quenching is the primary and most efficient mechanism of quenching in MBs, FRET also contributes. This is a distance- and overlap-dependent quenching mechanism. The distance dependence is shown in Equation 2.3.

$$k_T(R) = \frac{Q_D \kappa^2}{\tau_D r^6} \cdot \left(\frac{9000 \cdot ln\, 10}{128\pi^5 N n^4}\right) \cdot J(\lambda)$$

Equation 2.3: Dependence of FRET rate on distance
The terms are as follows: $k_t(R)$ is the rate of energy transfer with respect to distance; Q_D is the quantum yield of the donor in the absence of acceptor; n is the refractive index (typically assumed to be 1.4 for biomolecules in solution); κ^2 is a factor describing the relative orientation of the transition dipoles of the DAP (typically assumed to be 2/3); τ_d is the lifetime of the donor in the absence of acceptor; r is the distance between the DAP; N is Avogadro's number; and J is the overlap integral described in Equation 2.4 (Lakowicz, 2013). Equation 2.3 is also used in calculating efficiency in Equation 2.5.

$$J(\lambda) = \int_0^\infty F_D(\lambda)\varepsilon(\lambda)\lambda^4 d\lambda$$

Equation 2.4: Overlap integral between fluorescence of donor and absorbance of acceptor
The terms are as follows: J(λ) is the overlap integral as a function of wavelength, and is normally in units M^{-1} cm^{-1} nm^4; $F_D(\lambda)$ is the corrected fluorescence intensity of the donor in the wavelength range λ to λ + Δλ (area under the curve) normalized to unity, and is dimensionless.; ε(λ) is the extinction coefficient of the acceptor at λ, normally in units of M^{-1} cm^{-1}; λ is wavelength of light in nm.

$$E = k_T \tau_M$$

Equation 2.5: Efficiency of FRET with respect to energy transfer rate
The terms are as follows: E is the efficiency as a decimal; k_T is the rate of energy transfer between the DAP; τ_M is the fluorescence lifetime of the donor in the presence of acceptor.

It is important to note that the reader is not expected to understand the arithmetic manipulations which go into solving for the rate of energy transfer. Rather, the reader must understand the importance of each of the terms. The orientation of the DAP (κ^2), the spectral overlap between the emission of the donor and absorbance of the acceptor (J), the quantum yield of the donor and extinction coefficient of the acceptor (Q_D and ε, respectively) are of great importance when

designing a MB and choosing compatible donors and acceptors. Of specific note is the distance between the DAP.

The distance term typically dominates the efficiency and rate of energy transfer (Lakowicz, 2013). According to Clegg, "FRET is the physical process by which energy is transferred nonradiatively from an excited molecular chromophore to another chromophore by means of intermolecular long-range dipole-dipole coupling," (Clegg, 1996). In FRET, no photons are emitted from the donor chromophore. While Clegg states that this is intermolecular coupling, perhaps a more apt term may be interchromophore coupling. S.A.E. Marras et al. have shown that this process occurs in MBs, which are considered a single macromolecule (Marras et al., 2002, Tyagi et al., 1998, Tyagi and Kramer, 1996). The two chromophore moieties are held in close proximity in the closed conformation of the MB, but the absorption and emission patterns of the free moieties are unchanged upon linkage to the MB (Tyagi and Kramer, 1996).

This process occurs over a distance typically between 10-100 Å (1-10 nm), but most Förster distances (the distance at which quenching is 50% effective through FRET alone) are in the range of 2-6 nm (Clegg, 1996). A distance-based justification for FRET can be found in the distance between DNA strands: the typical distance between sugar-phosphate backbones is 2.2-2.6 nm (Mandelkern et al., 1981).

2.9.3: Summary of Signal Generation and Quenching

The previous sections explained how signal is quenched and prevented throughout the assay. While FRET is commonly cited as the mechanism for quenching in MBs, contact quenching tends to dominate given the distances and orientation involved. Upon hybridization to a target strand, the beacon will extend

and separate the DAP located on the termini. The average distance along the sugar-phosphate backbone of each nucleotide is 0.33 nm, meaning the average molecular beacon (at 35 nt) produces roughly 12 nm separation between the DAP, preventing static quenching and much longer than a typical Förster distance. This allows signal generation in the presence of a target strand. In the presence of a nontarget strand differing by only a single nucleotide, some hybridization will occur, potentially separating the DAP slightly; this slight separation is likely still within the Förster distance, allowing for signal quenching, albeit at a reduced efficiency.

Between static quenching (and to a lesser degree, FRET) in a closed conformation MB, and FRET in a random-coil conformation MB, fluorescent results can discriminate between nontarget and target presence. An open conformation MB results in generation of fluorescence, and this signal—or lack thereof—provides the basis for allelic discrimination.

2.10: Allelic discrimination based on fluorescent results

2.10.1: Allelic discrimination by kinetic results

The primary method of allelic discrimination is the presence of an exponential growth phase of signal followed by linear growth, as shown in Figure 2.6, parts B and C. This exponential growth phase corresponds to the expected second-order kinetics of MB hybridization, first order with respect to both MB and target concentrations.

Assays are designed to detect fluorescence at specific wavelengths corresponding to specific fluorophores. These specific wavelengths of detection are also called channels. An assay for a single locus with a heterozygous sample will show exponential then linear growth on two channels, corresponding to each MB,

as each MB has a different conjugated fluorophore. This is due to the presence of both alleles (Kostrikis et al., 1998a). A homozygous sample will show exponential then linear growth on only one of the two channels investigated, however. An example of this is shown as Figure 2.9, Figure 2.10, and Figure 2.11.

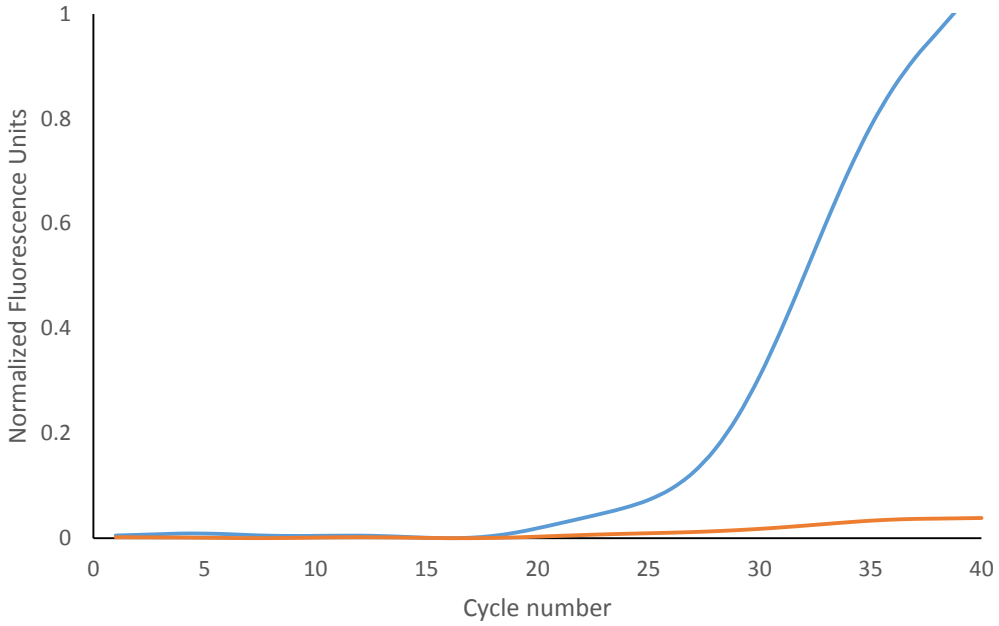

Figure 2.9: Example 1 of homozygous allelic discrimination
Blue line represents T-targeted probe, orange line represents C-targeted probe. Note the exponential and subsequent linear growth on only the T channel.

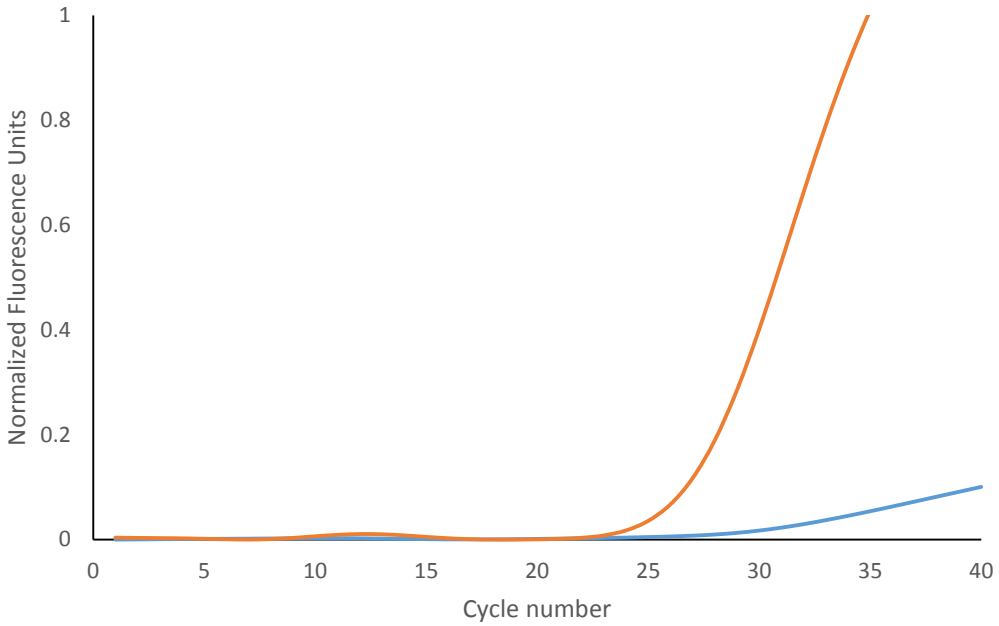

Figure 2.10: Example 2 of homozygous allelic discrimination
Blue line represents T-targeted probe, orange line represents C-targeted probe. Note the exponential and subsequent linear growth on only the T channel.

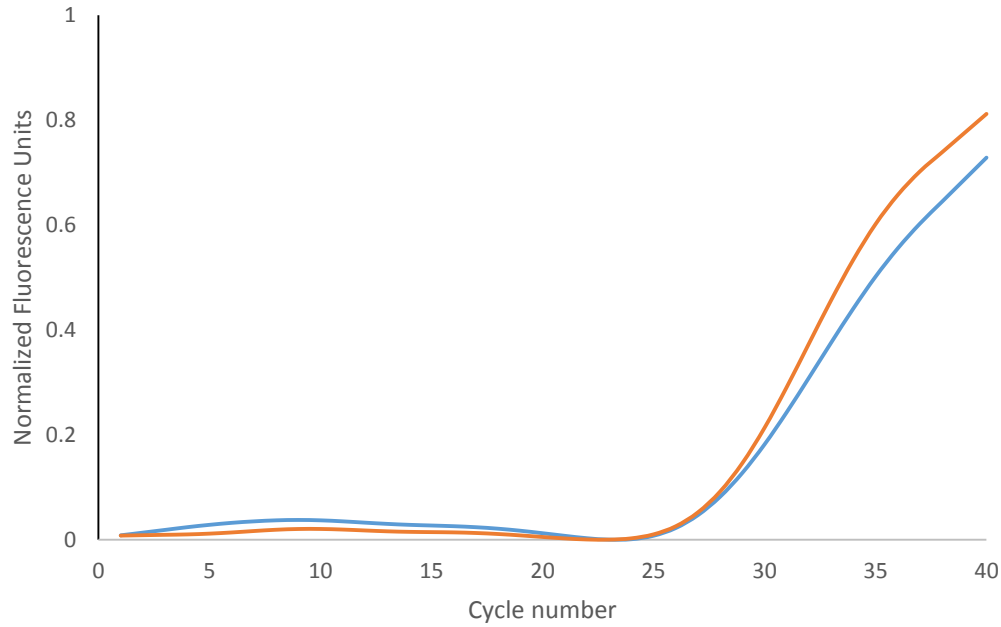

Figure 2.11: Example of heterozygous allelic discrimination
Blue line represents T-targeted probe, orange line represents C-targeted probe. Note the exponential and subsequent linear growth on only the T channel.

As stated previously in Section 2.3: Introduction to qPCR: terminology and mechanism, any qualitative PCR experiment such as this is inherently quantitative. The quantification alleleis determined by exponential growth begins. This point, also known as the cycle threshold (C_T) defines whether the exponential growth is significant or an artifact. The determination of this limit is assay-specific; a series dilution should be tested for the limit of detection (LOD) and limit of quantitation (LOQ) under all sampling conditions, and the LOD be used as the benchmark for the determination of artifacts. The C_T value is used to quantify DNA in solution by comparing to a calibration curve. This curve is formed using a serial dilution of known concentrations of DNA, amplified at the same locus with signal generated in the same way as the intended sample. Given the geometric growth nature of PCR, differences in the C_T have a logarithmic, not direct, relationship to the initial quantity of DNA in solution.

If possible, preliminary quantification of DNA in a sample is of value, as initial quantity of DNA has an inverse correlation to the C_T value. If the initial DNA concentration is lower than the LOD, the sample should be concentrated before amplification. If the sample has a significantly lower C_T than expected for its concentration, replicate assays should be considered to ensure confidence in the results. No template controls (NTCs) are also of extreme importance for intraassay validation and contamination detection in any type of qualitative discrimination.

2.10.2: Allelic discrimination by dissociation temperature

Allelic discrimination of PCR amplicons can also be accomplished through the detection of the dissociation temperature, or T_M. This discrimination method involves amplification followed by a temperature ramp during which fluorescence data is collected. As the temperature increases, the hybrid will reach a point at

which the hydrogen bonds are no longer stable; at this point, the hybrid will separate and the stem will form. This results in a marked decrease in fluorescence, which can be seen in a first-derivative plot of fluorescence with respect to temperature. The maximum of this plot is the T_M. The MB is tested in a five-fold excess of synthetic oligonucleotide target under separate conditions: with a perfectly matching target and its other allele. Given the targets interrogated are biallelic, testing both alleles is a simple method to gather data for the hybridization stability of a perfectly matching or single nucleotide difference target. This allows the determination of the dissociation temperature (Bonnet et al., 1999, Tyagi and Kramer, 1996, Wang et al., 2009).

Once the dissociation temperature is determined, amplification of the target is carried out with an extra step after cycling: a temperature-ramping melt. Reaction vessels are heated to 95°C and held for 30 seconds to ensure full denaturation of MBs and targets, then rapidly cooled to room temperature. Ideally, nonspecific annealing of the MB will occur at this temperature. Once the reaction vessel reaches room temperature, the temperature is held constant for 30 seconds, fluorescence data collected, and the temperature increased in 0.1-1°C steps followed by a hold time and data collection, up to 95°C (Tyagi et al., 1998).

The normalized fluorescence intensity is plotted as a function of temperature. Nonspecific annealing of MBs to nontargets will dissociate at a much lower temperature than MBs annealed to matching targets, with the magnitude of the difference governed by the placement of the mismatch in the structure of the probe sequence, the specific nucleotide mismatched, and the stem sequence (El-Hajj et al., 2009, Tyagi and Kramer, 2012).

If a mismatch is present, the difference between a perfectly matched and nonspecifically annealed MB T_M will likely be greater than 3°C with the melting

of a mismatch occurring at a temperature lower than the perfectly matched sequence (El-Hajj et al., 2009, Gundry et al., 2003). A measured sample T_M higher than the baseline T_M is not indicative of anything other than stronger bonding between the MB in the presence of a longer DNA target; this is likely due to a range of factors including sterics, electrostatics, and hydrophobicity. It is not uncommon for samples to exhibit higher T_M values compared to a complementary oligonucleotide only as long as the probe sequence.

Measuring the T_M of the sample-MB hybrid is another valuable method of verification. This data allows for less ambiguity in the inherently qualitative results of MB screening. Given the lack of quantifiable data used in the screening, T_M quantification allows an increased level of confidence in the results of allelic discrimination of SNPs by MBs. This data, however, is typically uneccessary given the specificity of the probes using an optimized thermocycle profile. This data is much more important for other hybridization probes, though.

2.11: Challenges of Direct PCR

2.11.1: Introduction to Direct PCR

Direct PCR is the amplification of DNA without purification or isolation. Most forensic sample analysis and *ex vivo* DNA cloning occurs after purification from the cell matrix. This purification introduces the possibility of contamination and is a very time-intensive process, so if methods and assays can be developed which are robust to the inhibitors commonly found in samples, purification may be unnecessary. Many commercial kits exist today for the purpose of direct PCR, but none are specific to the chemistry behind MBs (Thermo-Fisher, 2015c, Thermo-Fisher, 2015b, Thermo-Fisher, 2015a, BioLine Reagents, 2015). In fact, the

Thermo-Fisher Phusion kits contain an enzyme with exonuclease activity, very much incompatible with MB technology. The use of direct PCR is becoming more popular, however, with advances in enzyme technology and robustness against inhibitors.

2.11.2: Hair samples

Extraction of DNA from hair can occur using either the hair root or shaft (Dixon et al., 2005, Müller et al., 2007, Ohhara et al., 1994, Ottens et al., 2013a, Park et al., 2008, Robertson et al., 2007, Wilson, 1997, Wilson et al., 1995). The follicle is most commonly used, as the hair shaft does not have nucleated cells, and thus no nDNA; only mtDNA is present (Scientific Working Group on Materials Analysis (SWGMAT), 2005). Since mtDNA is out of the scope of this project, only freshly extracted hairs with attached follicles were analyzed. This allows for the analysis of gDNA SNPs. Extracted hairs can have 1-750 ng of DNA per hair, while shed hair can have as little as 1-10 ng of DNA per hair (Boonen et al., 2008, Ottens et al., 2013a, Ottens et al., 2013b, Robertson et al., 2007, Wilson et al., 1995, Lee and Ladd, 2001).

The wide range of available quantity of DNA from a hair with follicle allows for DNA analysis. Direct analysis of this sample type has PCR inhibitors involved, however (Butler, 2005). Melanin is in hair shafts and cells, and contributes to the color. This melanin will be released from the cell upon lysis, along with any nDNA. Eckhart *et al.* published a study on the reversible inhibition of thermostable DNA polymerases by melanin in 2000 (Eckhart et al., 2000). Their study shows that as either amplicon size or melanin concentration increases, PCR efficiency decreases.

2.11.3: Saliva samples

Liquid saliva samples can contain 1-10 µg of DNA per mL (Lee and Ladd, 2001). Saliva does not inherently contain DNA, however; rather, the source of DNA is free buccal cells. Buccal cells are similar to hair and other skin samples in that they contain melanin—specifically, buccal cells contain pheomelanin (Caspi et al., 2014).

An added challenge of liquid saliva samples, however, is the presence of any residual food or beverage. Saliva contains enzymes involved in lipid and carbohydrate digestion (Ligtenberg and Veerman, 2014). The presence of residual food, beverage, enzymes, and their products makes for a much more complex analytical matrix. These are all capable of inhibiting the polymerase, resulting in failed amplification.

2.11.4: Buccal swab samples

Buccal swab samples can contain 100-1500 ng of DNA per swab (Lee and Ladd, 2001). These buccal swabs are typically cotton- or foam-tipped applicators (Carracedo, 2005). Buccal swabs contain similar challenges to liquid saliva samples as they have the same source. Any additives in the swab itself can further complicate the matrix by inhibiting the polymerase or hybrid formation. Some swabs contain DNA binding agents or are treated with coatings such as ethylene oxide (Cell Projects, 2016).

2.11.5: Blood samples

Whole blood can contain 20-40 µg of DNA per mL (Lee and Ladd, 2001). Blood samples provide the most challenging sample type investigated in this

project. Reticulocytes, or mature red blood cells, purge the nucleus and all nDNA before entering the bloodstream (Alberts et al., 2015). Further, these reticulocytes contain hemoglobin, giving rise to their defining red color (Berg, 2012). Heme, a cofactor for hemoglobin, is a known PCR inhibitor (Berg, 2012, Akane et al., 1994). Additionally, heme has a coordinated iron ion in its divalent state upon oxygen uptake, increasing the divalent cation concentration of the matrix and thus lowering specificity of PCR (Berg, 2012, Alberts et al., 2015). Another major consideration is the clotting pathway, which could solidify the sample (Berg, 2012). Anticoagulants could potentially prevent this, but may also interfere with cation concentrations.

2.11.6: Solutions to Direct PCR Issues

The most common solution to the issues faced with direct PCR is simply purification. In this case, however, it is not desirable to purify, as this would introduce another potential source of contamination. Digestion with Proteinase K would eliminate many issues such as melanin and hemoglobin, but would also destroy the polymerase. Primarily, this project avoids these issues through thermal denaturation of proteins during the thermolytic phase of sample preparation and dilution of inhibitors as solutions to these potential issues.

2.12: Basis of DNA profiling

2.12.1: Introduction to Profiling

DNA profiling was first introduced in 1985 by Dr. Alec Jeffreys using mini-satellite regions of the human genome (Jeffreys et al., 1985). Profiling methods rely on polymorphisms in satellite regions which can make up as much as 30% of the

human genome (Singer, 1982). These hypervariable satellite regions consist of repeats of 2-60 nucleotides, depending on whether they are mini- or micro-satellites (National Library of Medicine, 2011b, National Library of Medicine, 2011a, Fredericks, 2011). These regions are of great importance to the forensic community as they are unique to each individual, even siblings (with the exception of monozygotic twins) (Schneider, 2007, Office of Justice Programs, 2012). Both VNTR and STR profiling are outside of the scope of this project, but a brief review is provided for the foundational understandings of DNA profiling. Other methods exist such as Y-chromosome and mtDNA profiling, but as they too are outside of the scope of this project, they will not be covered.

2.12.2: VNTR Profiling

Mini-satellite profiling was the first DNA "fingerprinting" method proposed. This method, however, is now obsolete with the rise of short tandem repeat (STR) profiling centered around micro-satellites. Variable number of tandem repeats (VNTR) profiling was crucial to the development of forensic biology as it provided the basis for identification. These hypervariable regions are typically 10-60 bases long, repeating 5-50 times in tandem (Tautz, 1989, Tautz, 1993, Chambers and MacAvoy, 2000). These mini-satellite regions are normally analyzed through restriction fragment length polymorphism (RFLP) analysis or Southern blotting, and to a lesser extent PCR (Hviid et al., 1992, Demers et al., 1995, Fredericks, 2011).

2.12.3: STR Profiling

Microsatellites such as STRs make up approximately 3% of the total human genome (International Human Genome Sequencing, 2001). Unlike VNTR profiling which utilizes large fragment sizes, STR profile fragments are small enough to analyze using capillary electrophoresis after multiplex PCR (Promega Corporation, 2015). The repeat size typically used for forensic identification is 2-5 bp (Urquhart et al., 1993, Urquhart et al., 1994). This is currently considered the standard method of forensic biology identification and is the only method currently accepted by the US Federal Bureau of Investigation for legal proceedings, although mtDNA and Y STR profiling are acceptable for missing persons investigations (Federal Bureau of Investigation, 2015).

This method of profiling involves the amplification of 13 loci simultaneously, and each locus has at least 15 known alleles with some loci having as many as 80 alleles (Butler, 2005, Federal Bureau of Investigation, 2015, Promega Corporation, 2015). Forensic STRs can be categorized in three ways: simple, compound, or complex repeats. Simple repeats are units of identical length and sequence, with the exception of the final repeat. The final repeat may be only a fractional repeat, known as a non-consensus allele or microvariant, and is possible even in simple STRs (Butler, 2005, Puers et al., 1993). An example is shown as Figure 2.12. A compound repeat consists of two or more adjacent simple repeats, with a motif such as the von Willebrand factor, 40^{th} intron (VWA) locus: $[TCTG]_n[TCTA]_m$. This is not an octanucleotide repeat, as the two tetranucleotide motifs are repeating adjacently; this is, however, considered a single microsatellite, useful for discrimination. An example is shown as Figure 2.13. Only one complex repeat is currently used for Combined DNA Index System (CODIS) identification;

these contain "several repeat blocks of variable unit length as well as variable intervening sequences," (Urquhart et al., 1993, Urquhart et al., 1994, Butler, 2005).

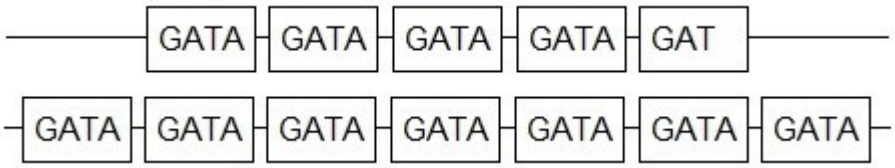

Figure 2.12: Example of Simple STR with Microvariation
The individual would have a profile of 4.3 and 7 repeats at this locus. Note the partial repeat in the top copy.

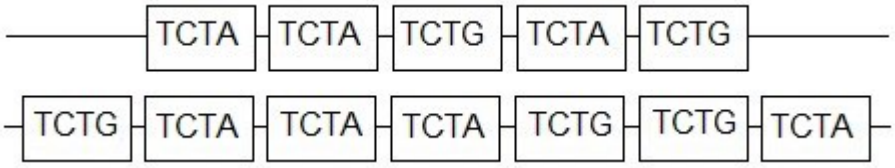

Figure 2.13: Example of Compound STR

The power of STR (and VNTR) profiling arises from the sheer number of alleles due to the variability of these regions, and the selection of independently assorting loci. Without the consideration of monozygotic twins, STR profiling commonly results in matching probabilities on the order of 1 in 10^{14} or rarer with only 13 loci chosen (Chakraborty, 1992, Chakraborty et al., 1999). Results are further narrowed using genotyping of the sex chromosomes, with different amelogenin gene sizes between X and Y chromosomes. The Y chromosome amelogenin copy contains a 6 bp deletion within intron 1 (Sullivan et al., 1993). If only one size is detected, the individual is XX (female); if two sizes are detected, the individual is XY (male) (Sullivan et al., 1993).

2.12.4: SNP Profiling

Single nucleotide polymorphism (SNP) profiling is based on the analysis of biallelic markers which are normally detected through sequence analysis methods such as capillary electrophoresis, Sanger sequencing, or pyrosequencing. The biallelic nature of SNPs is both an advantage and disadvantage to SNP profiling. The reduced number of alleles allows for faster screening, but the power of discrimination is greatly reduced. In order to obtain the match probabilities inherent in current STR analysis, between 25-50 SNP markers would need to be examined (Chakraborty et al., 1999, Gill, 2001). A study by Gill *et al.* in 2004 suggested between 50-100 SNPs would be required to match the power of discrimination and mixture resolution capabilities of the current STR methods (Gill et al., 2004). Recently, however, SNPs have been investigated as potential methods for familial and identification analyses (Amorim and Pereira, 2005, Costa et al., Huijsmans et al., 2007, Kidd et al., 2006). Current STR profiling methods require sample purification, PCR amplification, and subsequent capillary electrophoretic detection. These results are then interpreted by a forensic biologist to build a profile. A SNP panel using hybridization probes, though, would eliminate the capillary electrophoresis step; these reactions can be monitored in real-time and results obtained at the end of amplification. In the case of disaster victim identification (DVI), faster results and smaller amplicons are extremely favorable given the number of samples and potential for degradation (Ziętkiewicz et al., 2012). Additionally, the power of discrimination necessary for legal proceedings such as trials may not be necessary for DVI. Panels as small as 19-21 SNPs have proven to theoretically have sufficient probability of match values (Dixon et al., 2005, Kidd et al., 2006). A table comparing the advantages and disadvantages of the three profiling methods discussed is presented in Appendix B: Comparison of Selected

Profiling Methods, adapted from Fredericks (Fredericks, 2011). Insertion-deletion (InDel) profiling is virtually identical in practice to SNP profiling, with the exception of the genetic polymorphism tested. InDel profiling analyzes the sequence for a marker insertion or deletion, typically 4-8 nt in length (Pereira et al., 2010). Table 2.2 compares the differences in SNP and STR profiling.

Table 2.2: Comparison of STR and SNP markers

Characteristics	STRs	SNPs
Occurrence in genome	~1 in every 15 kb	~1 in every 1 kb
General informativeness	High	Low; 20-30% as informative as STRs
Marker type	2-5 nt repeat markers	Single nt difference
Alleles per marker	Typically >5	Typically 2
Detection methods	Gel/capillary electrophoresis	Sequence analysis, microchip assay
Multiplex capability	>10 markers	1000s using microchip assays
Major advantage	Power of discrimination	Small amplicons, higher success with degraded DNA

2.13: Summary of literature review

The advent of PCR and its associated technologies such as qPCR and hybridization probes has allowed great strides in the advancement of forensic biology. MBs are highly specific hybridization probes which allow for the detection of SNPs in a given sequence, requiring an amplicon only as long as the sum of the length of the primers and MB itself.

These probes are extremely useful when working with degraded DNA samples and result in simple, qualitative results. These results can also be made quantitative, depending on the needs of the end user. They generate signal in the presence of an exact target after thermodynamically favorable conformational changes. Under optimal conditions, MBs will provide specific and selective signal generation through these conformational changes, allowing allelic discrimination by either kinetic results or dissociation temperature measurements.

The use of MBs under direct PCR conditions for forensic applications is not well studied. A thorough literature review did not return any publications on this topic. There are currently many challenges to forensic applications of MBs in a direct PCR assay, with each sample presenting unique, inherent problems which must be solved.

The end goal of this project is to develop a working, proof-in-concept MB assay with minimal purification of the sample. This will help eliminate a possibility of contamination and provide faster collection-to-result times, ultimately resulting in a viable assay method to reduce the DNA backlog currently experienced by many law enforcement agencies.

CHAPTER 3: METHODS AND MATERIALS

CHAPTER 3: MATERIALS AND METHODS

3.1: Introduction

This chapter describes in detail the methodology and rationale of the techniques used for the project. Common forensic samples include saliva, buccal swabs, hair bulbs, and blood. All samples were subjected to the same thermal lysis protocol. In order to fulfill the requirements set out in the goals and scope (Section 1.5:), samples were obtained from living, consenting adults age 18-35. As this is a proof-in-concept, all samples were prepared from freshly extracted tissue or fluid. This ensures the matrix effects involved are as reduced as possible. In the future, studies will need to be done to ensure the viability and robustness of the protocols developed herein with various levels of degradation or preservatives.

3.2: Materials and equipment used

Many probe master mixes were tested. These contained different enzymes, cation concentrations, etc., and were tested to find the most sensitive, reliable, and robust kit for the sample types studied. BioLine SensiFAST™ Probes Master (BioLine, London, UK), BioLine SensiFAST SYBR Green I (BioLine, London, UK), LightCycler™ 480 Probes Master (Roche Applied Science, Penzberg, Germany), Phusion Human Specimen Direct PCR Kit (ThermoFisher Scientific, Massachusetts, USA) and Phusion Blood Diret PCR Kit (ThermoFisher Scientific, Massachusetts, USA) were used.

All oligonucleotides and MBs were custom ordered through Sigma-Aldrich (Missouri, USA). Oligonucleotides and MBs arrived as lyophilized pellets and were resuspended as a master stock at 100 μM. Aliquots of this master stock were then

diluted to a working stock of 10 μM. Specific sequences can be found in **Error! Reference source not found.**. Control gDNA samples were tested to find a heterozygote for the allele used as a model, and that heterozygote used throughout the project as a positive control. A heterozygous positive control is preferable, as it allows verification of a working assay for both MBs, rather than using a homozygote which would only give signal response for one MB in the assay. The control gDNA (Promega, Wisconsin, USA) for each positive control was 20 ng/reaction. A negative control was also employed using nuclease-free water (NFW) (Roche Applied Science, Penzberg, Germany) in place of gDNA. Equipment was routinely subjected to a CL-1000 UV crosslinker (UVP, California, USA) to prevent contamination.

Screening of participants prior to direct PCR amplification was necessary. Participant DNA was extracted using an Isolate II Genomic DNA Kit (BioLine, London, UK). This extraction was done according to the manufacturer's instruction without deviation. Extracted DNA was used to obtain baseline genotypes for this allele to ensure accuracy in subsequent experiments. All incubations for DNA isolation were carried out using an Accublock dry bath (Labnet, New Jersey, USA).

The Nanodrop 2000 (ThermoFisher Scientific, Massachusetts, USA) was used to approximate DNA quality and quantity. It is a microscale UV-Visible spectrophotometer, calibrated for aliquots as small as 0.5 μL. Using preconfigured and regularly calibrated standard curves, the Nanodrop 2000 is able to approximate dsDNA concentration in nanoscale samples and give information regarding purity. The Nanodrop 2000 was used for all DNA stock solution preparations. This purity information is generally given as a ratio of absorbances at 260 nm to 280 nm. Since DNA bases absorb at 260 nm, and proteins—the most common DNA contaminant—absorb at 280 nm, a ratio of the two absorbance values can give

relative amounts of each in a solution. However, since this is a spectroscopic method without derivatization, it is nonspecific. Any molecule absorbing 260 nm light will cause an increase in signal, whether it is DNA or not. While nonspecific, it is still useful in determining the potential viability of the sample. Most derivatization methods, such as SYBR quantification, are prohibitively expensive to test on unverified methods of DNA preparation; the Nanodrop 2000, conversely, requires only a micropipette and its included hardware in addition to the sample. All Nanodrop measurements were performed with the buffer used in sample preparation as a blank and according to manufacturer's guidelines. The Nanodrop 2000 has a published accuracy of 2% and precision of 0.002 AU.

3.3: Standard experimental concentrations and cycling

3.3.1: Assays without MBs

A master mix was prepared using 10.00 µL aliquot of the Roche LightCycler 480 Probes Master, 1.00 µL aliquot of each working stock for FOR100 and REV100 primers, and 7.00 µL aliquot of nuclease free water (NFW) per reaction, plus one additional reaction to account for pipetting loss, was placed into a 0.2 mL PCR tube. In the case of an assay with 6 samples, 70.0 µL of Probes Master, 7.00 µL of each primer, and 49.0 µL NFW would be used to prepare this master mix. This mixture was then vortexed. A 19.00 µL aliquot of this mixture was placed into each of the 0.2 µL PCR tubes, labelled appropriately. A 1.000 µL aliquot of NFW was placed in the – tube, and a 1.000 µL aliquot of control gDNA into the + tube. Resulting concentrations were as follows: Roche LightCycler Probes Master: 1X; Primers, each: 0.5 µM; control DNA: 20 ng/reaction (1.00 µg/mL). Amplification was performed using a C1000 Touch thermocycler

(BioRad, California, USA). A 95°C, 5 minute initial hold followed by 35 cycles of a 30 second hold at 95°C, then 48°C, then 72°C.

3.3.2: Assays with MBs

A master mix was prepared using 10.00 µL aliquot of the Probes Master, 1.00 µL aliquot of each 10 µM working stock for FOR100 and REV100 primers, 1.000 µL of an equimolar 10 µM MB mix, and 6.00 µL aliquot of nuclease free water (NFW) per reaction, plus one additional reaction to account for pipetting loss, was placed into a 0.2 mL PCR tube. In the case of an assay with 6 samples, 70.0 µL of Probes Master, 7.00 µL of each primer, 7.00 µL MB mixture, and 42.0 µL NFW would be used to prepare this master mix. This mixture was then vortexed. A 19.00 µL aliquot of this mixture was placed into each of the 0.2 µL PCR tubes, labelled appropriately. A 1.000 µL aliquot of NFW was placed in the – tube, and a 1.000 µL aliquot of control gDNA into the + tube. Resulting concentrations were as follows: Probes Master: 1X; Primers, each: 0.5 µM; control DNA: 20 ng/reaction (1.00 µg/mL). Amplification was performed using a Rotor Gene Q (Qiagen, Hilden, Germany). The thermal cycle followed a 95°C, 5 minute initial hold followed by 35-45 cycles of a 30 second hold at 95°C, then 48°C, then 72°C. Fluorescent data was acquired at the end of the annealing phase for both 510 and 555 nm channels and plotted as raw counts against cycle number, and then normalized for analysis. Software-assisted C_T values were determined for comparison.

3.3.3: Gel preparation

Agarose (Molecular grade, BioLine, London, UK) and a 0.089 M tris base, 0.089 M boric acid, 0.002 M EDTA (1X TBE) buffer (BioBasic, Ontario, Canada)

were used to make 2% $^{wt}/_{v}$ gels for electrophoresis. A GeneMate 120 g analytical balance (BioExpress, Pennsylvania, USA) was used to measure agarose before addition to the buffer. Gels were prepared by boiling using a 700 W microwave oven (Rival, Florida, USA) until optically clear, followed by the addition of ethidium bromide (EtBr) stain (BioBasic, Ontario, Canada) to a final EtBr concentration of 0.5 µg/mL in the gel.

A typical gel is prepared with 50 mL of 1X TBE, with 1.000 g agarose, followed by boiling. Upon cooling to slightly above room temperature, 2.500 µL EtBr stain is added and mixed thoroughly. The mixture is then poured into a casting plate with a comb insert to form the wells. Electrophoresis was performed using a Mini-300 gel electrophoresis system (Major Science, California, USA). Imaging was performed using a BioDoc-It UV Imaging system (UVP, California, USA). All gel electrophoresis procedures used the HyperLadder™ 50 bp molecular weight marker system (BioLine, London, UK) for comparison of fragment sizes, with an example image of the bands and their corresponding weights as Figure 3.1. Electrophoresis was performed for 75 minutes at 120 V. Images for gel electrophoresis of primer validation are presented in Section 4.2.1: . While the optimal annealing temperature for the primer set is much higher than 48°C, this is the operating temperature with molecular beacons and was used to check for nonspecific priming under operating conditions.

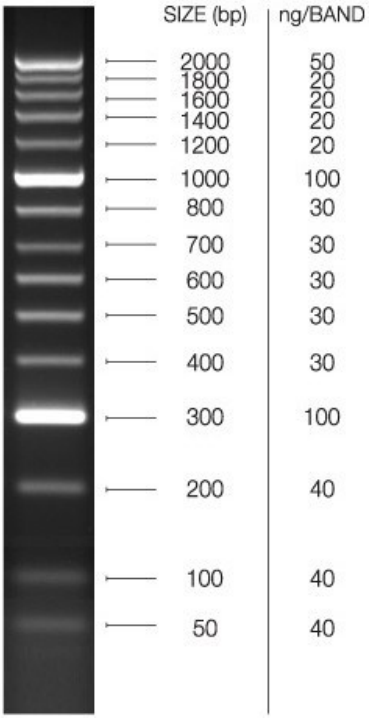

Figure 3.1: HyperLadder 50 bp run on 1.5% TBE/agarose gel, bands marked Image adapted from BioLine HyperLadder 50 bp product information web page, URL: http://www.bioline.com/us/hyperladder-50bp.html.

Table 3.1: PCR Conditions, Assays without MBs

	Tube	
Component	+	-
Probes Master	10.00 µL	10.00 µL
FOR100	1.000 µL	1.000 µL
REV100	1.000 µL	1.000 µL
Control gDNA	1.000 µL	N/A
NFW	7.00 µL	8.00 µL

Table 3.2: PCR Conditions, Assays with MBs

	Tube	
Component	+	-
Probes Master	10.00 µL	10.00 µL
FOR100	1.000 µL	1.000 µL
REV100	1.000 µL	1.000 µL
Control gDNA	1.000 µL	N/A
NFW	7.00 µL	8.00 µL
MB Mixture	1.000 µL	1.000 µL

3.4: Phase I: Verification of amplification and detection

3.4.1: Verification of FOR100 and REV100 primers

The MBs and an original set of known, amplifying primers were chosen from a previous research project by Dr. Jamie Fredericks (Fredericks, 2014). These primers amplify a 100 bp fragment around a SNP located on human chromosome 8, on an intron region of the *DLGAP2* gene with refSNP ID: RS763869. Neither allele has any known clinical significance, and as such has minimal ethical issues surrounding its genotyping. Knowledge of the genotype for this site has no clinical implications, making it ideal for research. There are no other known, common variations in the selected amplicon.

BioEdit was used to initially align the primer, MBs and target sequence. This is shown as Figure 3.2: Alignment using BioEdit, including FOR100, REV100, C probe, and T probe. The web-based nBLAST was used to confirm specificity in priming. The expect value is used to determine the likelihood of a random-chance match; the smaller the value, the less likely the match is due to

random chance. The expect value should be much larger for nontargets than it is for targets, or nonspecific annealing can occur. The expect values are as follows: FOR100, 0.002; REV100, 0.0005. This denotes high specificity in the primers for their targets.

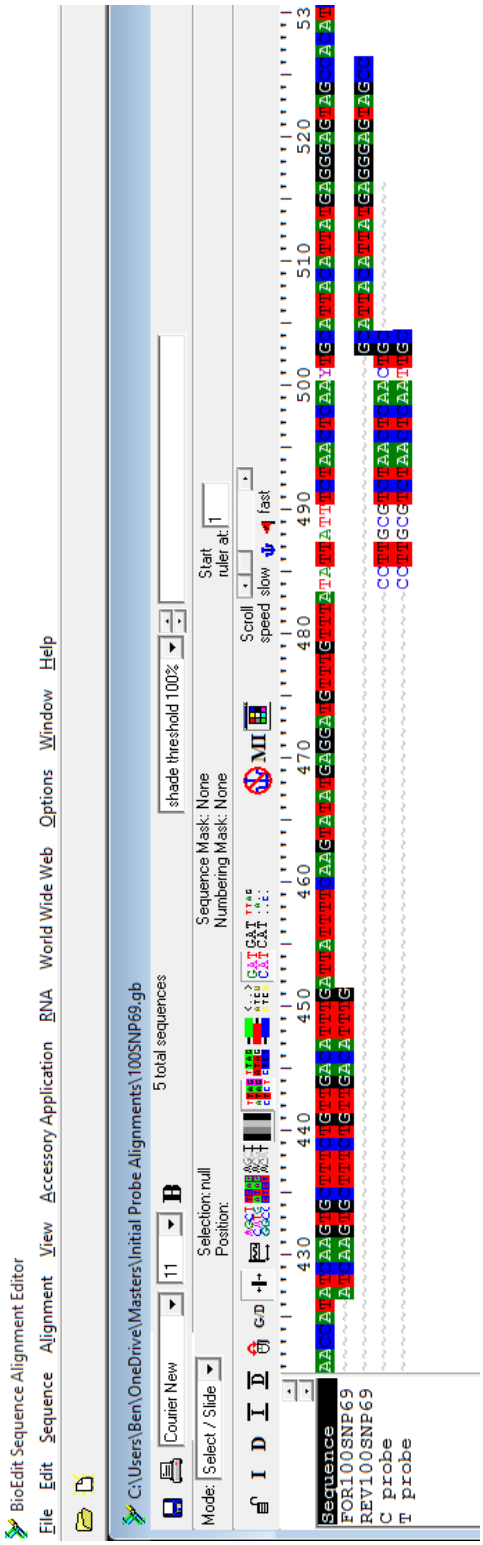

Figure 3.2: Alignment using BioEdit, including FOR100, REV100, C probe, and T probe

Letters represent one letter abbreviations for DNA bases. Bases with colored blocks match the target sequence; letters without blocks do not match target sequence.

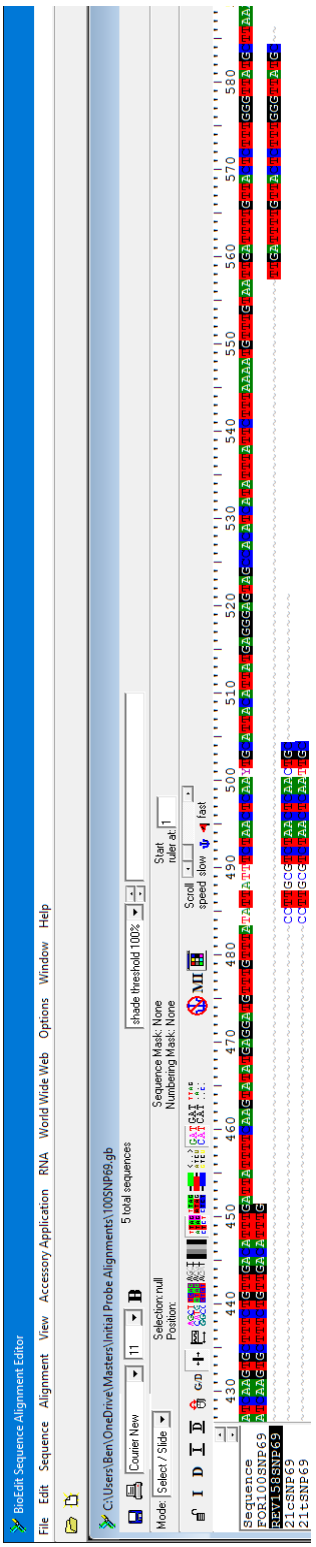

Figure 3.3: Alignment using BioEdit, including FOR100, REV158, C probe, and T probe.
Letters represent one letter abbreviations for DNA bases. Bases with colored blocks match the target sequence; letters without blocks do not match target sequence.

3.4.2: Thermal profile of MB probes

Complementary oligonucleotides and the MBs were ordered from Sigma Life Science (Missouri, USA). The complementary oligonucleotides were 50 bp, with one an exact match and another matching the opposite allele. These were ordered as 50 bp of the target sequence in order to fully cover the stem and simulate as best as possible the performance when in contact with the amplicon, as the new MB has a total length of 48 bp, stem and probe included. The thermal profile was generated using 1.000 μM MB, 10.00 μM complementary (or mismatch) strand, in 8.00 mM Tris-HCl, 280 nM $MgCl_2$, pH 8.0. The thermal profile used was a 5 minute, 95°C hold, followed by a 2 minute, 55°C hold, followed by a 1°C/min ramp from 25°C to 90°C. Fluorescence data was obtained for each MB and plotted as both raw counts against temperature and a first-derivative plot of fluorescence change with respect to temperature.

3.4.3: Verification of MB assay

The MB assays were verified using a standard 90 second cycling profile for at least 40 cycles. This verification first studied each MB probe separately in order to limit any interferences with the other probe. The T probe was conjugated with a FAM fluorophore, with λ_{max} at 520 nm; the C probe was conjugated with a HEX fluorophore, with λ_{max} at 556 nm (Biosearch Technologies, 2016). This experiment was performed to show the performance of each probe at both target wavelengths of fluorescence independently of the other MB, and validates the MBs are performing as desired. This assay followed the standard conditions for assays with MBs, with a modification. The MB set was not mixed; each tube had only of the

two MBs. The Roche LightCycler 480 Probes Master and REV100 primer were used in this assay.

After the MBs were validated separately, the MBs were validated in a multiplex format. It is important to note that this is not a true PCR multiplex which amplifies multiple loci in one reaction; rather, this is a probe multiplex, which uses multiple probes in the same reaction. Interferences and spectral effects are possible, and it was necessary to ensure valid performance of the probes when used in conjunction. This required adding both MBs in equal concentration to a single reaction mixture, then collecting the fluorescence data for each reaction on both channels. This experiment was performed to show the performance of each probe at both target wavelengths of fluorescence, and validates the MBs are performing as desired. This assay followed the standard conditions for assays with MBs, with no modification. The Roche LightCycler 480 Probes Master was used in this assay, with the REV100 primer.

3.4.4: Screening of genotypes

Assays were performed under standard assay conditions with MBs, with a modification: extracted DNA used as a sample in place of control gDNA. A control gDNA sample was used as a positive control, and a NFW sample used as a negative control. All other conditions were the same. The Roche LightCycler 480 Probes Master was used in this assay.

3.5: Phase II: Comparison and selection of kits

3.5.1: ThermoScientific Phusion kits

The thermocycle was set to the manufacturer's specifications. All concentrations were the same as standard MB assay conditions.

3.5.2: Roche LightCycler 480 and BioLine SensiFAST comparison

Assay conditions were the same as standard MB assays. Both Probes Master kits were used in this assay in separate reactions. Additionally, control gDNA concentrations varied using a series dilution. The reactions used 20, 10, 5, 2.5, and 0.25 ng/reaction for each the Roche and BioLine kits. The REV100 primer was used for this assay.

3.6: Phase III: Direct MB-PCR assay development

3.6.1: Sample preparation

What follows is a complete, detailed account of the sample preparation methods for the four sample types studied, along with rationale for the chosen methodologies. For convenience, Appendix C: Summary of Sample Preparation has a summary of the sample types and necessary materials.

3.6.1.1: Hair sample preparation

Hair samples were extracted using a pair of stainless steel dissection forceps. Typically, five to ten hairs were extracted from each pull. This ensured at least one hair would have the entirety of the follicle still attached.

After extraction, a suitable hair with attached follicle was chosen and trimmed approximately 1-2 mm above the follicle. The follicle was then placed at the bottom of a 0.2 mL tube. A 10.00 µL aliquot of the modified TE buffer described in Section 0 was added to each hair sample before thermal lysis. The buffer allows collection of the DNA and a suitable means by which to transfer the collected DNA to the reaction. No post-lytic treatment was required.

3.6.1.2: Saliva sample preparation

Saliva samples were obtained after a 30 second rinse using a cold 0.9% $^{w}/_{v}$ saline wash. The wash was expelled before sample collection. Participants were asked to expel saliva into a 1.5 mL microfuge tube. Approximately 1 mL of saliva was collected from each participant. Microfuge tubes were then centrifuged at 2000 G using a Prism Mini centrifuge (Labnet, New Jersey, USA) for approximately 10 seconds to separate any colloids in the saliva. The colloidal solids are primarily composed of proteins, which can interfere with PCR. For each sample, a 100.0 µL aliquot of supernatant was transferred to a 0.2 mL tube for thermal lysis. Once lysis was complete, this sample type required centrifugation at 2000 G using a Prism Mini centrifuge for approximately 10 seconds to again separate the colloidal solids from the saliva. Aliquots of the supernatant were used as the sample.

3.6.1.3: Buccal swab sample preparation

Buccal swab samples were obtained after saliva samples. Participants had already rinsed with a 0.9% $^{w}/_{v}$ saline wash prior to saliva collection. The buccal swab collection was done after saliva collection as the swab is abrasive and able to extract buccal cells even after the expulsion of saliva, containing loose buccal cells.

Had the buccal cell collection been done first, the resulting saliva would have had a greatly reduced concentration of loose buccal cells, and thus a lower concentration of DNA.

Participants were given a foam-tipped applicator and asked to brush the buccal cavity—the area between the gums and cheek—for at least 30 seconds to ensure cell collection. The swab was then placed in a 0.2 mL tube with 20.00 µL of the modified TE buffer described in Section 0, and the swab's handle removed. The buffer served as a means for collection and transfer of DNA after thermal lysis. Once the handle was removed from the swab, the sample was ready for thermal lysis. No post-lytic treatment was required. The buccal swab collection is done last to ensure the resulting saliva has a sufficient cell concentration for DNA extraction.

3.6.1.4: Blood sample preparation

A single-use, disposable, 20-gauge hypodermic needle (BD PrecisionGlide, Franklin Lakes, New Jersey) or 16-gauge disposable, single use, self-contained Unistik 3 Normal diabetic lancet (Owen Mumford, Georgia, USA) was used for blood extraction. Once the skin was pierced using either a hypodermic needle or lancet, a 2.000 µL sample was obtained using a micropipette and diluted to 10.00 µL with the modified TE buffer described in Section 0**Error! Not a valid bookmark self-reference.**—all in a 0.2 mL tube for thermal lysis. Once lysis was complete, this sample type required centrifugation at 2000 G using a Prism Mini centrifuge for approximately 10 seconds to separate the red blood cells from the plasma. The plasma still contained the leukocytes necessary for gDNA analysis, as the centrifuge used is not powerful enough to pull the leukocytes into their own layer. Given the methods and instrumentation, erythrocytes would interfere with

the optical signal and must be removed. These erythrocytes can be removed without reducing the quantity and quality of DNA present.

3.6.1.5: Thermal lysis

All samples were subjected to the same thermal lysis conditions. Each 0.2 mL tube was placed in a C1000 Touch thermocycler (BioRad, California, USA) at 98°C for 5 minutes, followed by a 4°C infinite hold until the samples were retrieved. The samples were held at 4°C for no longer than 10 minutes. This was especially important in the case of the blood samples, as no anticoagulants were used in this study. The 4°C hold was used to ensure a consistent time was used for thermal lysis, and thus its reproducibility.

3.6.2: Verification of amplification under direct conditions

To validate amplification, standard MB-free assay conditions were used, with a modification: direct samples as prepared above were used in place of either NFW or control gDNA as samples. The Roche LightCycler 480 kit was used in this assay as the Probes Master, with the REV100 primer.

3.6.3: Design of new primer set

During the course of the project, the negative control consistently showed heterozygosity; the signal did not rise above the limits for significance (three times the background), but was still noticeable. Given the specificity of the MBs, the likely etiology was a stem-primer hybrid forming with the reverse primer. The MBs for both alleles were identical with the exception of the central base, which was allele-specific. With an identical stem, both MBs would likely form a stem-primer

hybrid as the two were complementary. This caused the observed false heterozygosity, and a new reverse primer was designed in order to afford greater sensitivity and specificity.

An ideal primer is highly specific for the given target, has a G-C content at least 50%, and is between 15-20 bp. Given the DNA sequence surrounding the primers, the new selected site of the reverse primer had to be at least 50 bp further away to ensure a similar T_M while still preventing a stem-primer hybrid formation. The sequence contains an extremely A-T rich region adjacent to the previous primer with poly-A and poly-T regions, which if included in a primer could allow for false priming due to increased homology at other sites. This repeating region, if included, would greatly reduce the PCR efficiency. Further, since an A-T pair has fewer hydrogen bonds than a G-C pair, the T_M is lower for primers with a lower G-C content. A primer with such a low G-C content would require a much longer primer, accompanied by a longer annealing time due to the kinetic considerations. Primer REV158 was designed with the T_M of FOR100 in mind, so only one primer would need to be changed. Changing multiple primers would require additional verification assays; simply lengthening the fragment size by 50 bp does not appreciably affect hybridization of the MBs. Additionally, this primer was designed to be as close as feasible to a 150 bp fragment, allowing easy verification by gel electrophoresis using a 50 bp ladder. Primers for a MB assay should limit the amplicon size to no higher than 200 bp due to the kinetics of hybrid formation. Since the target sequence must first line up with the probe sequence, hybridization takes much longer as fragment size increases, thus lowering the likelihood of hybrid formation.

Primer design was aided by the use of BioEdit Sequence Alignment Editor (Ibis Therapeutics, California, USA), web-based Nucleotide Basic Search

Alignment Tool (nBLAST) (NCBI, Maryland, USA), and web-based Primer Basic Search Alignment Tool (PrimerBLAST) (NCBI, Maryland USA). These tools were helpful in ensuring the designed primers would not amplify outside of the target amplicon. The nBLAST and PrimerBLAST tools check for specificity against all known sequences using probability-based algorithms. Additionally, web-based primer design tools for technical data (Sigma Life Science, Missouri, USA), (Integrated DNA Technologies, Iowa, USA) were employed to ensure T_M compatibility between the FOR100 and the designed primer by using algorithms to predict the new primer's T_M.

Oligo7 software was then used to search for any secondary structures in the primers. Secondary structures (such as hairpin formation, self-dimers, and heterodimers) can prevent the primer from annealing to the target, thus preventing amplification. This program works by predicting any complementarity which could cause the formation of secondary structures, given theoretical ΔG calculations for specified conditions. Neither FOR100 nor REV100 formed any hairpins nor significant self-dimers. The primer REV158 did, however, form a self-dimer. This self-dimer can be neglected due to the number of base-base mismatches; in solution, and at the lowest temperature of the cycle (48°C), the self-dimer is not likely to form. The primer also formed a hairpin up to 21.4°C; at a temperature above this such as the annealing temperature of the cycle, the hairpin will separate and linearity will be restored. Web-based tools from Sigma Life Science and Integrated DNA Technologies were used to screen for heterodimers between a forward and reverse primer; none were observed. Oligo7 and the web-based tools were also used to screen the primers for thermal compatibility. All primers were within 0.5°C of each other, which is ideal for PCR.

3.6.4: Comparison of new primer set

Comparison assays followed standard, MB-free conditions with a modification: a positive and negative control was used for the FOR100/REV100 set, and also for the FOR100/REV158 set. The reactions with REV158 did not include an aliquot of REV100. The BioLine SensiFAST kit was used in this assay as the Probes Master.

3.6.5: Analysis of primers in real-time

Real-time analysis of the primers followed standard, MB-free assay conditions using the BioLine SensiFAST SYBR Green I Master Mix. The thermocycle settings were set to manufacturer's recommendations. One reaction used the REV100 primer, and another used the REV158 primer. All other conditions were the same.

3.6.6: Verification of direct amplification, REV158

Assay conditions were the same as in 3.6.2, using the REV158 primer instead of REV100; additionally, the BioLine SensiFAST Probes Master was used in this assay.

3.6.7: Thermocycle optimization

Thermocycle optimization followed the standard MB assay conditions with modification. The times for each step of the cycle were modified for each assay. Initially, the standard cycling conditions were used. Then, the extension time was changed. Once the extension time was optimized, the denaturation time was

optimized. Finally, the annealing time was optimized. Once a step was optimized, that optimized time was carried forward to the next assays. The BioLine SensiFAST Probes Master and REV158 primer were used in these assays.

3.6.8: Buffer development

The use of a sample collection buffer was necessary in this project. Water, nuclease-free or distilled, is unsuitable due to its lack of a buffering capacity. Blood especially requires a buffer, as it contains carbonic acid and carbonate ions, capable of changing the pH of the sample. Additionally, cellular matrices contain DNAses—enzymes designed to cleave free nucleic acids. In PCR, it is important to deactivate DNAses, or risk reducing the efficiency of PCR. This is commonly done through Proteinase K incubation to denature the enzyme and the addition of EDTA to chelate any metal cofactors necessary for DNAse activity. Proteinase K incubation is not an option in a direct PCR, as it would destroy the polymerase and result in no amplification. The addition of too much EDTA can result in lower PCR efficiency due to chelation of necessary Mg^{2+}.

Using either NFW or 1X TE as a sample buffer, false positive results with lower but significant fluorescence were observed. Comparison to saliva samples with no false positive results, with no buffer, made clear that the buffer itself was causing these false positive results. Given that this is an allelic discrimination method, it is important to avoid these false positives. While the buffer components for the kits used are not published, the most common buffer system for DNA chemical biology studies is TE. There was concern, however, that the dilution would lower the buffering capacity, which could cause pH imbalance, especially in blood samples. A 9.000 mL aliquot of 10 mM Tris buffer, pH 8.0 (BioBasic, Ontario, Canada) was used to dilute 1.000 mL 1X TE, pH 8.0 buffer (MCLab,

California, USA). This resulted in a 10 mM Tris, 100 μM EDTA buffer at pH 8.0. Assays followed standard conditions for MB assays. The BioLine SensiFAST Probes Master and REV158 primer were used in these assays.

3.7: Phase IV: Allelic discrimination using MB-DPCR assays

Assays initially followed standard conditions for MB assays with direct sample addition as a modification. Once the assay was verified under standard conditions, the assay was then subjected to optimized time. The BioLine SensiFAST Probes Master and REV158 primer were used in these assays.

3.8: Design of new MB

The design of a MB is based on the target assay conditions. For an in-depth review of the mechanics of MBs, refer to 2.4: Overview of molecular beacons: an introduction to the biochemical and analytical bases of signal generation. A properly designed MB should remain closed and dark in the absence of an exact target at the desired annealing temperature, but be able to spontaneously form hybrids, resulting in extension and signal generation, in the presence of target. The spontaneity of the hybrid formation is governed by the length of the probe sequence; the probability of remaining dark in the absence of target is governed by the GC content and length of the stem. A 21 bp probe sequence may only anneal spontaneously and remain stable up to 48°C, while a 30 bp probe may retain its hybridization up to as high as 65°C. Stem content and length is an equally important consideration, with a higher GC or longer stem able to remain dark at increased temperatures and better compete with nontarget hybridization. This section has been adapted from molecular-beacons.org (Public Health Research Institute, 2016).

3.8.1: Probe sequence design

The first step in designing a MB is selecting the site and length of the probe sequence. The probe sequence should be centered about the polymorphism considered, increasing its discriminatory ability. This region should be within but not overlapping the primer sequences. This probe sequence should be long enough that, at the desired annealing temperature of the assay, the MB will still remain hybridized. For SNP discrimination, the dissociation temperature of the MB should be 7-10°C above the annealing temperature. Prediction of the dissociation temperature is accomplished through software estimates based on percent GC and nearest neighbor rules; these algorithms are not discussed in the review, but are available in literature. The probe length is normally between 15-30 bp.

3.8.2: Stem sequence design

The stem sequence is designed with a high GC content. This ensures a strong thermodynamic drive for the formation of a secondary structure, which is capable of quenching the fluorophore. The stem should melt at 7-10°C above the annealing temperature of the assay. The length of the stem is dependent upon the desired specificity of the probe; too long a stem can, however, lead to complications in the assay. As stem length increases, ΔG_{stem} becomes more negative than ΔG_{hybrid}, and the probe may not hybridize even in the presence of target. Also, guanine residues should not be used at the 5' terminus adjacent to the fluorophore; guanine has innate quenching ability, and will reduce SNR if placed adjacent to a fluorophore. Guanine should, however, be used at the 3' end to enhance quenching and increase SNR.

The dissociation temperature cannot be determined through typical percent GC or nearest neighbor algorithms, as the stem forms through intramolecular hybridization. Software such as Oligo7 is typically employed to predict the dissociation temperature of the stem through theoretical ΔG calculations under specified conditions. This study heavily relied on the data presented by Oligo7. An example of this data is shown as Figure 3.4. Considerations of final MB design

Once the desired probe and stem sequences are determined, the MB as a whole must be examined much like a primer for any additional secondary structures. With MBs, any secondary structure formation other than the intended stem can cause nonfunctional assays. Figure 3.4 shows the formation of the desired stem sequence plus an additional two secondary structures. These secondary structures can be safely ignored as they dissociate well below the intended annealing temperature of the assay. Since these will dissociate at above 36.4°C, and the assay runs at 48°C, the other secondary structures will not form spontaneously and thus may be neglected.

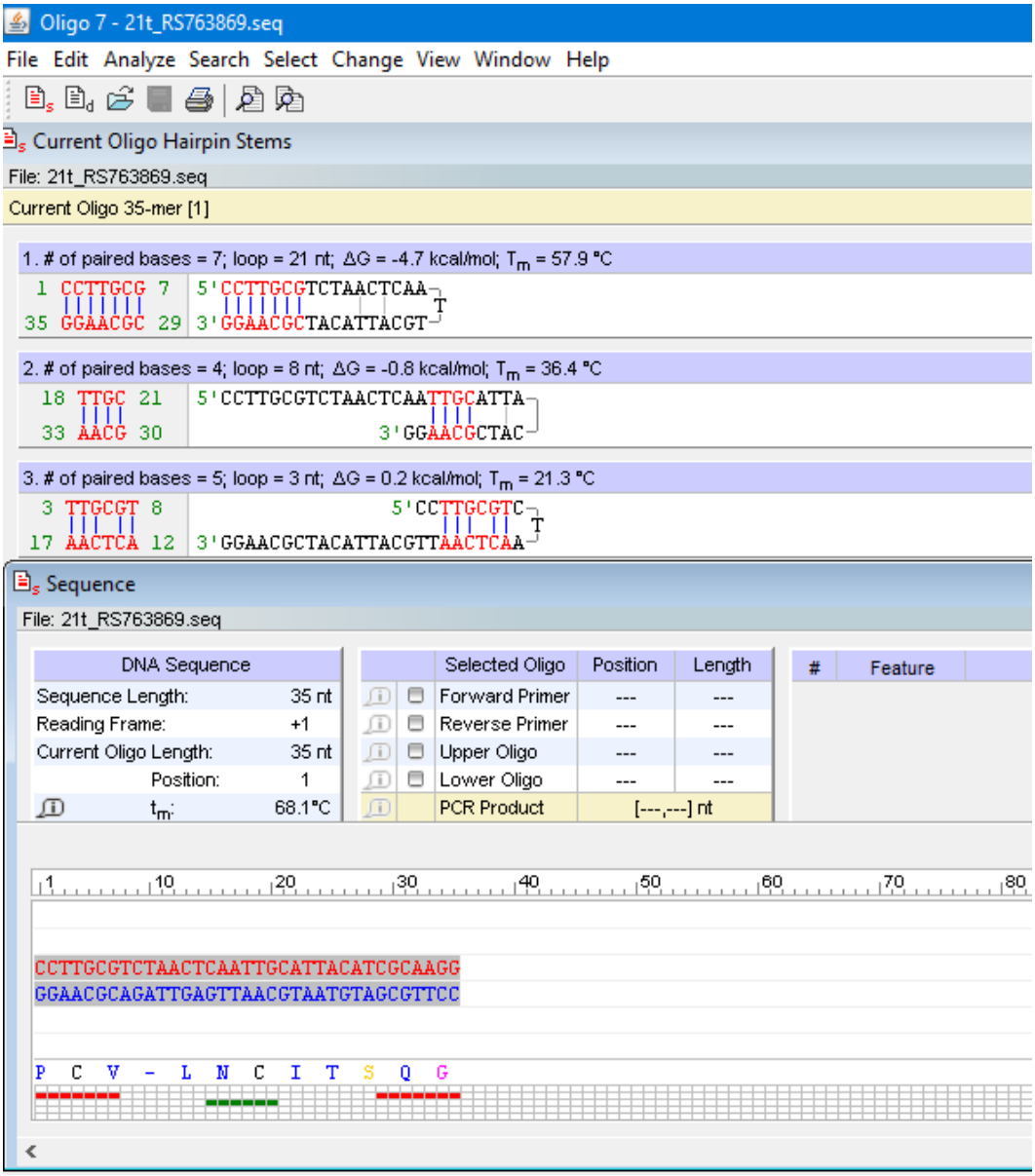

Figure 3.4: Data from Oligo7 demonstrating stem dissociation temperature prediction, T probe
This figure shows the predicted T_M of the whole MB using proprietary algorithms. The stems formed and their predicted dissociation temperatures are presented in the top half of the figure, and the entire sequence in the bottom half of the figure.

3.9: Normalization methods

As with most optical characterization methods, allelic discrimination through MBs requires normalization of the fluorescent data in order to easily interpret the results. A characteristic of MB fluorescent data is a constant, positive, linear slope of raw fluorescence units throughout the experiment, even in the no template control (NTC). This is due to the pseudo-first order kinetics observed between the MB and target DNA hybridization before the DNA concentration is roughly equivalent to that of the MB concentration. This is brought to a zero slope through the use of Microsoft Excel's linear regression (LINEST) function for the first 20 cycles, using the fluorescence unit value as known y-values and cycle number as known x-values. This gives the slope of the first 20 cycles, before the exponential phase occurs. This slope, multiplied by the cycle number minus one, is subtracted from the raw fluorescence unit value to obtain a near-zero slope for the baseline. The minimum of the sample's fluorescence unit value is then subtracted from this slope-adjusted fluorescence unit value to ensure all slope values are positive. Next, the adjusted fluorescence unit value is divided by the maximum of the positive control fluorescence unit value to obtain the normalized fluorescence units. The normalized fluorescence unit values are then plotted using a scatter plot with smooth lines, no data points. This scatter plot allows easy interpretation of allelic discrimination data. Further, the baseline is very close to horizontal, and this method of normalization does not alter the overall shape of the curve, only rotating it slightly. Samples are normalized to the positive control, as this is the standard for the reaction. A normalized fluorescence unit value either higher or lower than the positive control is not indicative of any specific result; only the presence (or lack thereof) an exponential phase followed by linear growth is indicative of a specific allele. After normalization, the average of the baseline is calculated. If the

maximum signal generated is at least three times as large as the average of the baseline, the signal is said to be significant and above the LOD. Below this, the signal is unreliable, and the sample must be retested.

CHAPTER 4: RESULTS

CHAPTER 4: RESULTS

4.1: Introduction

This chapter presents the results of the project in their entirety. As described in Chapter 1: Introduction, the specific aims of this project were to develop a protocol capable of allelic discrimination using MBs, apply that protocol in an unpurified, direct PCR format, and optimize the protocol such that its time is as short as possible without compromising results. The project followed a methodical approach to test each individual component, followed by stringent assays to select an appropriate kit, before ultimately testing the assay in a direct format. The division of these three phases was essential to the success of the project.

4.2: Phase I: Verification of amplification and detection

4.2.1: Verification of FOR100 and REV100 primers

The primers were tested using a verification protocol (Section **Error! Reference source not found.Error! Reference source not found.**) and analyzed using gel electrophoresis. This protocol used an annealing temperature gradient for analysis of optimal annealing temperature. This primer set was designed to amplify a 100 bp region. Any fragment sizes not in agreement with this size can be attributed to non-specificity and primer-dimer formation, neither of which is desirable in a PCR assay. The subsequent gel electrophoresis image allows the determination of the temperature at which primer-dimer formation and nonspecific amplification no longer occur, while still amplifying the target region. An image of the verification gel electrophoresis assay is shown as Figure 4.1.

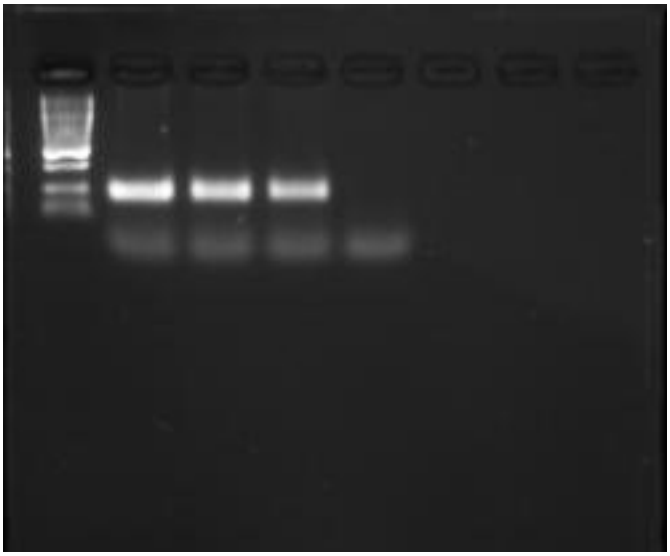

Figure 4.1: Verification of FOR100 and REV100 amplification.
Lanes, from left to right: 50 bp ladder, positive control, positive control, positive control, negative control. Note the amplification at 100 bp in all positive controls, but none in the NTC. All samples exhibit dimer formation. Image by Dr. Fredericks.

4.2.2: Thermal Profile of MB probes

The fluorescence as a function of temperature, or thermal profiles, of MB probes were determined before incorporation into an assay. This involved using a complementary oligonucleotide in a multifold excess of MB probe, and subsequently analyzing the normalized fluorescence with respect to temperature. The assay was then repeated with a mismatched oligonucleotide in order to gauge the specificity of the probe at a given temperature. The thermal profiles were then used to determine the optimal annealing temperature for the PCR assay, based on the difference in fluorescence between complementary and mismatched profiling assays for all MBs and targets.

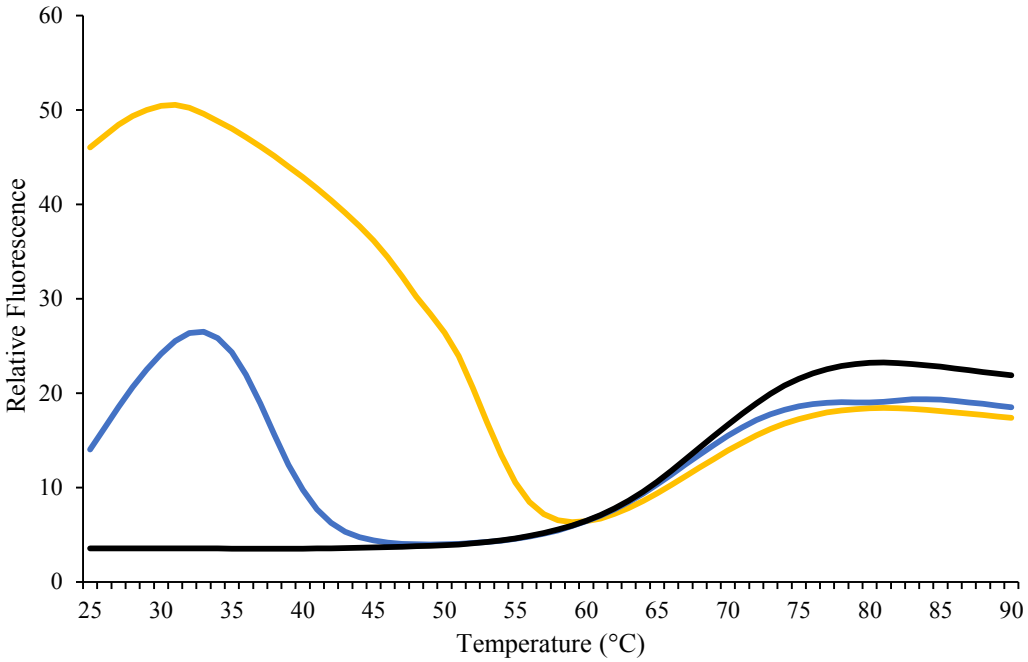

Figure 4.2: Thermal profile of MB probes using C probe complement
Blue line represents T probe, yellow line represents C probe, black line represents control.

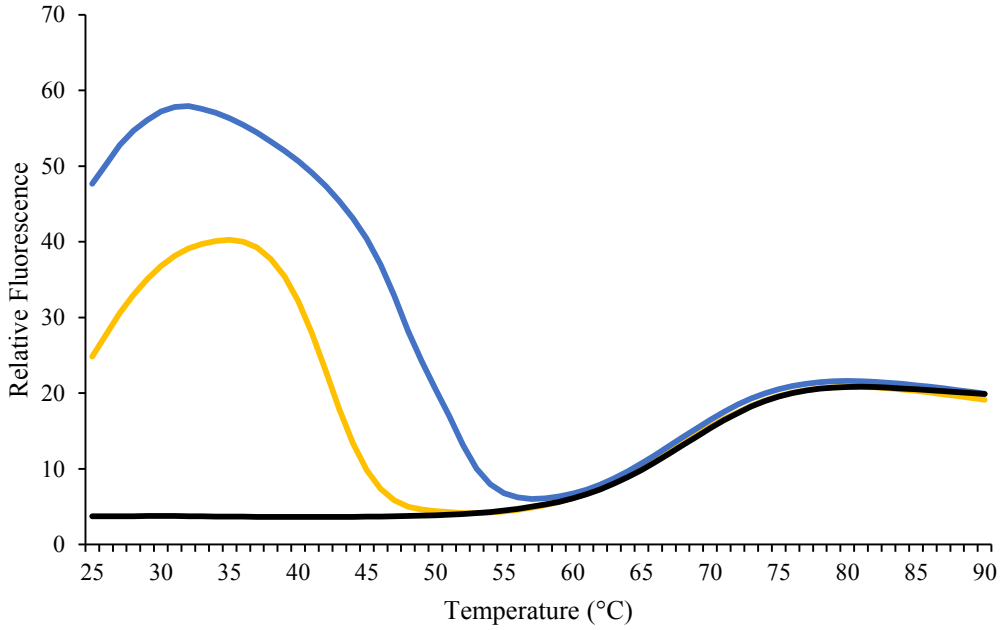

Figure 4.3: Thermal profile of MB probes using T probe complement
Blue line represents T probe, yellow line represents C probe, black line represents control.

4.2.3: Verification of MB assay

The MB assays were verified using a standard 90 second cycling profile for at least 40 cycles. This verification first studied each MB probe separately in order to limit any interferences with the other probe. The T probe was conjugated with a FAM fluorophore, with λ_{max} at 520 nm; the C probe was conjugated with a HEX fluorophore, with λ_{max} at 556 nm (Biosearch Technologies, 2016). This data is designed to show the performance of each probe at both target wavelengths of fluorescence, and validates the MBs are performing as desired.

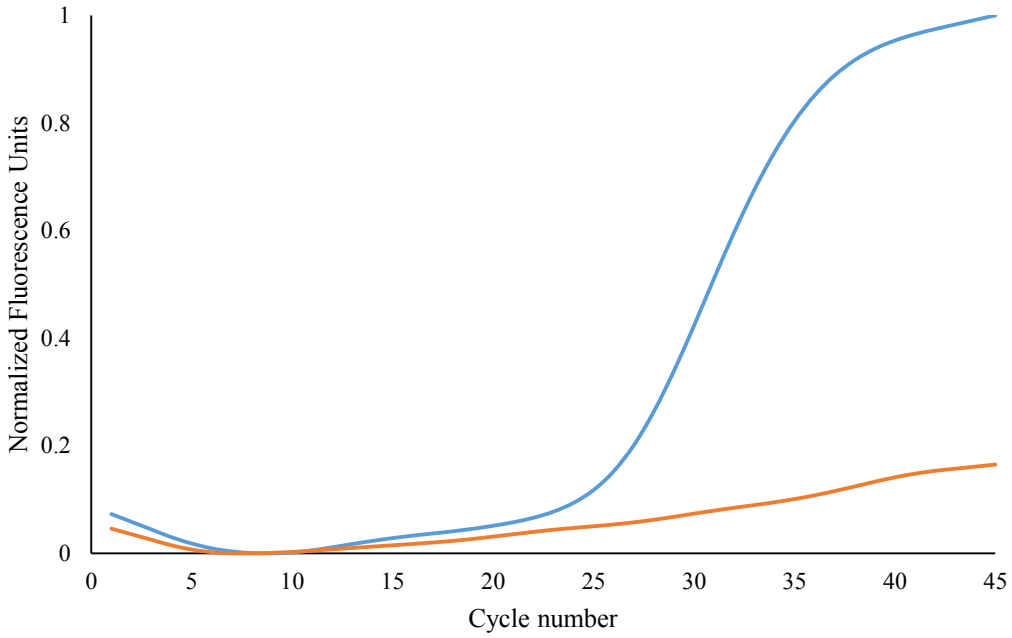

Figure 4.4: Verification of MB assay, T probe only, heterozygous gDNA target, 510 nm acquisition
Blue line represents T probe with gDNA, orange line represents T probe without gDNA.

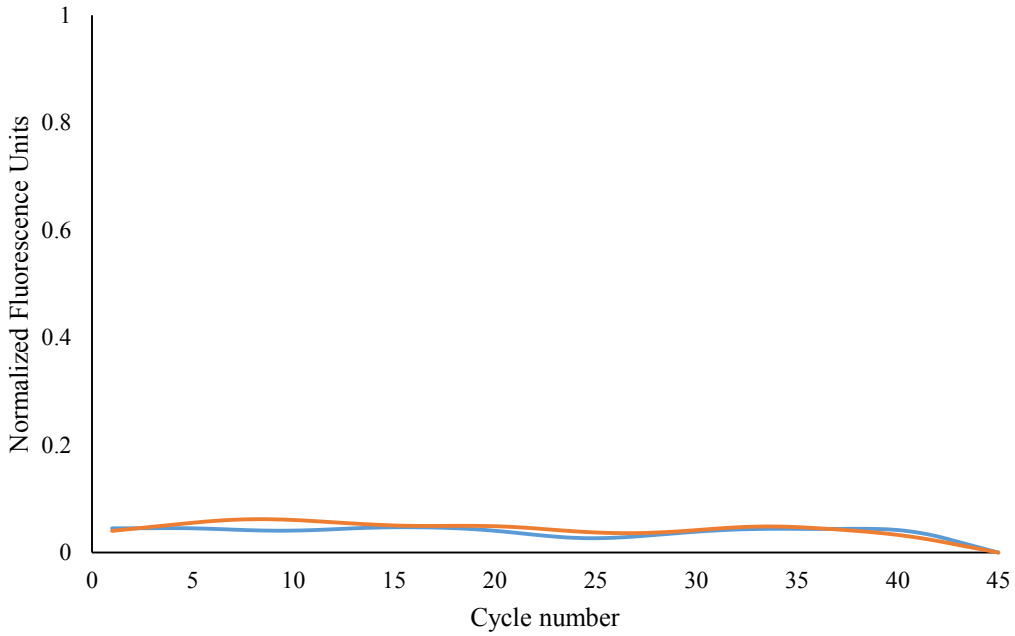

Figure 4.5: Verification of MB assay, T probe only, heterozygous gDNA target, 555 nm acquisition
Blue line represents T probe with gDNA, orange line represents T probe without gDNA.

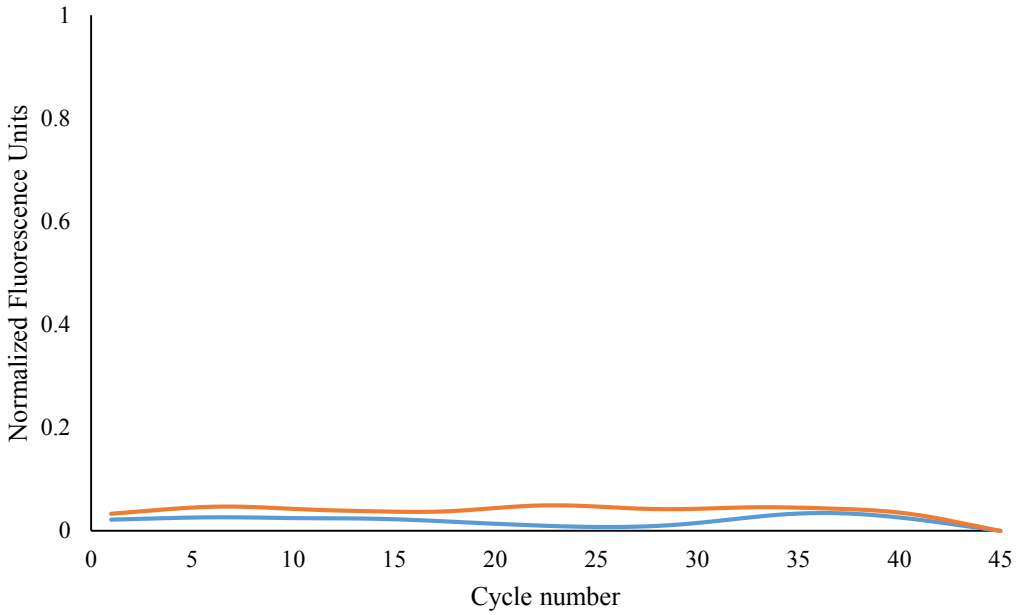

Figure 4.6: Verification of MB assay, C probe only, heterozygous gDNA target, 510 nm acquisition
Blue line represents C probe with gDNA, orange line represents C probe without gDNA.

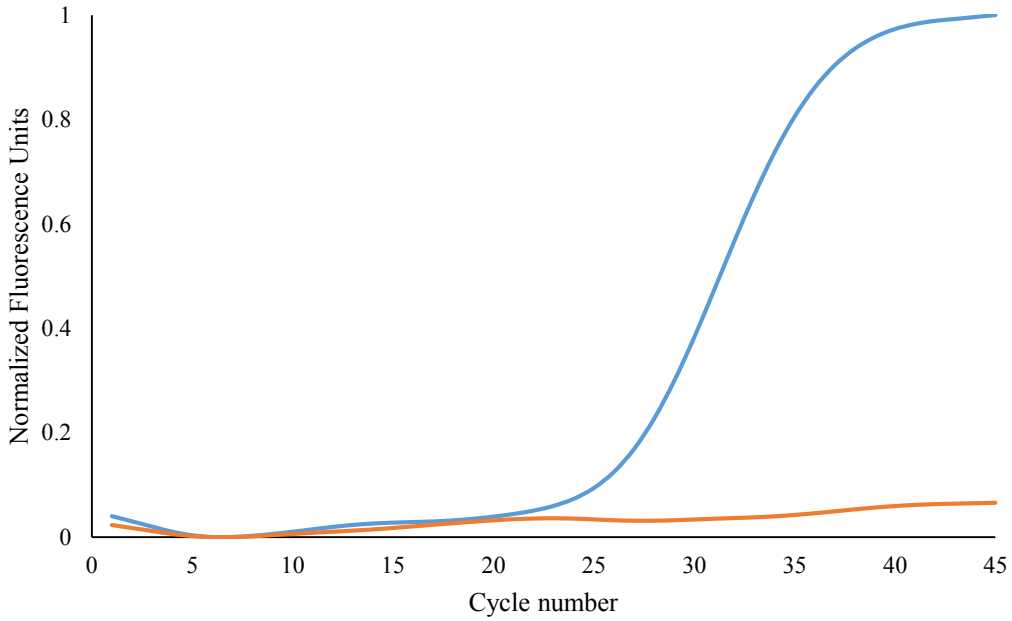

Figure 4.7: Verification of MB assay, C probe only, heterozygous gDNA target, 555 nm acquisition
Blue line represents C probe with gDNA, orange line represents C probe without gDNA

After the MBs were validated separately, the MBs had to be validated in a multiplex format. It is important to note that this is not a true PCR multiplex which amplifies multiple loci in one reaction; rather, this is a probe multiplex, which uses multiple probes in the same reaction. Interferences and spectral effects are possible, and it is necessary to ensure valid performance of the probes when used in conjunction. This requires adding both MBs in equal concentration to a single reaction mixture, then collecting the fluorescence data for each reaction on both channels.

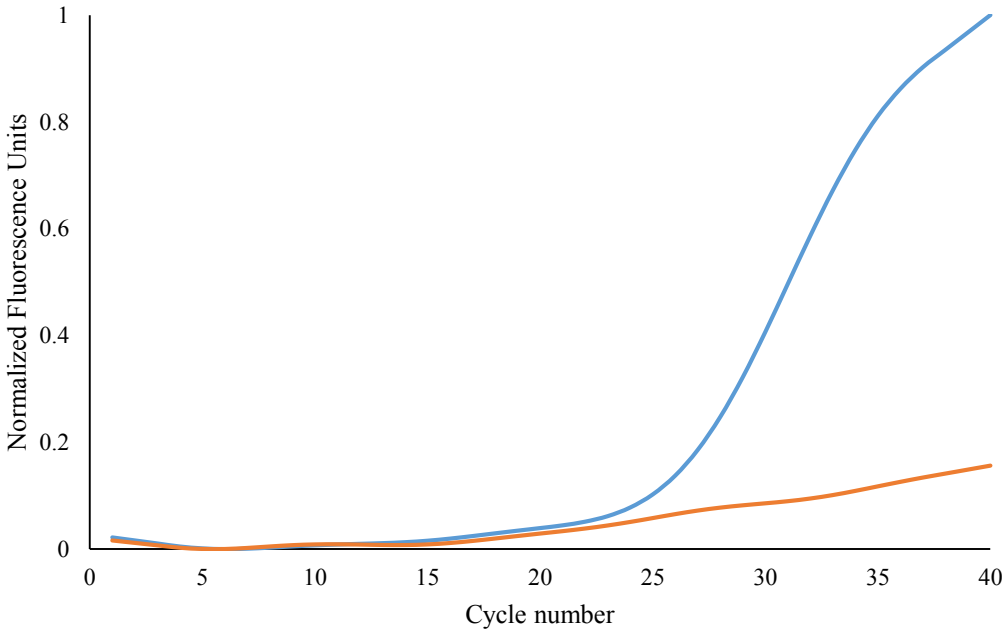

Figure 4.8: Multiplex MB assay, heterozygous gDNA target, 510 nm acquisition Blue line represents positive control, orange line represents negative control.

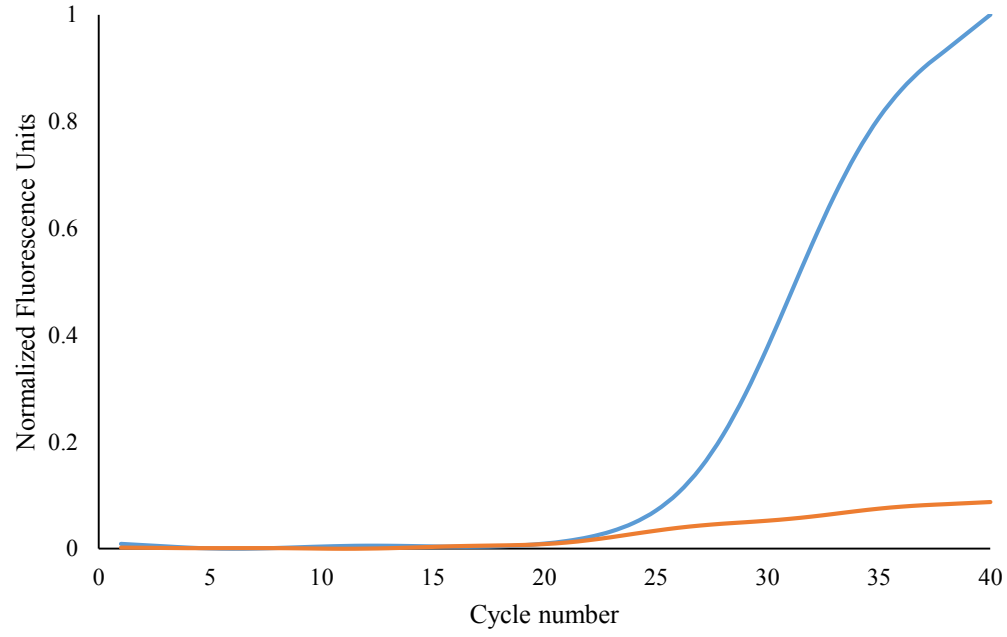

Figure 4.9: Multiplex MB assay, heterozygous gDNA target, 555 nm acquisition Blue line represents positive control, orange line represents negative control.

4.2.4: Screening of genotypes

Upon verification with a purified heterozygous C/T sample, the assay was then used to discriminate genotypes in order to screen subjects for the study. This was also useful in determining the ability of the MBs to discriminate homozygous samples, as only heterozygotes had been tested up to this point. Figure 4.10 and Figure 4.11 show the results from the genotyping test conducted.

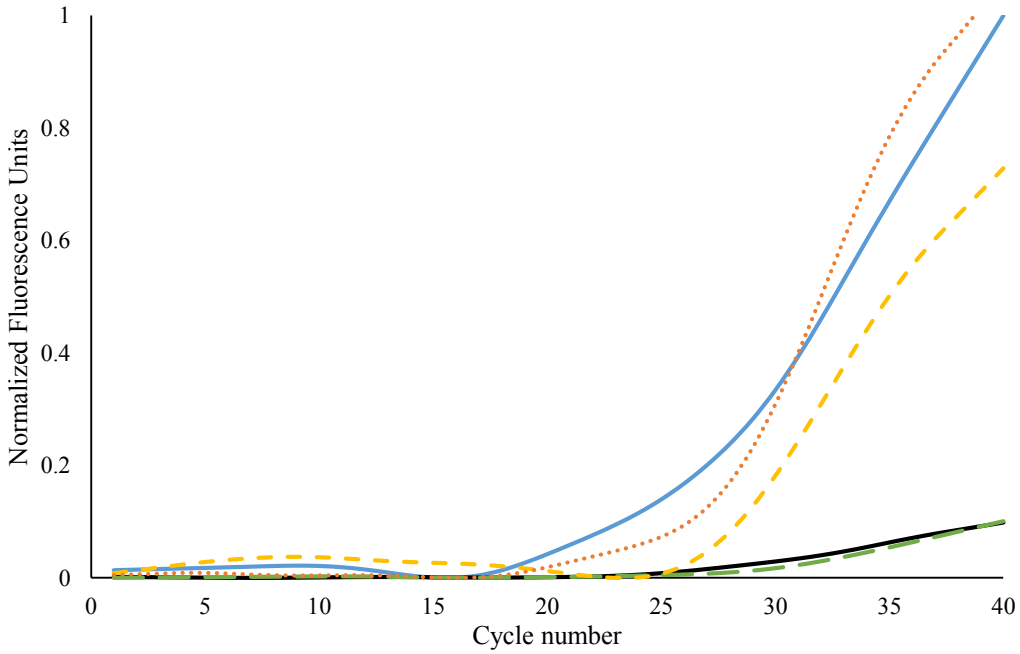

Figure 4.10: Genotyping test, purified DNA, T probe
Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a homozygous C sample, yellow short dash line represents a heterozygous sample, red dotted line represents a homozygous T sample.

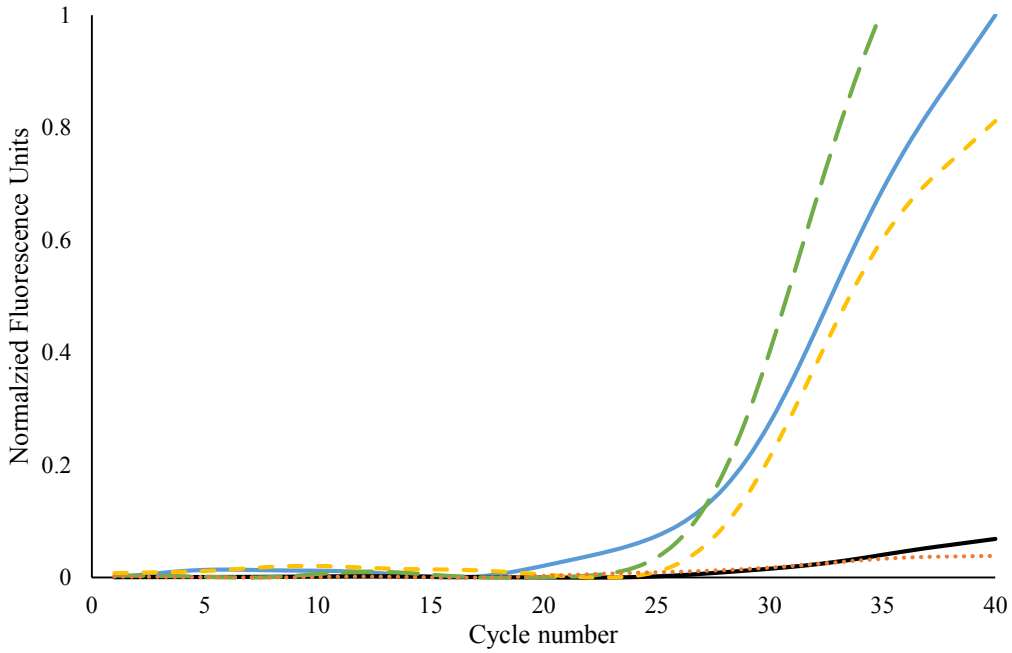

Figure 4.11: Genotyping test, purified DNA, C probe
Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a homozygous C sample, yellow short dash line represents a heterozygous sample, red dotted line represents a homozygous T sample.

4.3: Phase II: Comparison and selection of PCR master mix

4.3.1: Kits considered

This study required the consideration of multiple kits, as direct PCR has not yet been adapted for use in a real-time format, specifically for MB assays. As such, it was important to test multiple PCR master mixes for efficiency in direct formats coupled with MB assays. Direct PCR kits such as ThermoScientific's Phusion master mixes typically utilize a *Pyrococcus furiosus*-based polymerase with innate 3'-5' exonuclease activity as a proofreading mechanism given the difficulty associated with direct PCR. Master mixes designed for real-time PCR, including the Roche LightCycler 480 Probes Master and BioLine SensiFAST Probes Master,

typically employ a *Thermus aquaticus*-based polymerase, without proofreading ability.

4.3.2: ThermoScientific Phusion kits

The ThermoScientific Phusion kits were incompatible with the MB assays studied. This can be seen graphically as Figure 4.12 and Figure 4.12. Adjusting the figures for slope would distort the data such that it is no longer representative. The kits do not allow the MBs to retain the DAP in close proximity, and it is likely that either the fluorophore or quencher are being cleaved each cycle. Such rapid increase in fluorescence each cycle is not expected with a MB assay. As such, the Phusion kits were removed from consideration. While this kit is designed for direct PCR applications, it is not useful for this particular application, even under purified DNA conditions.

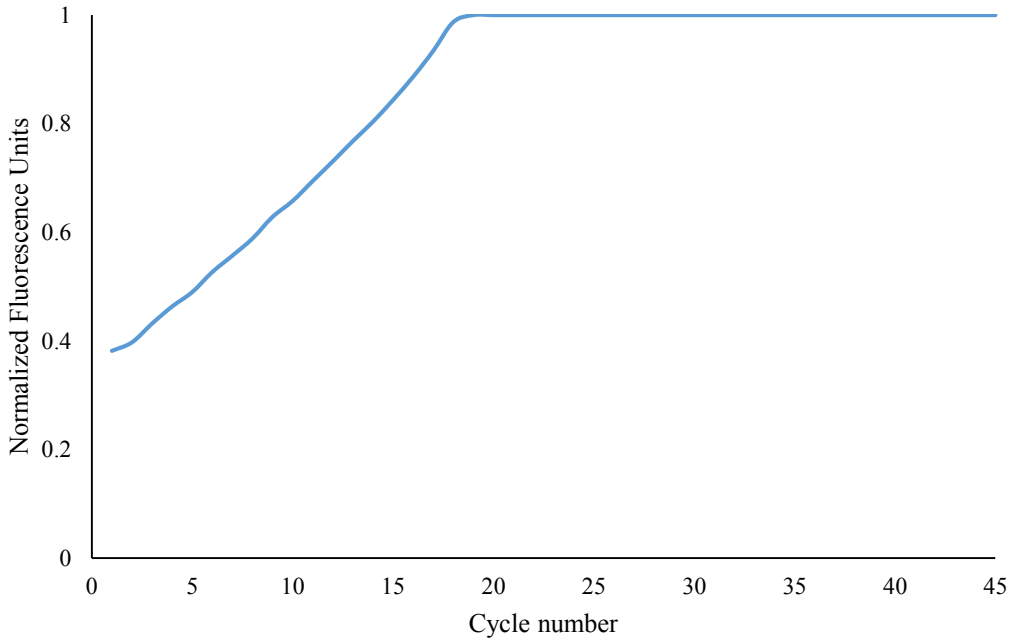

Figure 4.12: MB assay, T probe, Phusion kit

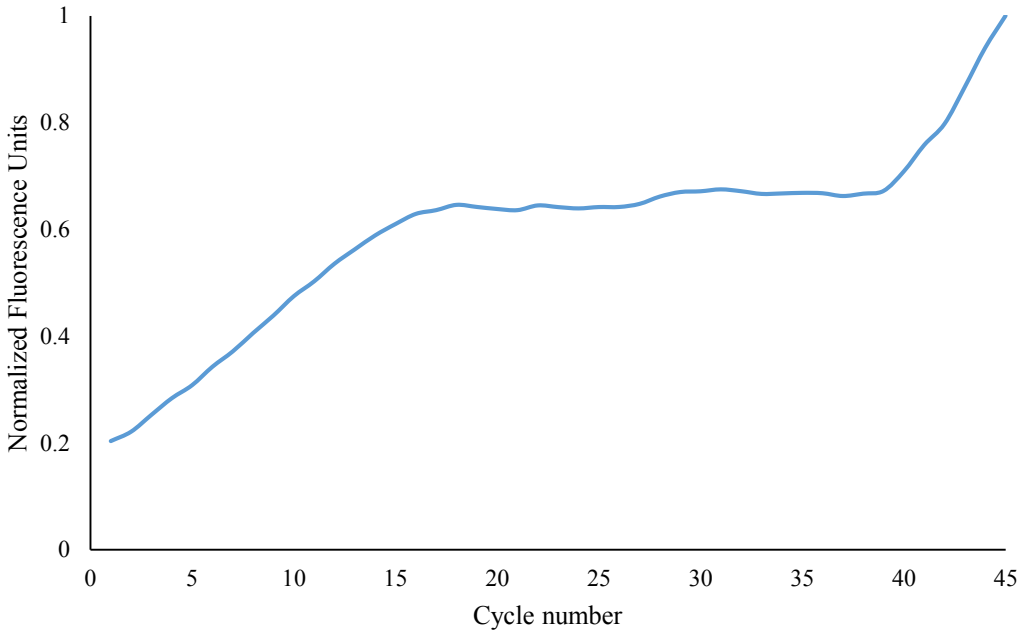

Figure 4.13: MB assay, C probe, Phusion kit

4.3.3: Roche LightCycler 480 and BioLine SensiFAST comparison

The Roche LightCycler 480 and BioLine SensiFAST kits both initially worked well with a MB assay under purified conditions. Comparison of C_T values and limit of detection (LOD) testing was used to determine which kit would be most compatible with and sensitive to the applications of the project.

4.3.3.1: C_T values

When studying a qPCR assay, the most useful and reliable method of comparing kits is to use a serial dilution of DNA in replicate samples, set the thermocycle to the manufacturer's recommendation, and compare the point at which the exponential signal phase rises significantly higher than the background standard deviation. This is typically done through software, and the point is known

as the cycle threshold, or C_T. A lower concentration of DNA initially will lead to a higher C_T, as it takes more cycles to replicate enough DNA to reach roughly equivalent concentration of MBs.

Table 4.1: Average C_T Values, Comparison of Roche and BioLine kits

	T probe	C probe
Roche, 20 ng	30.70	27.23
Roche, 10 ng	30.90	27.39
Roche, 5 ng	34.32	29.29
Roche, 2.5 ng	35.35	29.90
Roche, 0.25 ng	30.05	N/A
BioLine, 20 ng	25.02	22.80
BioLine, 10 ng	25.37	23.06
BioLine, 5 ng	27.86	25.19
BioLine 2.5 ng	28.32	25.78
BioLine, 0.25 ng	33.02	30.69

4.3.3.2: Reliability comparison

As seen when comparing the BioLine kits (Figure 4.16 and Figure 4.17) to the Roche kits (Figure 4.14 and Figure 4.15), the BioLine kits exhibit increased sensitivity. The Roche kit was unable to reliably amplify down to 0.25 ng of DNA per reaction, while the BioLine kit was consistently able to amplify down to 0.25 ng per reaction. Note allelic dropout occurs in the Roche kit when comparing Figure 4.14 and Figure 4.15; a heterozygous sample was used, and exhibits heterozygous response in the higher concentrations.

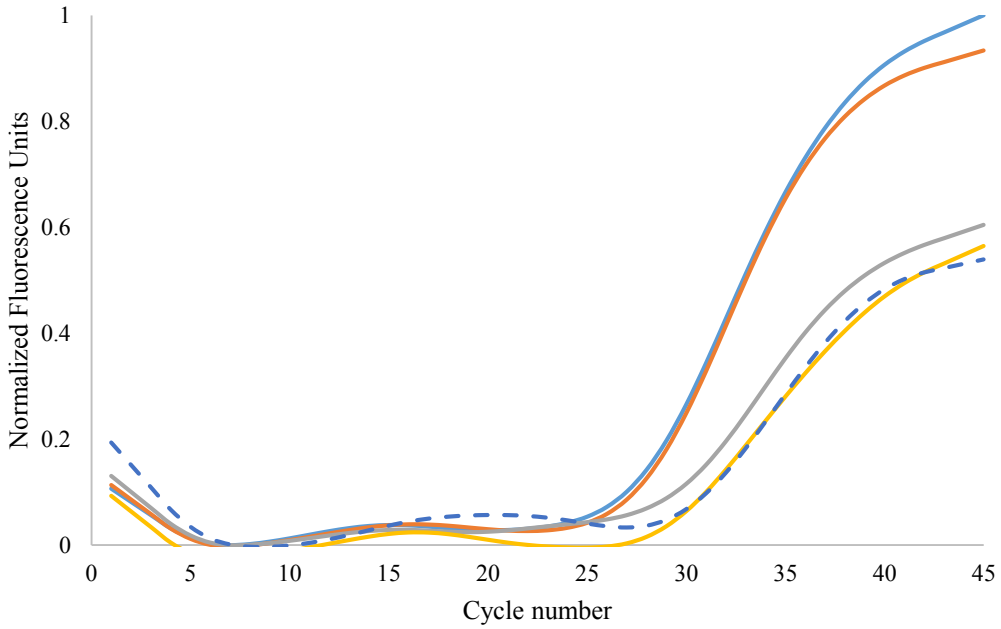

Figure 4.14: Sensitivity comparison, Roche kit, T probe
Blue line represents 20 ng, orange line represents 10 ng, grey line represents 5 ng, yellow line represents 2.5 ng, blue dashed line represents 0.25 ng.

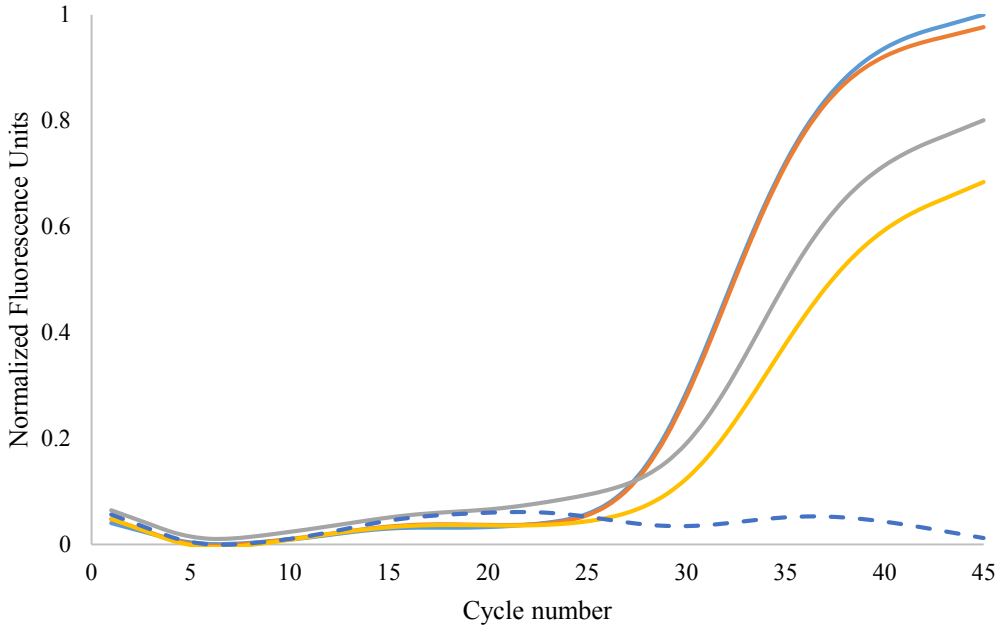

Figure 4.15: Sensitivity comparison, Roche kit, C probe
Blue line represents 20 ng, orange line represents 10 ng, grey line represents 5 ng, yellow line represents 2.5 ng, blue dashed line represents 0.25 ng.

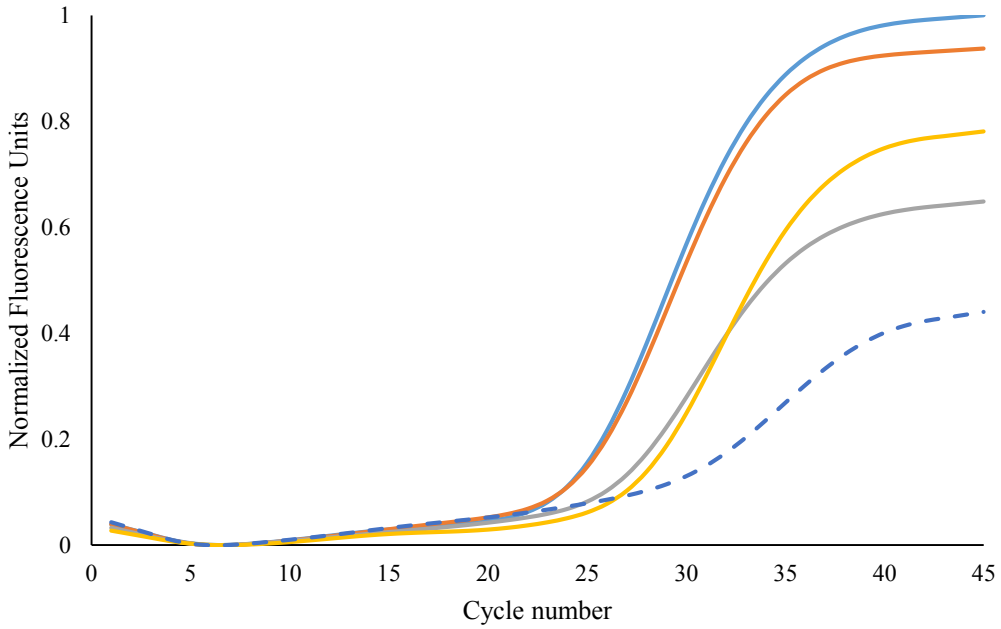

Figure 4.16: Sensitivity comparison, BioLine kit, T probe
Blue line represents 20 ng, orange line represents 10 ng, grey line represents 5 ng, yellow line represents 2.5 ng, blue dashed line represents 0.25 ng.

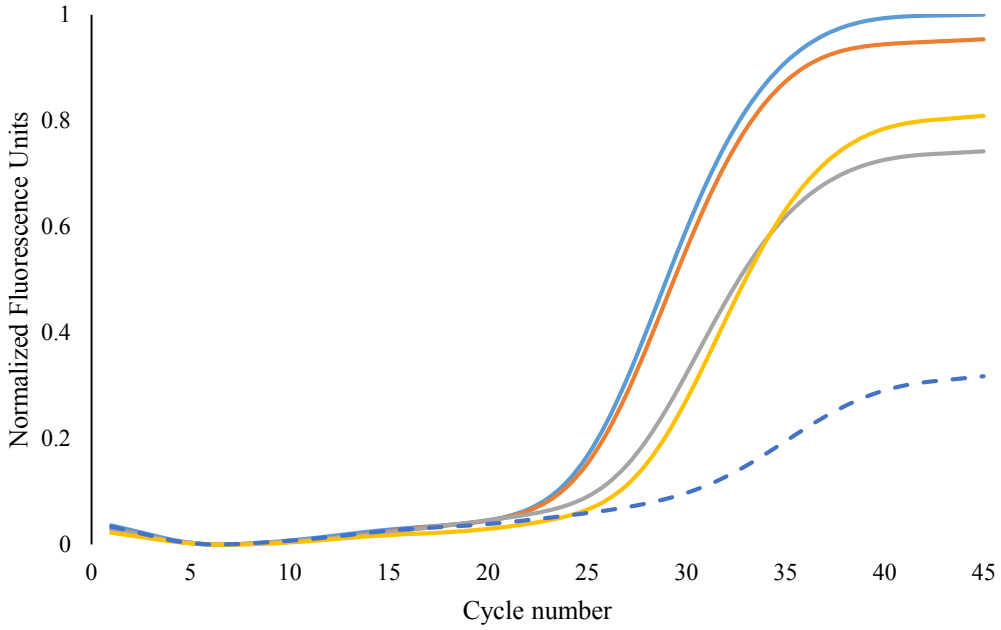

Figure 4.17: Sensitivity comparison, BioLine kit, C probe
Blue line represents 20 ng, orange line represents 10 ng, grey line represents 5 ng, yellow line represents 2.5 ng, blue dashed line represents 0.25 ng.

Table 4.2: Summary of Sensitivity Comparison

	Roche	BioLine
20 ng	Detected	Detected
10 ng	Detected	Detected
5 ng	Detected	Detected
2.5 ng	Detected	Detected
0.25 ng	Negative	Detected

4.4: Phase III: Direct MB-PCR Assay Development

4.4.1: Verification of amplification under direct conditions

Verification of amplification in an assay is crucial before the use of more expensive, labeled probes such as MBs. Amplification under purified conditions using these primers had been confirmed in Section 4.2.1: Verification of FOR100 and REV100 primers. Confirmation of amplification by the original primers under unpurified, direct PCR conditions was required to ensure kit compatibility. This was verified by gel electrophoresis, a size-exclusion method for the separation of DNA fragments. An image of the gel is shown as Figure 4.18. Amplification was carried out using the verification protocol (Section **Error! Reference source not found.Error! Reference source not found.**), using Roche LightCycler™ 480 Probes Master instead of the BioLine SensiFAST™ Probes Master. An image of the molecular weight marker used is shown as Figure 3.1.

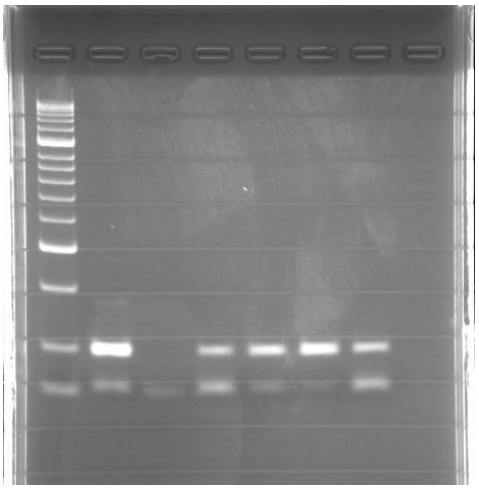

Figure 4.18: Gel electrophoresis of direct PCR fragments, after amplification Lanes, from left to right: 50 bp step ladder; positive control; no template control; hair sample; buccal swab sample; saliva sample; blood sample.

Referencing Figure 4.18, all lanes containing amplified DNA have a band at roughly 100 bp as expected. This confirms amplification under the conditions used. Lane 3, the no template control, shows a faint band at roughly 50 bp, as do all samples and the control. This is indicative of inefficient priming, likely caused by dimer formation of the REV100 primer. A new primer was developed and tested in response to this.

4.4.2: Design of new primer set

The new primer, REV158, was so named as it has an expected fragment size of 158 bp. The alignment is shown as Figure 3.3: Alignment using BioEdit, including FOR100, REV158, C probe, and T probe. This primer was designed according to the protocol described in Section **Error! Reference source not found.Error! Reference source not found.**. Since this is an unverified primer, amplification under standard, purified conditions must be confirmed before further testing. Verification of amplification of the new primer set was accomplished

through an annealing temperature gradient. This was used to determine the optimal annealing temperature for this primer set. This utilized the verification protocol described in Section **Error! Reference source not found.Error! Reference source not found.**, with NFW used in place of MB mix. An image of the subsequent gel after electrophoresis is shown as Figure 4.19. Given the thermal profiles of the MBs and the gel image after annealing temperature gradient, it was clear that nonspecific amplification would be unavoidable in a MB assay. The highest temperature the MBs can withstand while still annealing is 48°C, while primer specificity is not achieved until at least 52°C.

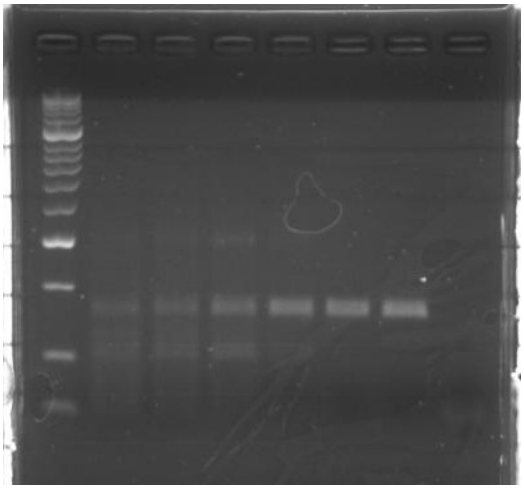

Figure 4.19: Gel electrophoresis image of REV158 primer annealing temperature gradient
Lanes, from left to right: 50 bp step ladder; 48.0°C annealing temperature; 49.5°C annealing temperature; 51.9°C annealing temperature; 54.7°C annealing temperature; 61.7°C annealing temperature; 63.7°C annealing temperature; NTC.

4.4.3: Comparison of new primer set

A comparison of the two primers, REV158 and REV100, was completed in order to determine which had better priming efficiency after introduction of MBs.

A positive control of 20 ng genomic, purified DNA and a NTC were used for both primers, using FOR100 as the forward primer for all reactions. Verification conditions are described in Section **Error! Reference source not found.Error! Reference source not found.**, modified for an annealing temperature of 55°C each cycle instead of 48°C. The 55°C annealing temperature was used to minimize the effects of nonspecific priming and observe the MBs' effects. An image of the agarose gel after electrophoresis is shown as Figure 4.20: Comparison of priming efficiency for REV100 and REV158 primers.

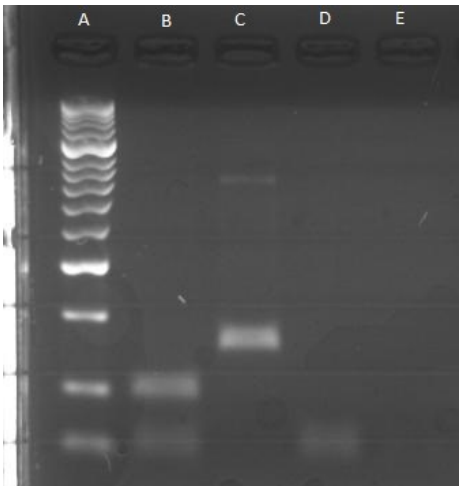

Figure 4.20: Comparison of priming efficiency for REV100 and REV158 primers Lanes: A—50 bp step ladder; B—REV100 positive control; C—REV158 positive control; D—REV100 NTC; E—REV158 NTC.

The positive controls for both primers show bands at roughly the expected sizes, confirming amplification. Further, both REV100 lanes show inefficient priming at roughly 50 bp, while neither REV158 lanes have this band. In fact, the REV158 NTC lane is completely clear of any visible bands, indicating a much higher priming efficiency when compared to REV100.

4.4.4: Analysis of primers in real-time using melt analysis

Further comparison information between primer sets was gathered using a SYBR Green-based qPCR assay with subsequent melt analysis. This analytical method employs a standard PCR assay with the addition of SYBR Green—an intercalating dye. The mechanism of the dye is described in 2.9: Signal generation and quenching: Dynamic and Static quenching. This dye binds to the minor groove of DNA with great specificity and forms a J-type aggregate, thus greatly increasing fluorescence. This dye will bind indiscriminately, and thus will bind to all fragment sizes after amplification. The melt phase of analysis then measures the change in fluorescence with respect to time. The temperature begins at 50°C and slowly rises. Initially, all fragments will have SYBR Green intercalated, and the raw fluorescence will be at its maximum at the lowest temperature of the melt. As the temperature rises, however, there will be sudden drops in fluorescence, followed by a period of a slower decrease in fluorescence. Using the first-derivative plot of this fluorescence with respect to time, peaks are analyzed to show at what temperatures the fragments dissociate. The number of peaks corresponds to the number of fragments present.

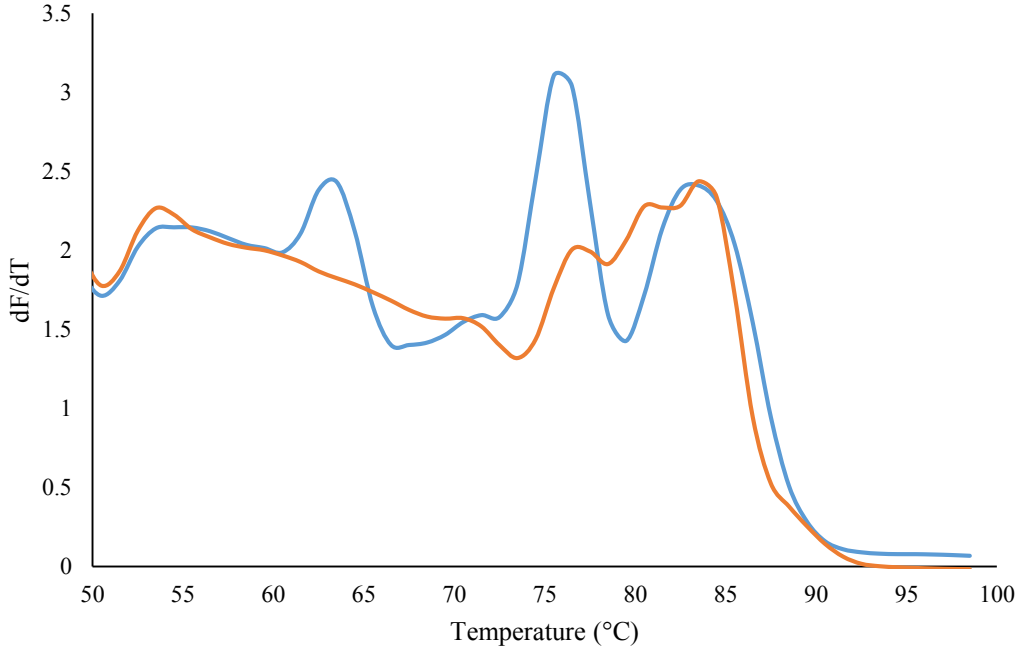

Figure 4.21: Melt analysis of qPCR comparison of REV100 and REV158 primers after 48°C annealing
Blue line corresponds to REV100, orange line corresponds to REV158.

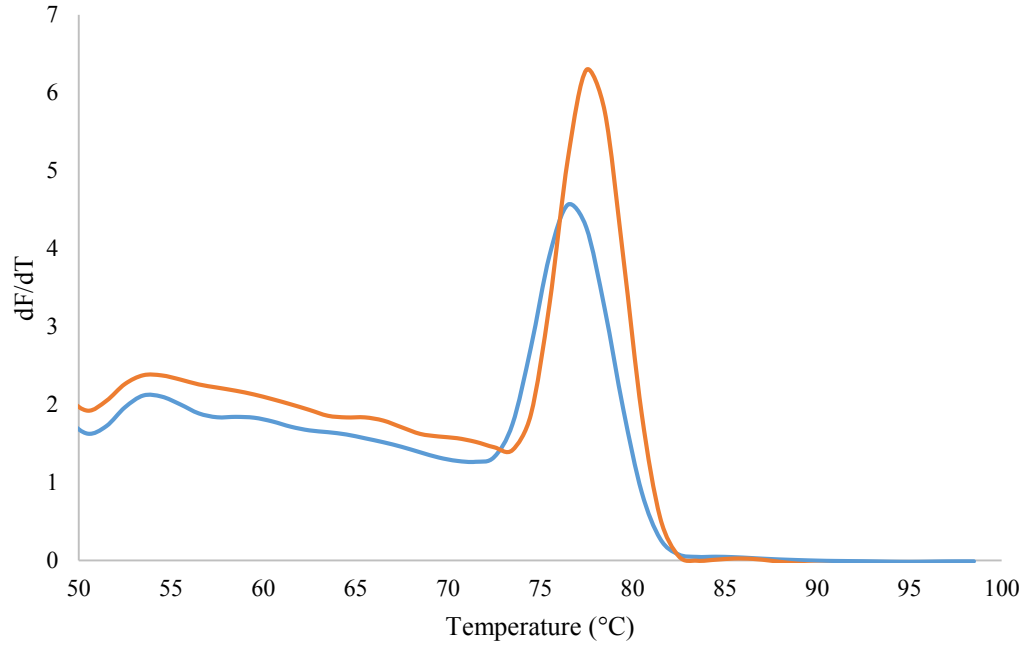

Figure 4.22: Melt analysis of qPCR comparison of REV100 and REV158 primers after 58°C annealing
Blue line corresponds to REV100, orange line corresponds to REV158.

4.4.5: Verification of direct amplification using REV158

The primer set was then tested for amplification of direct, unpurified DNA using the verification protocol described in Section **Error! Reference source not found.Error! Reference source not found.** with sample additions for direct, unpurified DNA sources as described in **Error! Reference source not found.**An image of this gel is shown as Figure 4.23: Verification of amplification of direct samples using REV158 primer.

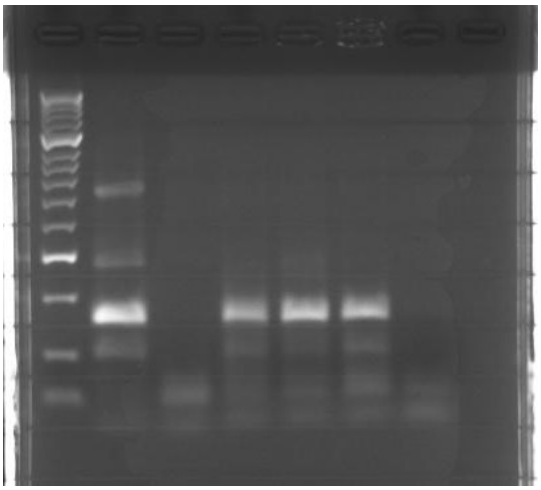

Figure 4.23: Verification of amplification of direct samples using REV158 primer Lanes, from left to right: 50 bp step ladder; positive control; NTC; hair; saliva; buccal swab; blood.

4.4.6: Assay Optimization

Optimization is necessary to ensure the fastest possible time from sample collection to results, and the most efficient temperature is used. As explained in Chapter 1: Introduction, the purpose of this project is to develop a method which is faster than industry standard, capable of providing reliable preliminary results pertaining to the potential success of amplification and possibly a preliminary

profile. The method must be robust and reliable, while still sufficiently improving the speed at which results can be produced. A standard 90 second cycle is not sufficient for this purpose.

Thermocycle optimization was carried out through the manipulation of times within the cycling events. Initial denaturation was performed according to the manufacturer's instructions so as to ensure complete cleavage of the heat-labile blocking groups which prevent enzyme activity below 75°C. Extension time was first tested for optimization, followed by denaturation time, and finally annealing time. Results show that the optimal thermocycle for speed and data integrity is a 5 second melt, 30 second anneal without extension. Further optimization allowed the determination of the minimum MB concentration required to elicit a reliable discriminatory response. After confirmation of amplification under the optimized conditions, the assay was tested for sensitivity limits down to 0.1 ng of purified genomic heterozygotic DNA.

To ensure no loss of data integrity occurred, the cycle threshold (C_T) values were used as a quantitative comparison. These values are selected by using the maximum of the pre-exponential fluorescence as a background, then observing when the exponential phase crosses this background. This threshold is dependent upon the concentration of DNA in the reaction, and has an inverse relationship. If less DNA is present, the reaction will have a higher C_T value. All reactions for this investigation were prepared simultaneously and aliquoted from a master mix containing all components: PCR kit, primers, water, and DNA. All aliquots are expected to have negligible differences, so any differences in C_T can be attributed to PCR efficiency. Higher C_T values in this case are caused by variations in PCR efficiency due to cycling differences. Slight variations are to be expected as with

any experiment. No optimized reactions performed worse than the baseline 90 second cycle.

Table 4.3: C_T Values of Assay Optimization

	Probe		
	C_T, T	C_T, C	C_T, NTC
30-30-30	23.58	22.20	N/A
30-30-15	21.25	21.88	N/A
30-30-5	22.62	23.30	N/A
30-30-0	21.88	20.97	N/A
15-30-0	22.54	22.18	N/A
10-30-0	21.79	20.77	N/A
5-30-0	23.42	23.83	N/A
5-15-0	21.76	19.58	N/A
1-15-0	22.64	20.77	N/A

Values are averages of the triplicate measurements.

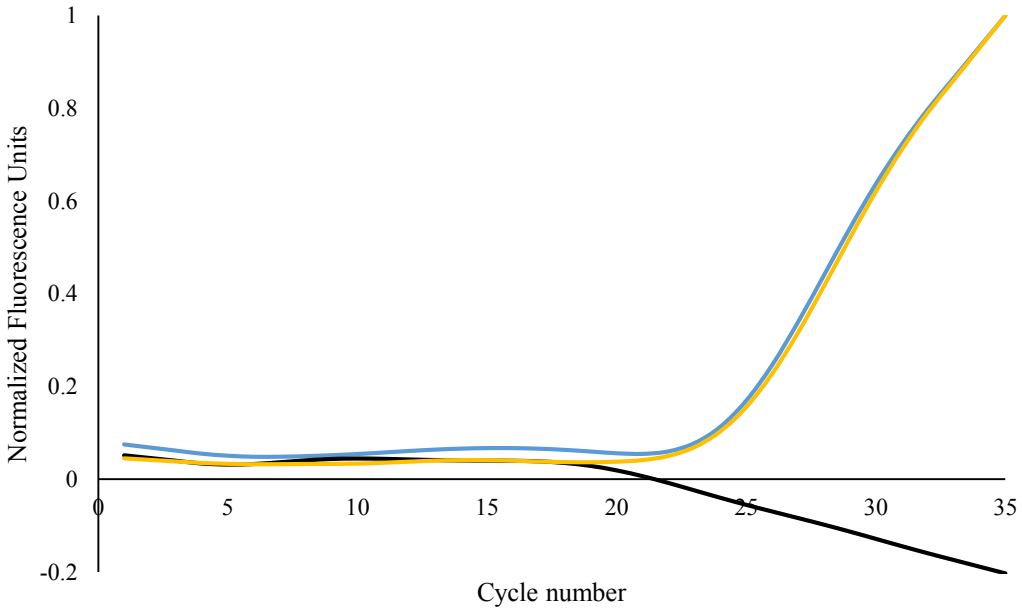

Figure 4.24: Assay optimization, 30 s denaturation, 30 s annealing, 30 s extension. Blue line represents T probe, yellow line represents C probe, black line represents NTC.

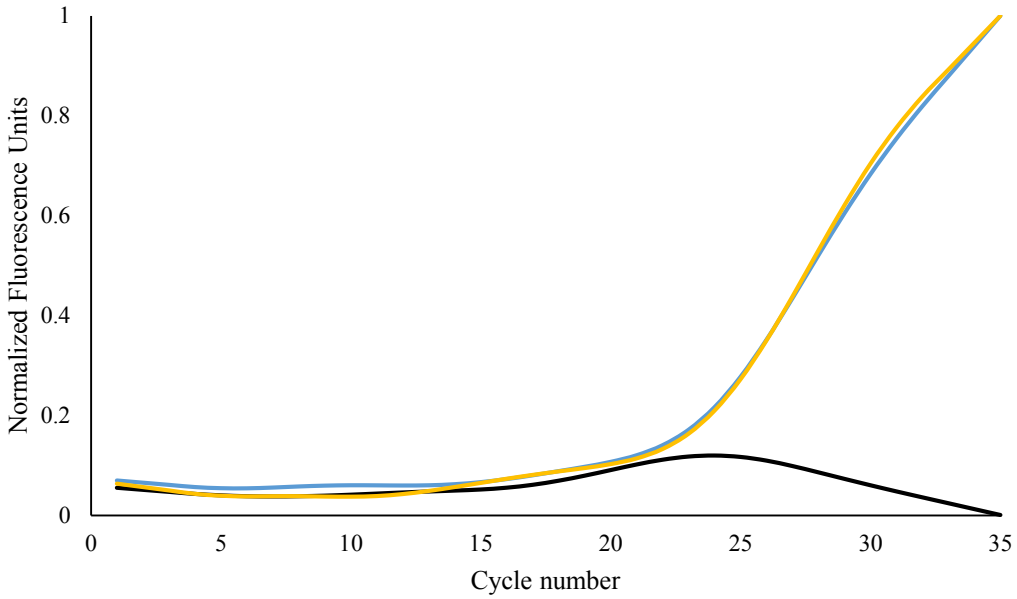

Figure 4.25: Assay optimization, 30 s denaturation, 30 s annealing, 15 s extension. Blue line represents T probe, yellow line represents C probe, black line represents NTC.

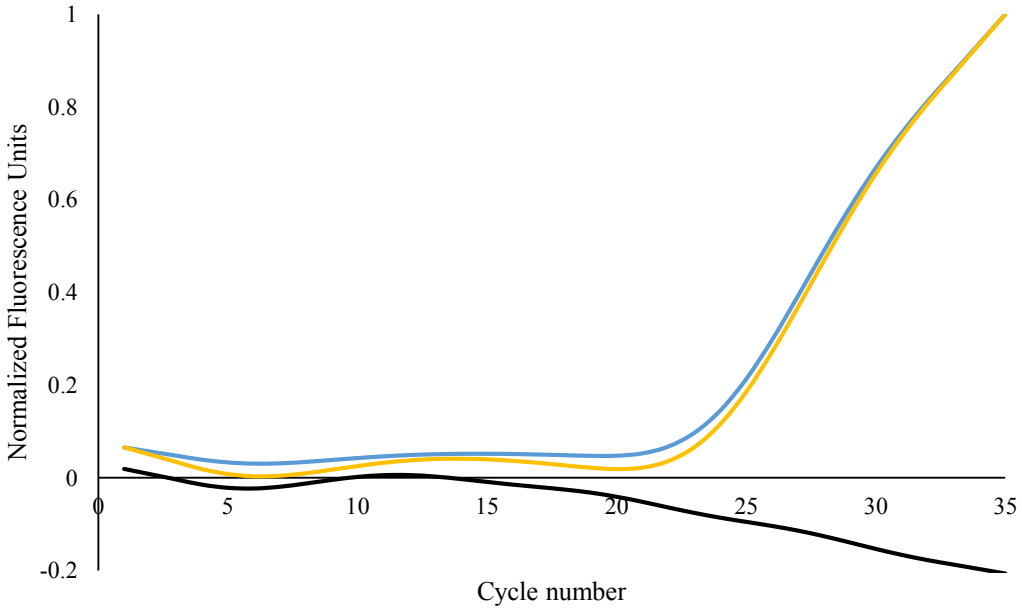

Figure 4.26: Assay optimization, 30 s denaturation, 30 s annealing, 5 s extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

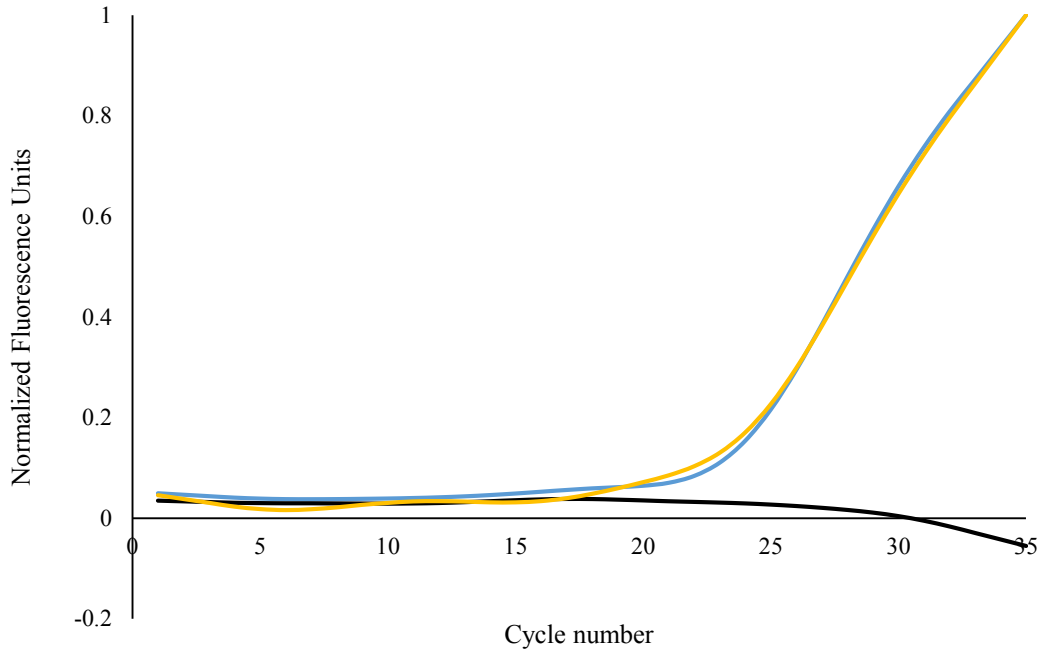

Figure 4.27: Assay optimization, 30 s denaturation, 30 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

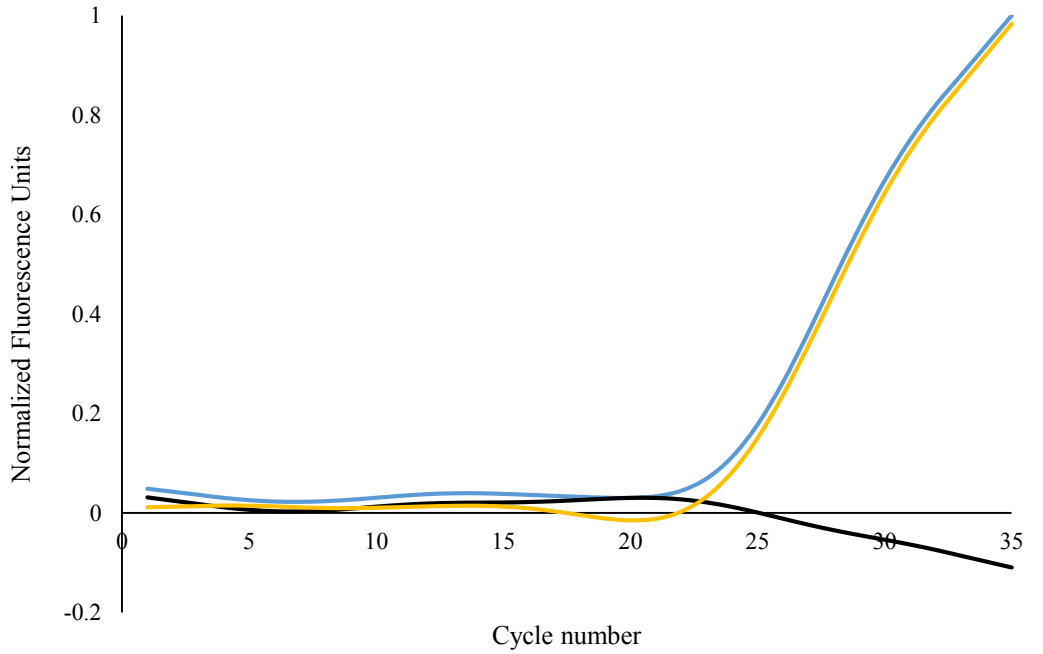

Figure 4.28: Assay optimization, 15 s denaturation, 30 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

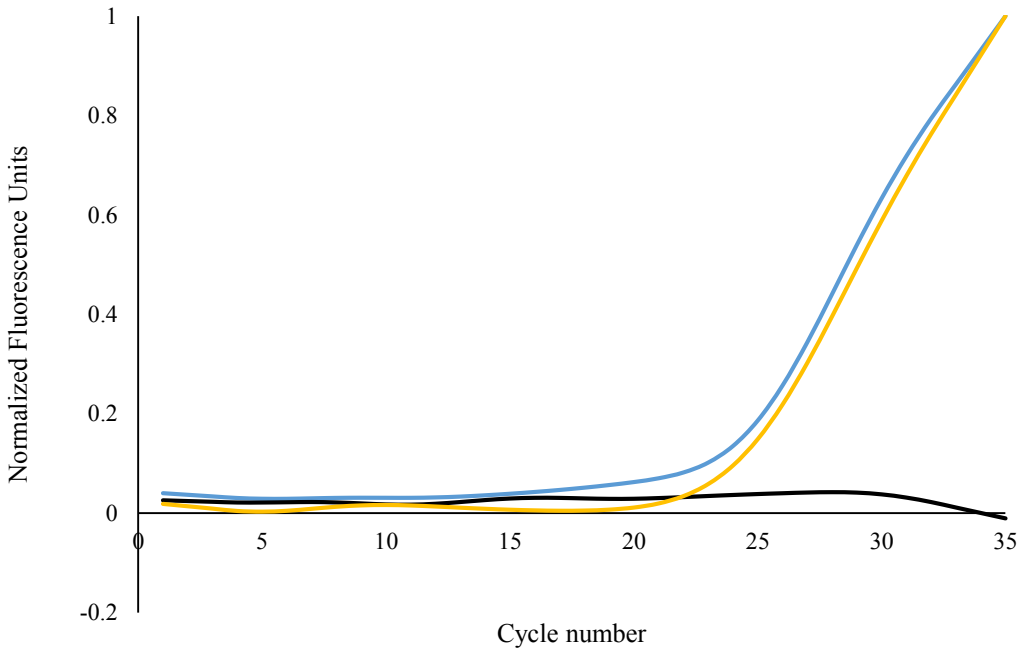

Figure 4.29: Assay optimization, 10 s denaturation, 30 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

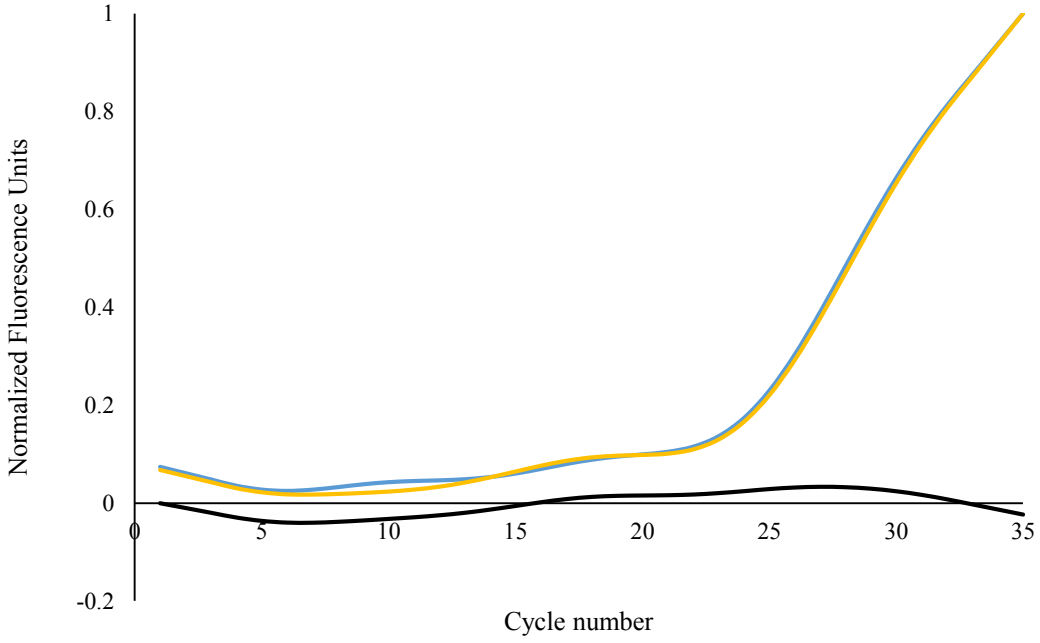

Figure 4.30: Assay optimization, 5 s denaturation, 30 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

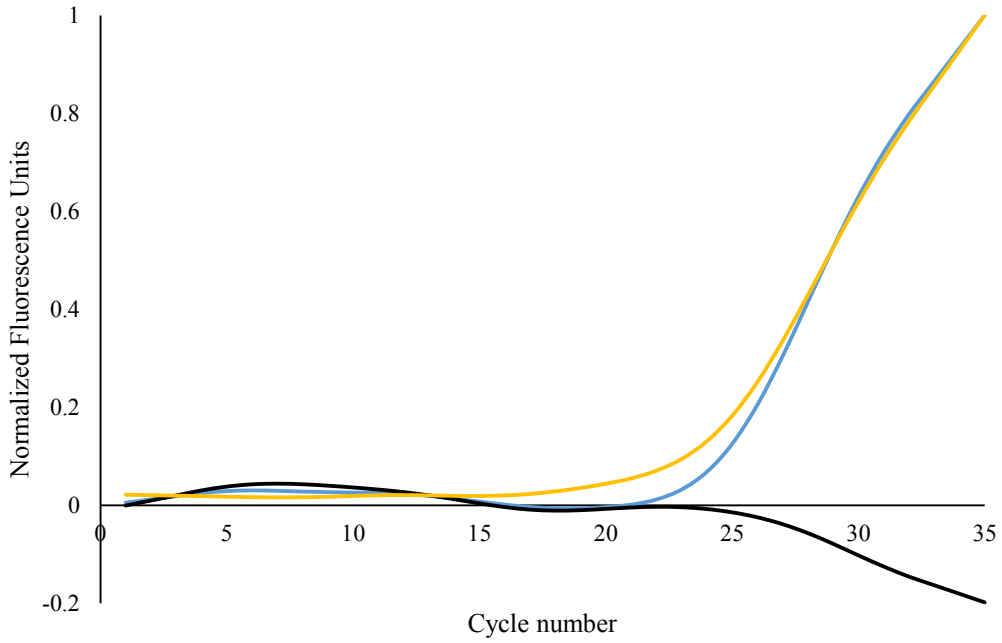

Figure 4.31: Assay optimization, 5 s denaturation, 15 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

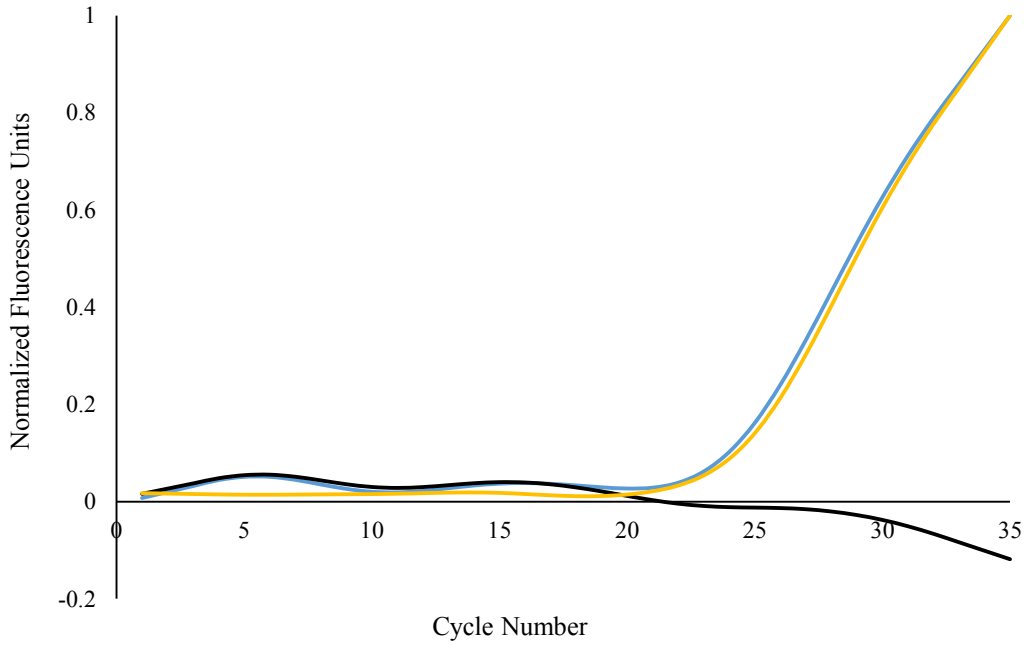

Figure 4.32: Assay optimization, 1 s denaturation, 15 s annealing, no extension Blue line represents T probe, yellow line represents C probe, black line represents NTC.

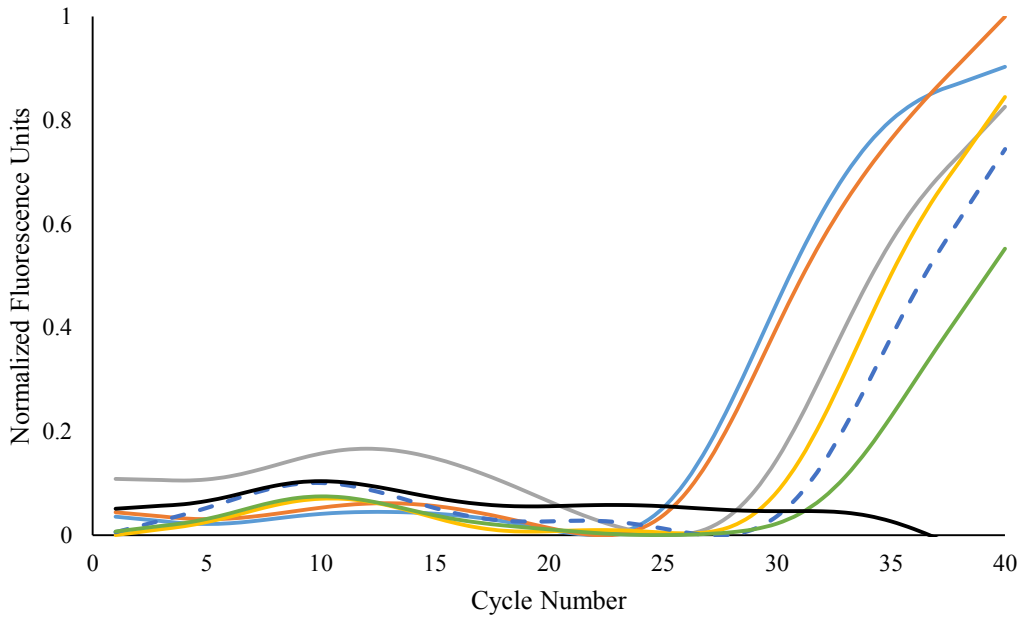

Figure 4.33: Initial sensitivity test, T probe, after optimization
Solid blue line is 20 ng, orange line is 10 ng, grey line is 5 ng, yellow line is 1 ng, dashed blue line is 0.5 ng, green line is 0.25 ng, black line is NTC. Can also be read from left to right as 20 ng, 10 ng, 5 ng, 1 ng, 0.5 ng, 0.25 ng, with NTC at bottom.

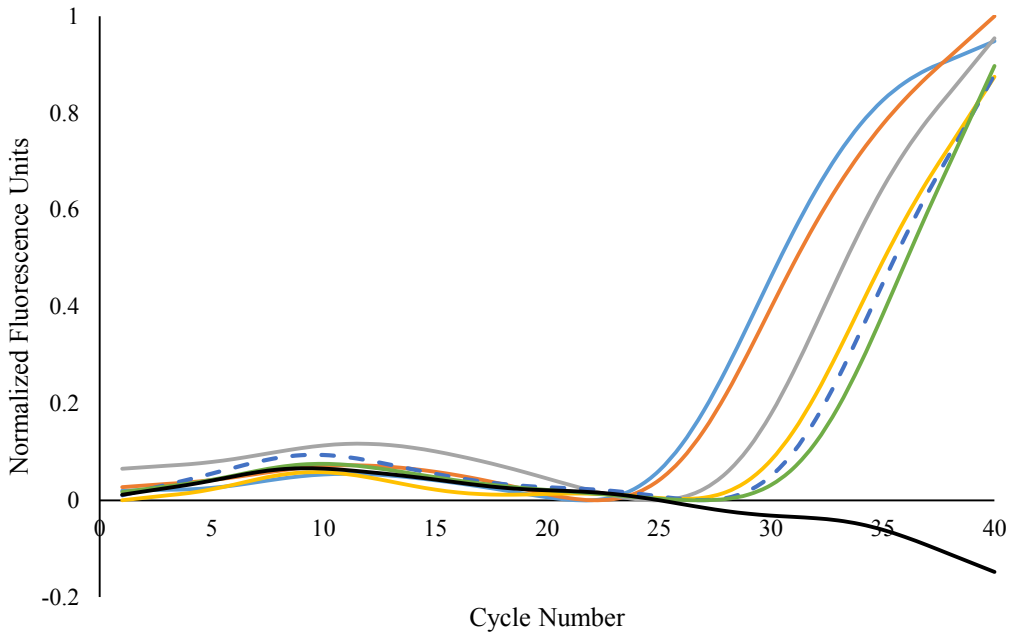

Figure 4.34: Initial sensitivity test, C probe, after optimization
Solid blue line is 20 ng, orange line is 10 ng, grey line is 5 ng, yellow line is 1 ng, dashed blue line is 0.5 ng, green line is 0.25 ng, black line is NTC. Can also be read from left to right as 20 ng, 10 ng, 5 ng, 1 ng, 0.5 ng, 0.25 ng, with NTC at bottom.

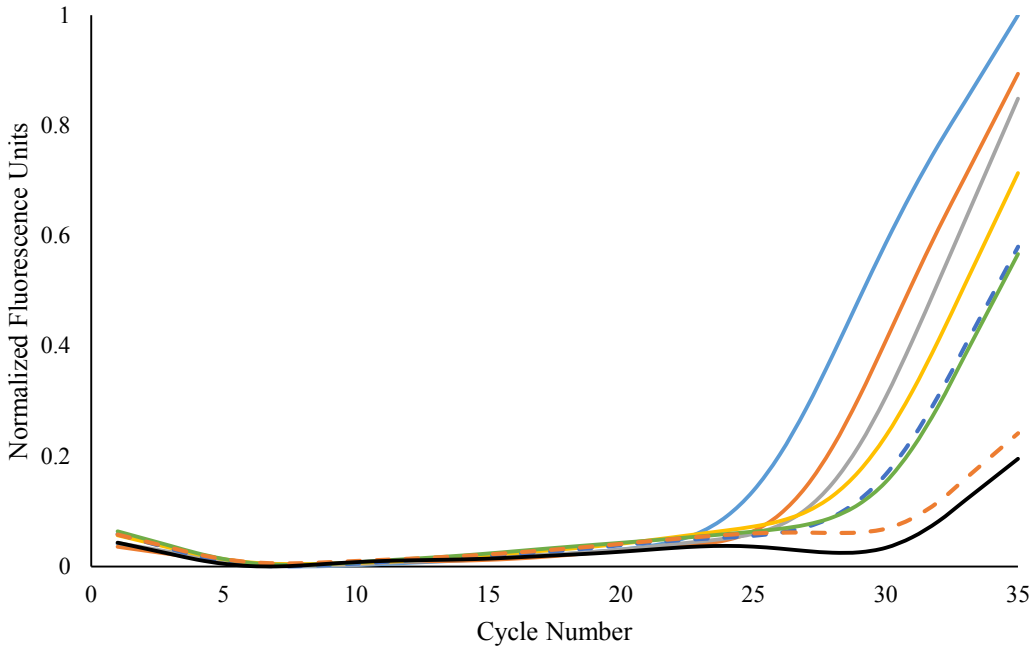

Figure 4.35: MB concentration test, T probe, 10 ng to 0.1 ng, 200 nM MB Solid blue line is 10 ng, orange line is 5 ng, grey line is 2 ng, yellow line is 1 ng, dashed blue line is 0.5 ng, green line is 0.25 ng, dashed orange line is 0.1 ng, black line is NTC. Can also be read from left to right as 10 ng, 5 ng, 2 ng, 1 ng, 0.5 ng, 0.25 ng, 0.1 ng with NTC at bottom.

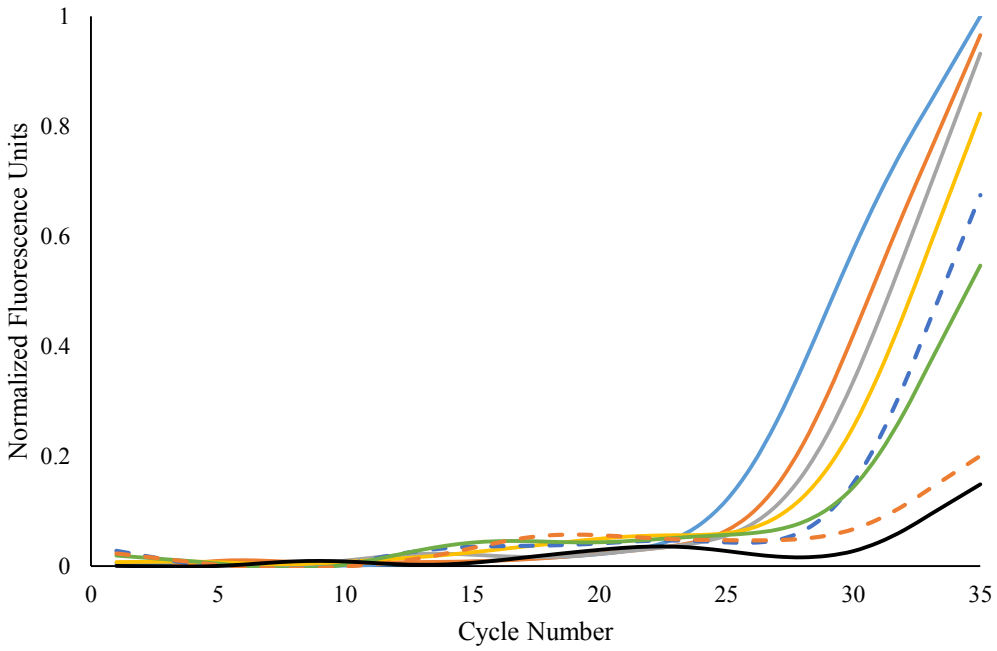

Figure 4.36: MB concentration test, C probe, 10 ng to 0.1 ng, 200 nM MB Solid blue line is 10 ng, orange line is 5 ng, grey line is 2 ng, yellow line is 1 ng, dashed blue line is 0.5 ng, green line is 0.25 ng, dashed orange line is 0.1 ng, black line is NTC. Can also be read from left to right as 10 ng, 5 ng, 2 ng, 1 ng, 0.5 ng, 0.25 ng, 0.1 ng with NTC at bottom.

4.4.7: Buffer development

Over the course of the project, numerous false positives were observed even after the development of a replacement primer. This mainly occurred in a direct sample. The sample preparation methods at the time involved using either NFW or 1X TE as a collection buffer, and upon review it was determined the buffer itself was the cause. Figure 4.37, Figure 4.38, Figure 4.39 and show the results for both the T probe and C probe MBs, where the sample is known to be homozygous T at this site and should only show signal response for the T probe, with no signal response for C probe. As can be seen in Figure 4.38, false positive results are present using a standard 1X TE buffer. However, Figure 4.40 does not show a false positive result. A small exponential growth phase is visible, but does not rise appreciably above the background and is not indicative of a positive result for this allele. Figure 4.38 shows the results using 1X TE, while Figure 4.40 shows results using 0.1X TE. There was concern, however, that diluting the buffer system by a factor of ten would reduce its buffering capacity enough to cause pH changes upon addition of sample. The buffer system was modified such that the concentration of buffer was the same, while the concentration of EDTA was reduced by a factor of ten. This buffer was tested, showed no false positive results, and used throughout the remainder of the project.

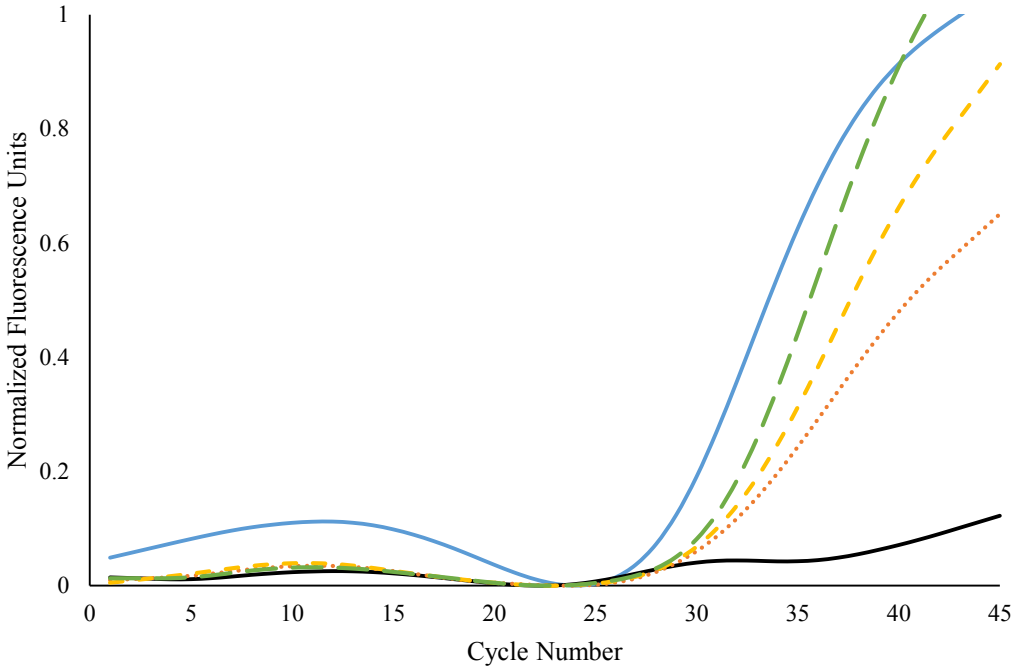

Figure 4.37: T probe genotyping, homozygous T sample, 1X TE sample prep Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a 2 µL addition of sample, yellow short dash line represents a 1 µL addition of sample, red dotted line represents a 0.5µL addition of sample.

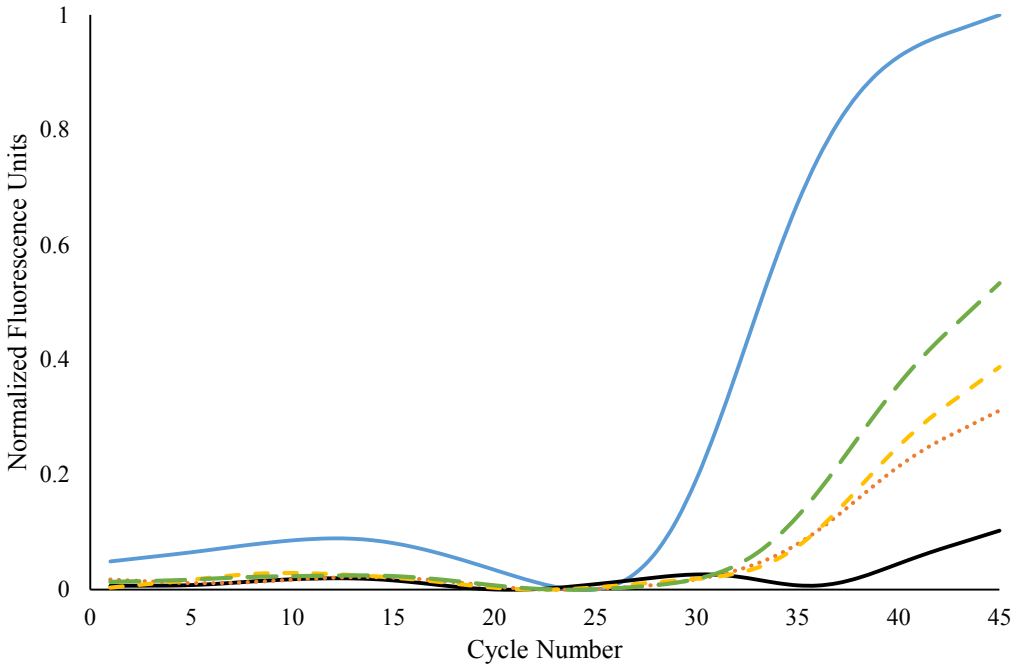

Figure 4.38: C probe genotyping, homozygous T sample, 1X TE sample prep Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a 2 µL addition of sample, yellow short dash line represents a 1 µL addition of sample, red dotted line represents a 0.5µL addition of sample.

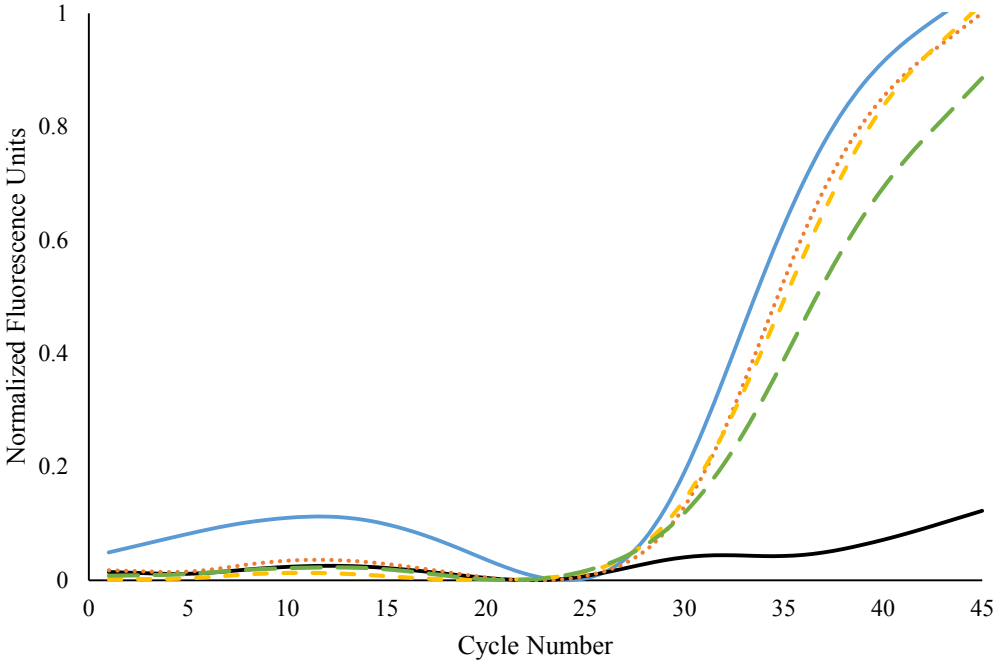

Figure 4.39: T probe genotyping, homozygous T sample, 0.1X TE sample prep Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a 2 µL addition of sample, yellow short dash line represents a 1 µL addition of sample, red dotted line represents a 0.5µL addition of sample.

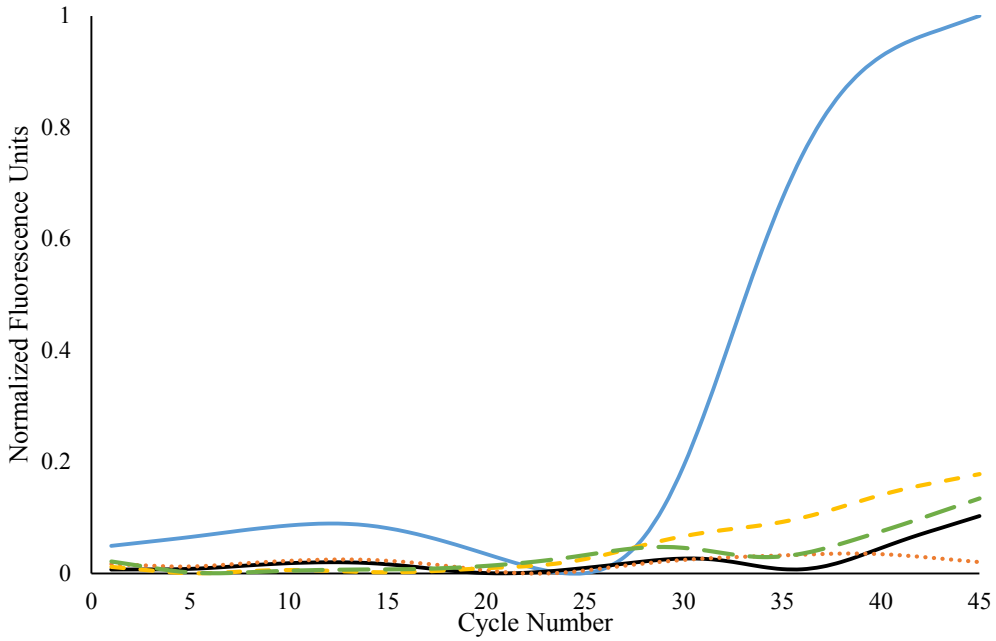

Figure 4.40: C probe genotyping, homozygous T sample, 0.1X TE sample prep Blue solid line represents positive control, black solid line represents NTC, green long dash line represents a 2 µL addition of sample, yellow short dash line represents a 1 µL addition of sample, red dotted line represents a 0.5µL addition of sample.

4.5: Phase IV: Allelic discrimination using MB-DPCR assays

4.5.1: Discrimination based on a 3-step format

The final phase of the project was to use the assay developed and test direct samples. Initially, a standard 90 second cycling profile was used with a single set of samples from a known T/T homozygote in order to test the ability to distinguish heterozygosity under direct conditions. These are shown as Figure 4.41 and Figure 4.42.

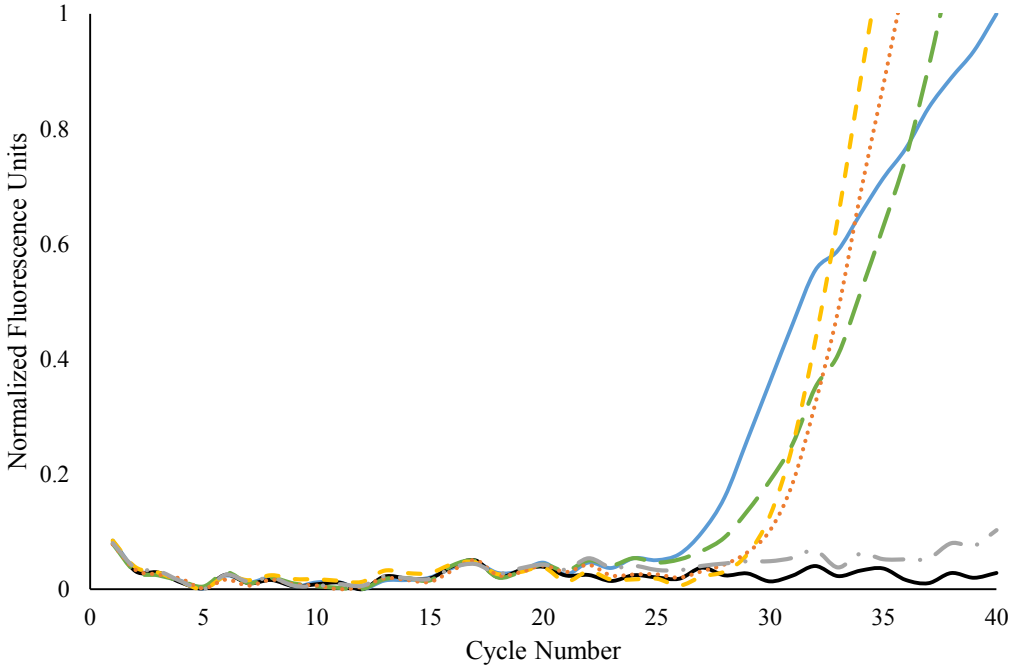

Figure 4.41: Verification 3-step protocol of amplification of direct samples, T homozygote, T probe
Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample, dashed-dotted grey line represents blood sample

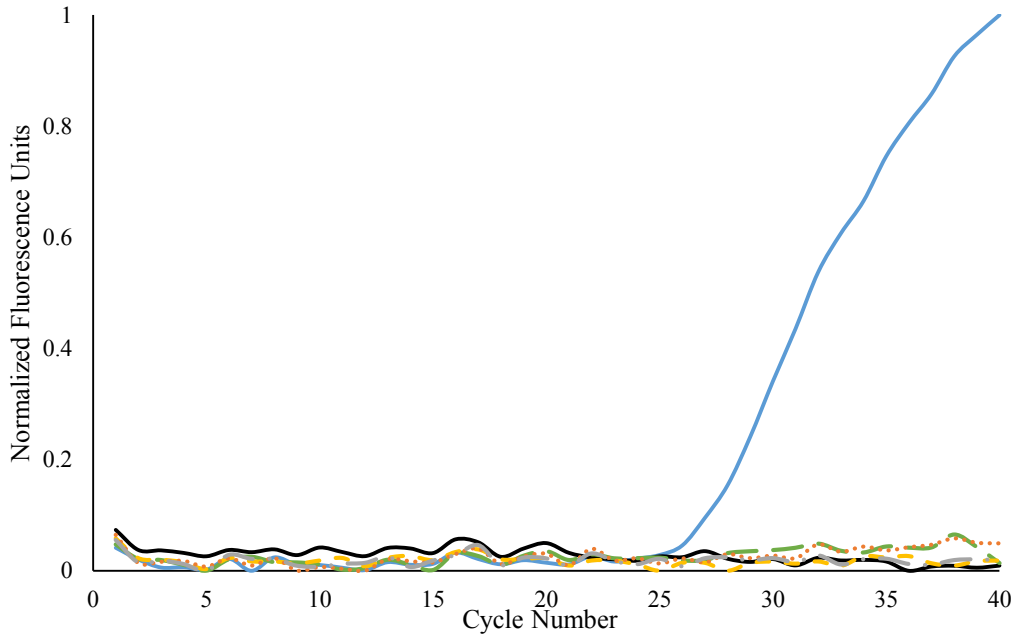

Figure 4.42: Verification 3-step protocol of amplification of direct samples, T homozygote sample, C probe
Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample, dashed-dotted grey line represents blood sample

4.5.2: Discrimination based on a 2-step format

The assay was then tested in a 2-step format following the successful discrimination of the homozygous T/T sample using a standard 90 second profile. New samples were collected from the same homozygous T/T individual and prepared nominally. The assay was conducted using a 5 second denaturation followed by a 30 second annealing phase, with optimized concentrations of MBs. Upon initial success, as noted in Figure 4.43 and Figure 4.44, all genotypes were tested using this assay. These can be seen as Figure 4.45 through Figure 4.50.

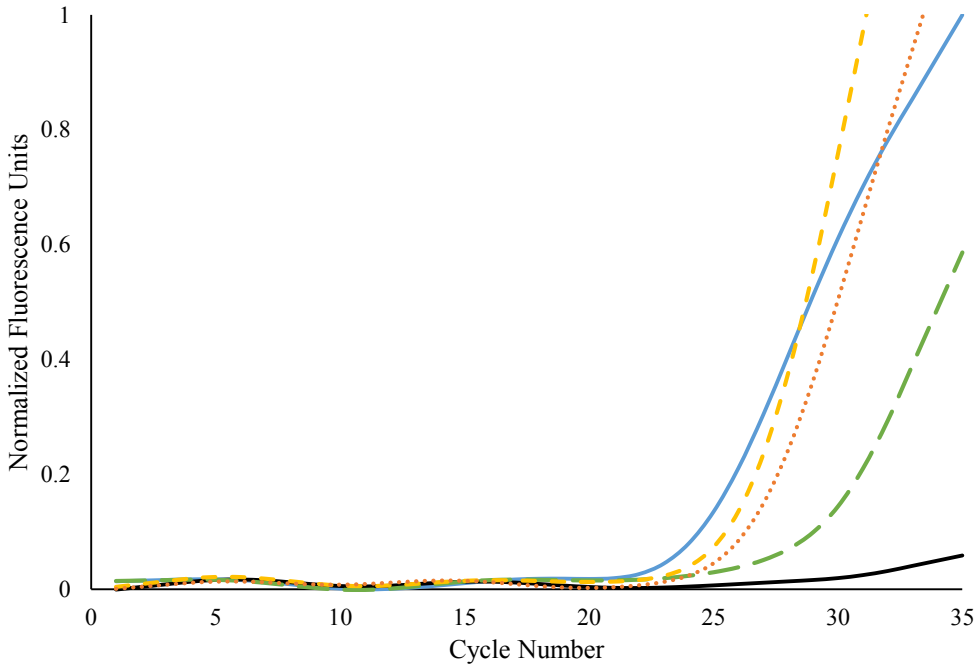

Figure 4.43: Initial allelic discrimination test, 200 nM MB, T probe, known homozygous T sample, 2-step protocol
Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

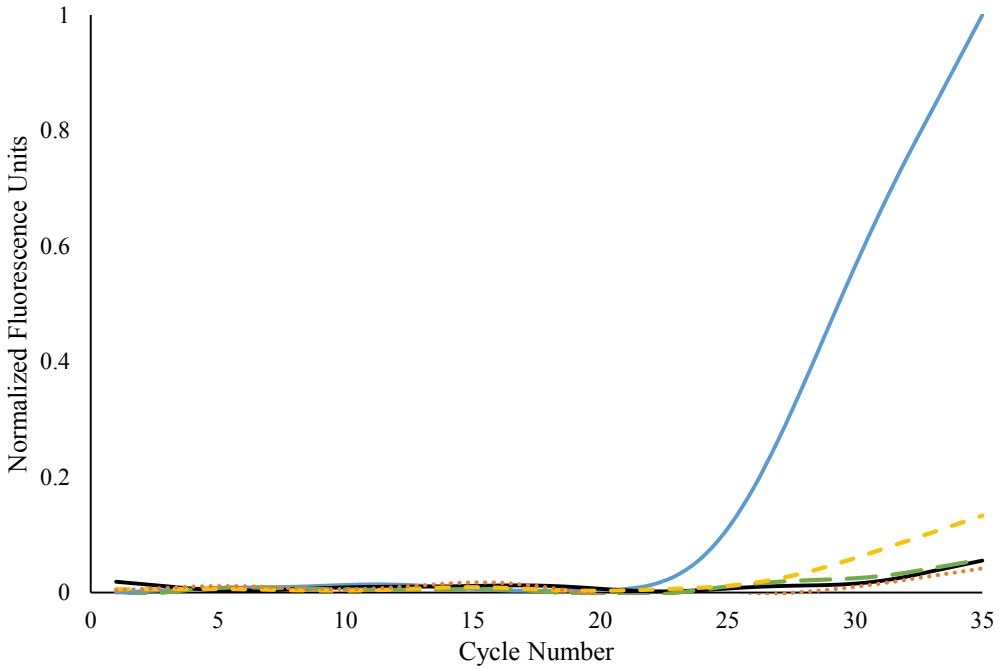

Figure 4.44: Initial allelic discrimination test, 200 nM MB, C probe, known homozygous T sample, 2-step protocol
Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

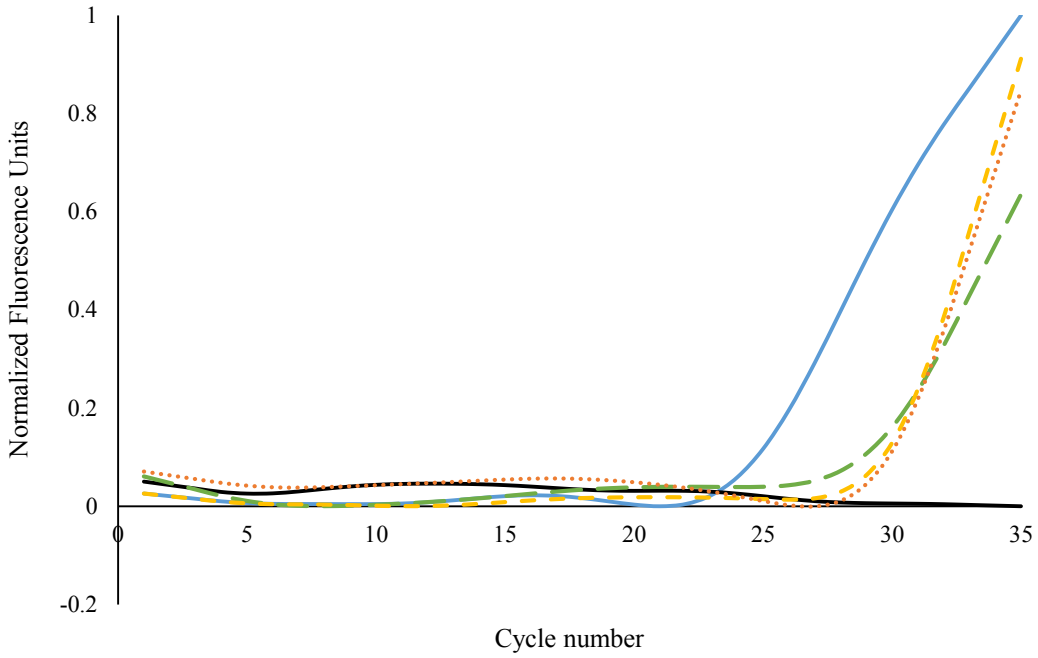

Figure 4.45: Final allelic discrimination test, 2 step, homozygous T, T probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

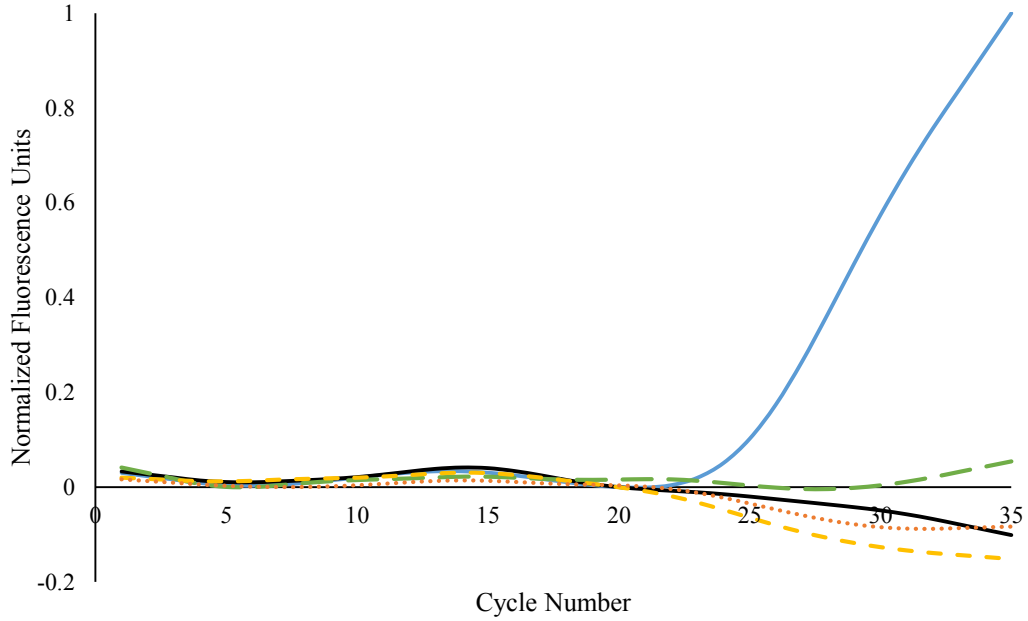

Figure 4.46: Final allelic discrimination test, 2 step, homozygous T, C probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

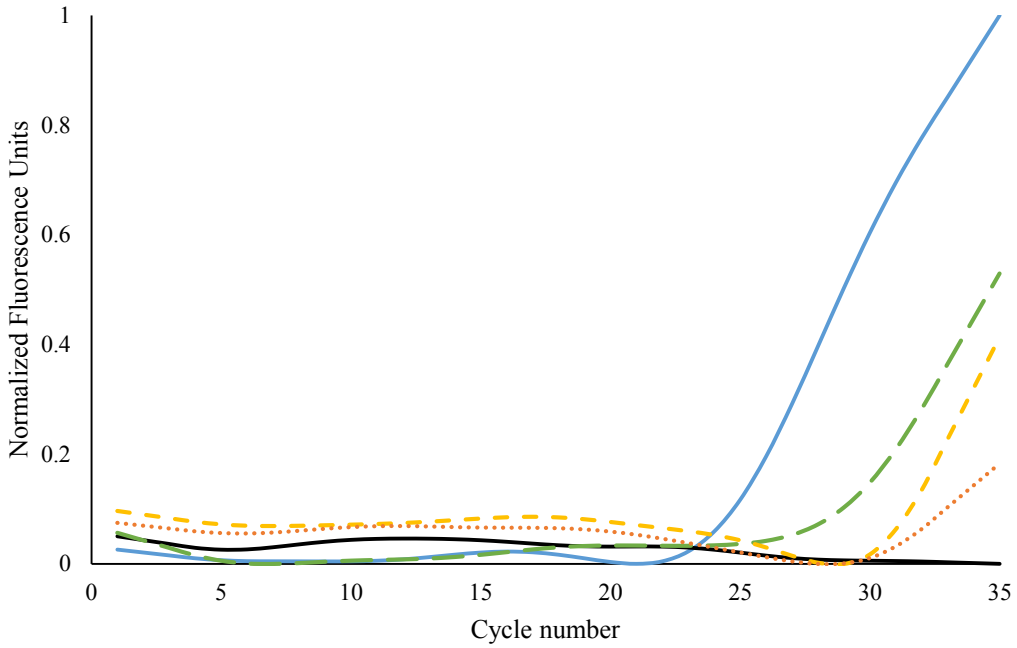

Figure 4.47: Final allelic discrimination test, 2 step, heterozygote, T probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

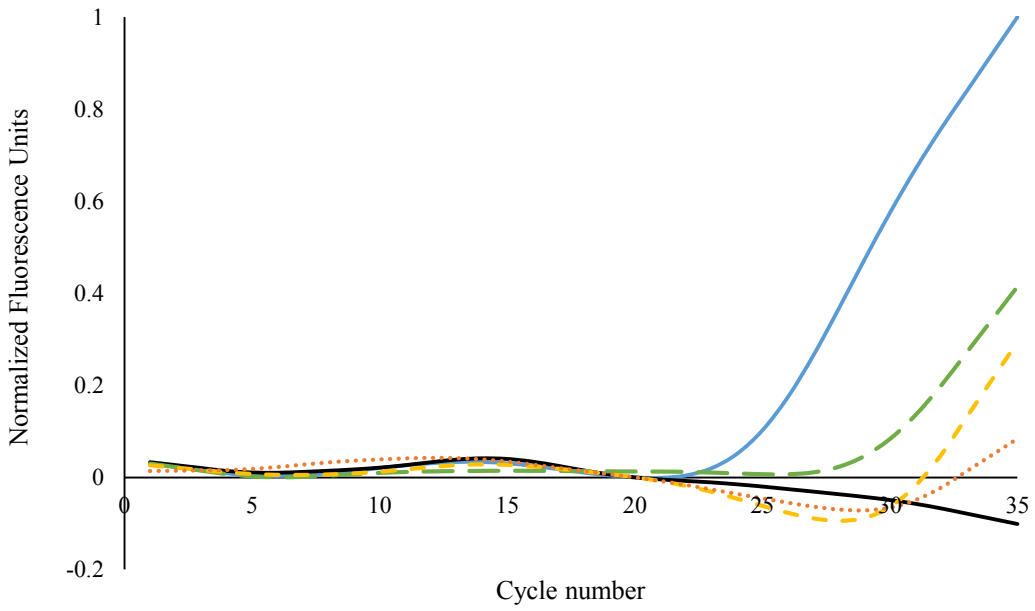

Figure 4.48: Final allelic discrimination test, 2 step, heterozygote, C probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample.

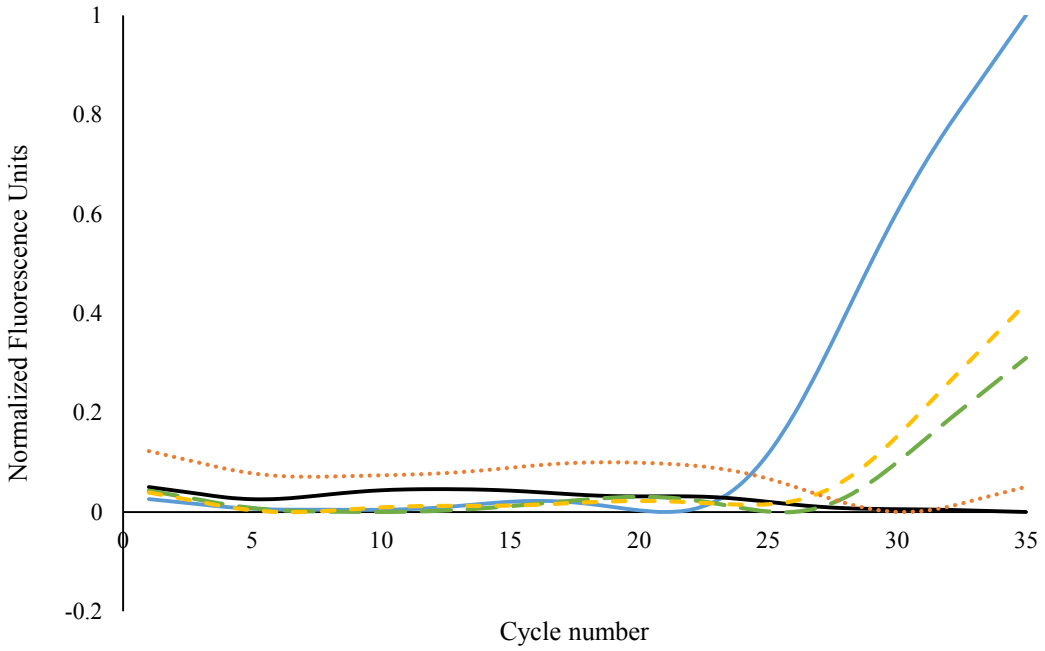

Figure 4.49: Final allelic discrimination test, 2 step, homozygous C, T probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample

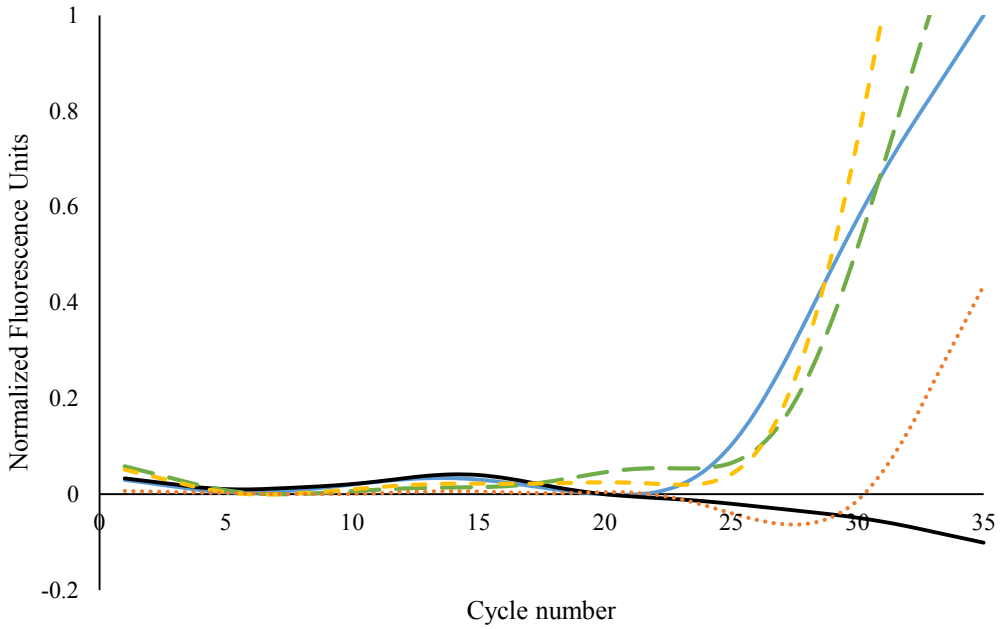

Figure 4.50: Final allelic discrimination test, 2 step, homozygous C, C probe Solid blue line represents positive control, solid black line represents NTC, long-dashed green line represents hair sample, short-dashed yellow line represents saliva sample, dotted red line represents buccal swab sample

4.6: Design of new MB

4.6.1: Software predictions and theoretical design

The false positive results from the homozygous C/C sample prompted the need for a higher-dissociating MB. Since 48°C is the highest the current beacons can be used, and this is an extremely low temperature for PCR, many nonspecific fragments were produced. This can be seen in Figure 4.21: Melt analysis of qPCR comparison of REV100 and REV158 primers after 48°C annealing. The identities of the nonspecific fragments are unknown, and could cause nonspecific hybridization of the beacon. This nonspecific hybridization can cause false signal generation.

A new MB was designed with a dissociation temperature of roughly 60°C, the same annealing temperature of the current primers. This closer match ensures better compatibility. In order to do this, the probe region was selected first. This was extended to 34 bp from the original 21 bp. This additional length of the probe region allows the beacon to remain hybridized and stable at increased temperatures. The predicted temperature is 58.0°C, with the data of the probe region shown as Figure 4.51.

The thermal profile of the current MBs, Figure 4.2 and Figure 4.3, shows that the current stem design would not exhibit effective quenching at the desired annealing temperature; additionally, the stem sequence was about half-complementary as seen in Figure 3.2 and Figure 3.3, which would greatly reduce the thermodynamic competition of the stem formation. A new stem was designed to correct for this. Upon stem design, however, Oligo7 predicted that the probe region would become self-complementary. The additional bases forming a stem places complementary nucleotides in the probe domain in proximity such that they

are capable of weakly associating. This weak association is actually beneficial in this case, as it increases the thermodynamic competition of the stem, further aiding SNP discrimination. With SNP discrimination, as discussed in 2.12.4: SNP Profiling, there is an increased requirement for competition due to the similarity of the alleles. An image of the secondary structures and their respective predicted T_M values is shown as Figure 4.52.

With the probe and stem sequences designed, the entire beacon was checked with nBLAST and found to have an expect value of 10^{-8}, with the next closest expect value being 2.5. This means the probe is extremely specific to the target sequence, and chances of random alignment in any other known genome is negligible. The alignment for the new MB is shown as Figure 4.53.

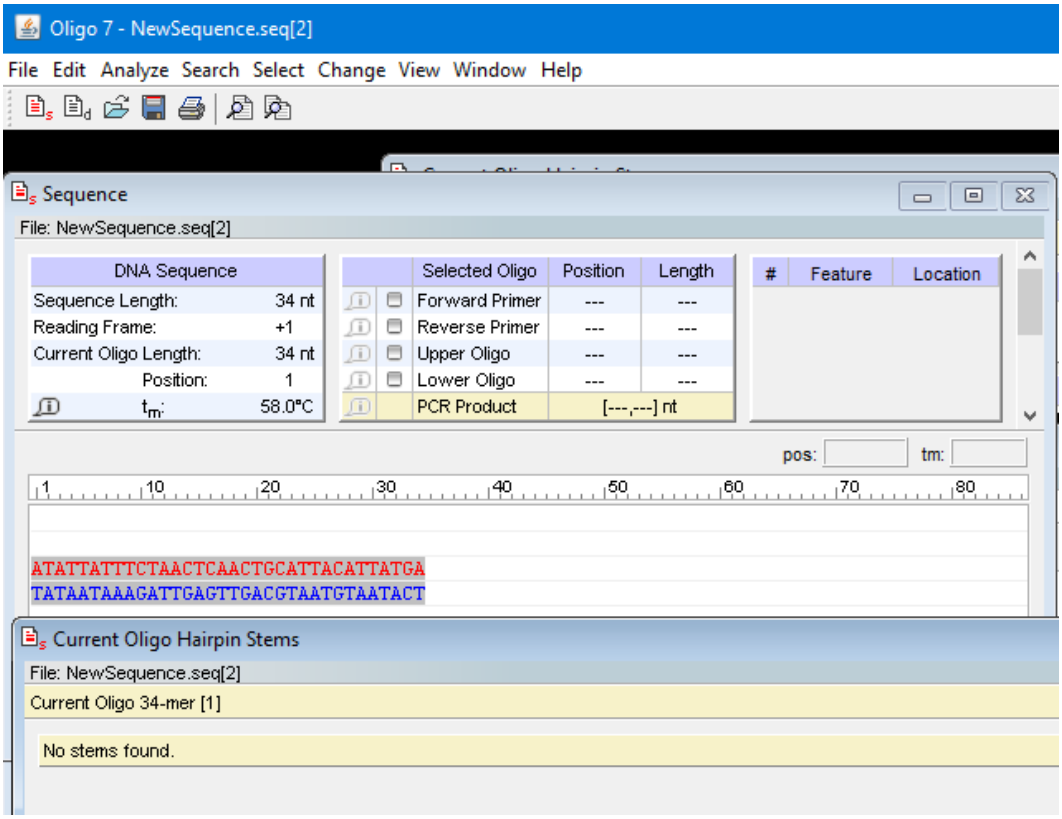

Figure 4.51: Prediction of T_M and secondary structure formation of new C probe, Oligo7

This figure depicts the probe segment only of the newly designed MB and its theoretical T_M. This also shows that there are no predicted secondary structures formed by the probe sequence.

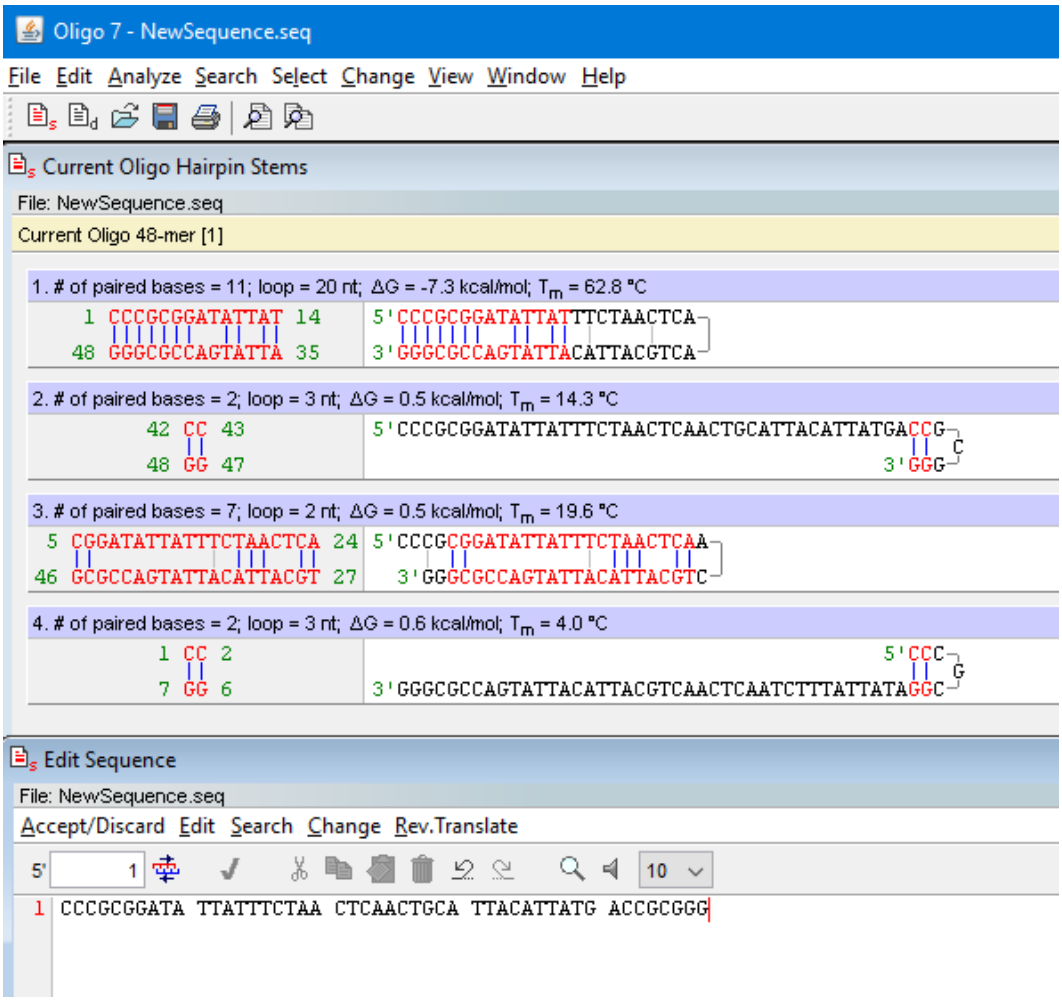

Figure 4.52: Predicted secondary structures of new C MB with stem, Oligo7
This figure shows the sequence and predicted secondary hairpin structures formed by the full MB with the stem included. The resulting stem melting temperatures are also predicted.

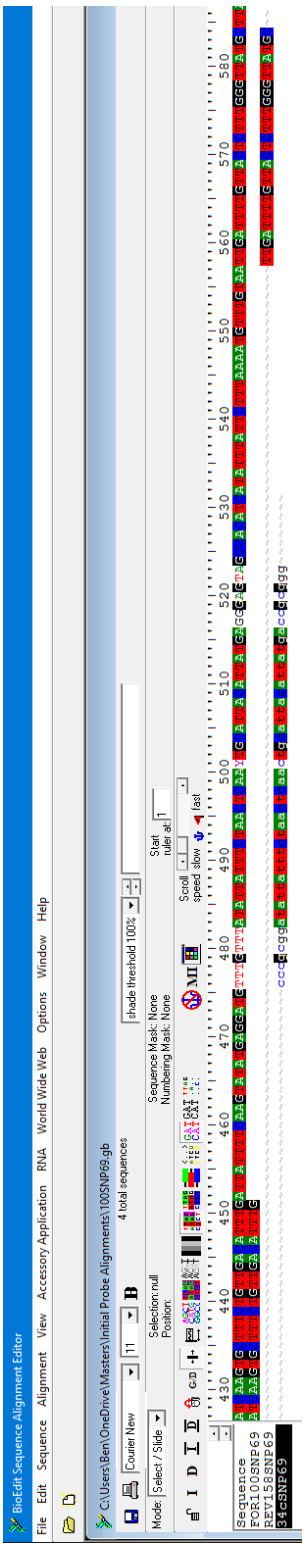

Figure 4.53: Alignment of new C MB, stem included, with FOR100 and REV158 primers

Letters represent one letter abbreviations for DNA bases. Bases with colored blocks match the target sequence; letters without blocks do not match target sequence.

4.6.2: Thermal profile and initial testing of new MB

Due to the inherent cost of MBs, only one of the two targets was designed and ordered for initial testing. Complementary oligonucleotides and the new MB were ordered from Sigma Life Science (Missouri, USA). The complementary oligonucleotides were 50 bp, with one an exact match and another matching the opposite allele. These were ordered as 50 bp of the target sequence in order to fully cover the stem and simulate as best as possible the performance when in contact with the amplicon, as the new MB has a total length of 48 bp, stem and probe included. The thermal profile was generated using 1.000 μM MB, 10.00 μM complementary (or mismatch) strand, in 8.00 mM Tris-HCl, 280 nM $MgCl_2$, pH 8.0. The resulting thermal profile is shown as Figure 4.54.

An initial verification assay was tested upon determination of the thermal profile of the new MB. The thermal profile showed the lowest normalized fluorescence of a mismatch strand at between 60-65°C, with 65°C being essentially zero normalized fluorescence. The annealing temperature of the primer set is known to perform well up to 63.7°C, as seen in Figure 4.19: Gel electrophoresis image of REV158 primer annealing temperature gradient. With this in mind, the assay was tested at 60°C for performance under verification conditions. This assay used a standard 90 second cycling profile, a known homozygous C/C sample to screen for amplification, a known homozygous T/T sample to screen for nonspecific signal generation, and an NTC to screen for contamination and random coil formation resulting in signal in the absence of any DNA. This is shown as Figure 4.55. All DNA samples were purified using BioLine's Isolate II Genomic Purification Kit prior to analysis.

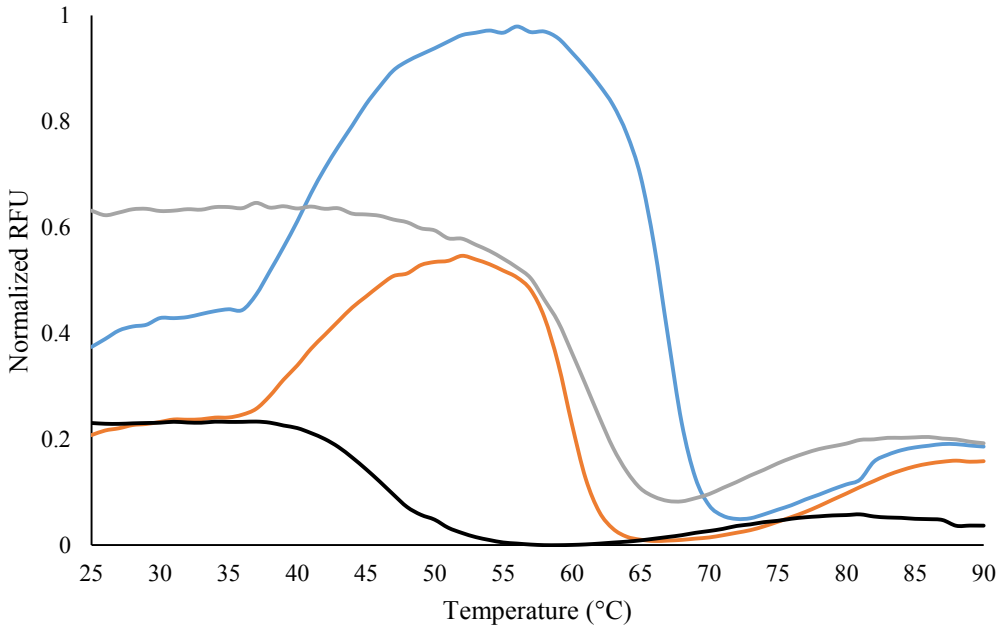

Figure 4.54: Thermal profile comparison, new MB and original C probe
Blue line represents new MB with complementary strand, orange line represents new MB with mismatch strand, grey line represents original C probe and complementary strand, black line represents original C probe and mismatch strand.

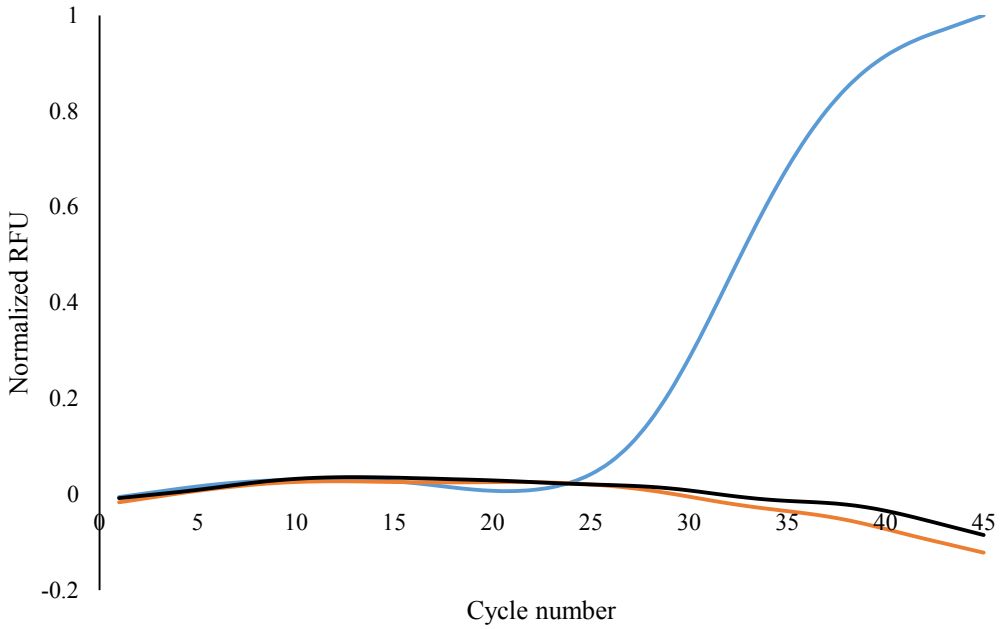

Figure 4.55: Initial verification of new MB, standard conditions
Blue line represents homozygous C/C sample, orange line represents homozygous T/T sample, black line represents NTC.

CHAPTER 5: DISCUSSION

CHAPTER 5: DISCUSSION

5.1: Introduction

This chapter discusses results of the project. The results are explained in detail, along with their implications. As described in Chapter 1: Introduction, the specific aims of this project were to develop a protocol capable of allelic discrimination using MBs, apply that protocol in an unpurified, direct PCR format, and optimize the protocol such that its time is as short as possible without compromising results. The project was divided methodically, so as to ensure success.

5.2: Phase I: Verification of amplification and detection

5.2.1: Verification of FOR100 and REV100 primers

Verification of amplification using direct samples for the first set of primers was crucial. This ensured the PCR was amplifying DNA, and provided a means to ensure the PCR fragment was the expected size. The FOR100/REV100 primer set produces a 100 bp fragment, and as Figure 4.18 shows, all samples were amplified using a direct protocol under verification conditions. The NTC did not show a 100 bp fragment, confirming a lack of contamination. All lanes showed amplification at roughly 50 bp, confirming the presence of unintended PCR products.

The REV100 primer actually overlaps the stem portion of the MBs used in this assay, causing the formation of a dimer. It is unclear exactly what the fragment at 50 bp is, but it is clear that this is an erroneous, unintended product as it is show in the NTC as well. Identification of nonspecific fragments could be confirmed by mass spectrum or RFLP analysis, but is unnecessary. Instead, this necessitated the

development of an additional reverse primer to pair with the existing FOR100 primer. While the gel image is not enough to confirm the exact identity of the unintended fragment, subsequent false-positives in the real-time PCR were attributed to the primer annealing to the MBs.

This nonspecific priming is also indicative of inefficiency in priming, which can result in PCR inhibition and a reduction of overall efficiency. The reduction of overall PCR efficiency would cause increased C_T values, leading to problems with low copy number (LCN) samples. The development of a primer positioned such that the MB has no contact with no guanine residues within 3 bp of the 5' end of the MB allowed an increased PCR efficiency.

5.2.2: Thermal profile of MB probes

The thermal profile of the MB probes was determined prior to the study (Fredericks, 2014). These profiles (shown as Figure 4.2: Thermal profile of MB probes using C probe complement and Figure 4.3: Thermal profile of MB probes using T probe complement) were generated using an assay consisting of the MB and a fold excess of a synthetic oligonucleotide matching the target and extending past the stem. The reaction vessel was then heated to melt any dimers and cooled to 25°C, followed by slowly heating up to 90°C. The temperature was held constant during signal acquisition. The assay measured the normalized fluorescence of each probe when in contact with a target or non-target oligonucleotide in order to simulate the mechanics of the probe when exposed to gDNA.

Referencing Figure 4.2, the normalized fluorescence of the C probe is much higher than the T probe, because the C complement was used in the assay. This is expected, as hybridization is greatly reduced in the presence of a non-target. A similar trend is seen in Figure 4.3, as the T probe shows higher fluorescence than

the C probe when using the T complement. These curves were used to identify an optimal annealing temperature for the cycle. At roughly 48°C, there is a large relative difference in fluorescence in both assays. The MB and non-target oligonucleotide show fluorescence at a similar temperature to the control. This shows that at 48°C, the probes will selectively anneal. Selective annealing prevents non-target signal generation. This temperature is still low enough such that the probes will anneal to their targets, enabling signal generation. This temperature, 48°C, is extremely low for an annealing temperature in a PCR. This can cause non-specific priming, leading to greatly reduced PCR efficiency and signal generation. For the sake of this study, this pair of MBs is sufficient, but an improved set is necessary for future studies.

5.2.3: Verification of MB assay

It was important to verify the performance of the MBs individually before testing in conjunction with each other. Depending on the intensity of the signal, spectral overlap between the fluorophores can cause signal carryover; this artifact produces a seemingly positive response to a negative sample due to the intensity of a fluorophore at the detection wavelength. Given the emission spectrum of FAM, there is a possibility of significant emission at 555 nm, despite its 520 nm λ_{max}. If significant emission occurs at that wavelength, the instrument will not be able to differentiate signal due to a FAM-labeled probe from that of a HEX-labeled probe, preventing allelic discrimination. In this case, a homozygote could appear heterozygous. The primary goal of this verification assay was to ensure the proper fluorophore was conjugated to the probe and the probes were performing adequately, i.e. efficient quenching and limited noise due to overlap. Referencing Figure 4.4 through Figure 4.7, positive controls responded only at the desired

wavelength, with only marginal response from negative controls at any wavelength. This validates the performance of the MBs separately, and allowed studying the MB pair in a multiplex format.

Once verified as working correctly, the MBs were tested in a probe multiplex format. The stems of these probes were identical, which may cause nonspecific expansion of the probes resulting in increased noise. As seen in Figure 4.8 and Figure 4.9, however, no increased noise was observed. The data supports a lack of increased noise or false positives. The negative control still did not show any signal, while the positive control responded on both channels. This response is expected, as the positive control is heterozygous. A heterozygous control was used to ensure both MBs were working.

5.2.4: Screening of genotypes

The MBs had been validated with a heterozygous sample, but no homozygous samples had yet been tested. Samples were tested from a number of consenting, college-age adults for genotypes for further use in the study according to the IRB for the study (IRB 16-033). This also allowed the investigation of the ability of this MB pair to discriminate profiles in real time. Three individuals were identified as research subjects for the study, with one each as a homozygous C/C, heterozygous C/T, and homozygous T/T. As shown in Figure 4.10 and Figure 4.11, the MB pair is capable of discriminating genotypes in real-time from purified DNA samples in a standard 90 second cycling profile. This discrimination is confirmed through corresponding positive and negative signal generation in the presence of respective genotypes. Homozygosity was confirmed upon the observation of significant signal from one MB, with no response from the other. Significant signal is defined as a resulting signal, after normalization, at least three times the

maximum of the background. The background is the average of the normalized fluorescence of the first 15 cycles, prior to the takeoff point. These profiles used were noted, and the same individuals agreed to participate in the study. Individuals donated samples as requested, per IRB 16-033.

5.2.5: Summary of verification and detection

Multiple assays were performed in order to ensure all components were performing as intended. This phase of the project did not so much accomplish novel work as it did lay the foundation for testing a novel application. The primers were tested first; without working primers, no amplification can occur. Amplification is essential for signal generation, so primers are the most crucial component of the assay. The thermal profile of the MBs were generated in order to determine the optimal detection temperature, which was used as the annealing temperature of the PCR cycling. This data was then used to develop an initial assay for the screening of genotypes. Genotypes can be confirmed through Sanger sequencing.

5.3: Phase II: Comparison and selection of PCR master mix

5.3.1: Kits considered

As discussed in 2.11: Challenges of Direct PCR, there are many obstacles when using direct PCR. Specialized kits have been commercially developed for direct PCR, however these are only for amplification. These kits are not designed with any optical considerations. Real-time applications are generally avoided with direct PCR. Real-time PCR tends to be quantitative in nature, and consistent results with low standard deviation would be extremely difficult to obtain when used in a

direct format, given the concentration of inhibitors. The project took care to investigate many different kits before selecting a final kit for further investigation.

5.3.2: ThermoScientific Phusion kits

The complexity of direct PCR made the use of the Phusion kits an attractive option. These kits are engineered to successfully amplify direct samples with high fidelity. This allows the reliable amplification of even the most challenging samples, such as blood. The fidelity of these kits is due to their exonuclease activity. Exonucleases are enzymes which are capable of recognizing a base-base mismatch in a sequence and cleave the terminal base of the nascent strand. In the Phusion kits, the polymerase employed has innate exonuclease activity.

The annealing phase of the PCR cycle is used to hybridize the primers and probes to the template strand, and allow the polymerase to coordinate with the hybrids. It is then conceivable that, upon probe hybridization, a polymerase can coordinate with the MB-DNA hybrid and begin proofreading before polymerizing. Once the exonuclease domain reaches the stem—which is non-complementary to the target—the exonuclease will recognize a base-base mismatch.

The likely cause of incompatibility, then, is the cleavage of either the terminal nucleic acid of the stem along with the reporter or quencher, or cleavage of only the reporter or quencher. As seen in Figure 4.12: MB assay, T probe, Phusion kit and Figure 4.13: MB assay, C probe, Phusion kit, there is significant signal increase with each cycle. This rapid signal increase is likely due to the cleavage of one of the dyes, located at the termini of the MB. The polymerase is causing the MB to behave more like a TaqMan probe by destroying the MB during the assay. In MB assays, though, it is imperative that the MB remain intact. Thus,

the Phusion kits are incompatible with the project, which relies solely on MB assays.

5.3.3: Roche LightCycler 480 and BioLine SensiFAST comparison

5.3.3.1: C_T values

The most reliable method of comparing the sensitivity of two kits is the comparison of C_T values. This method relies on using a serial dilution of DNA concentration, typically determined through the use of a Nanodrop 2000—a specialized UV-Vis spectrometer. The C_T values are inversely related to the initial DNA concentration; as DNA concentration decreases, C_T increases. This is because it takes more cycles to replicate the smaller DNA concentration such that it is in roughly equivalent concentration to the MBs. When comparing kits, it is crucial to ensure replicate samples are used, and the samples are consistent. Variations in concentrations can easily skew results. Results obtained from the analysis of serially diluted samples enables the software to calculate the threshold of the data and produce a table of C_T values for the entire set.

Referencing Table 4.1: Average CT Values, Comparison of Roche and BioLine kits, the BioLine kit consistently had lower C_T values, indicating increased sensitivity over the Roche kit selected. Upon averaging replicate samples, the data showed significant differences between kits, with the BioLine kits performing roughly 5 cycles faster than the Roche kit. The BioLine kit is optimized to work on multiple instruments, while the Roche kit is designed specifically with the LightCycler instrument in mind. Thus, the Roche kit may be less suited for more conventional fluorometric thermocyclers.

5.3.3.2: Reliability comparison

Observation of the assay performance beyond C_T values is important as well. While C_T is the most reliable means of comparison, it is also important to test LOD and observe effects of LCN samples. Referencing Figure 4.14: Sensitivity comparison, Roche kit, T probe and Figure 4.15: Sensitivity comparison, Roche kit, C probe, note that Figure 4.15 does not show amplification at 2.5 ng/reaction. These reactions were performed with the same DNA source, a known heterozygote. The assay must be able to reliably report the genotype, a task which becomes increasingly difficult as the sample concentration decreases. A known issue with LCN samples is allelic dropout—an artifact caused by early, preferential amplification of one allele, leaving the other allele in such low concentration at later cycles that it is unable to be detected. Given that the BioLine kit did not exhibit allelic dropout at any concentration studied, while the Roche kit exhibited it as low as 0.25 ng/reaction, the BioLine kit is more reliable than the Roche for LCN samples.

5.3.4: Summary of kit comparison and selection

The Phusion kits are incompatible with MB assays even under purified conditions; without the use of a different polymerase, or the same polymerase without an exonuclease domain, these kits cannot be used to gain reliable data for this project. The Roche LightCycler 480 Probes Master was less sensitive and less reliable than the BioLine SensiFAST Probes Master. Ultimately, the BioLine SensiFAST Probes Master proved to be the most compatible with the assays performed.

5.4: Phase III: Direct MB-PCR assay development

5.4.1: Verification of amplification under direct conditions

The components of the project had been tested extensively up to this point. The primers, MBs, genotypes of samples, and kits were subjected to rigorous validation. Even with all of the validation work done, however, no data had been collected on the ability of the MBs to generate a profile without a DNA purification step. Thus, the third and final phase of the project began much like the initial steps of the project: standard PCR without real-time data generation, followed by gel electrophoresis for confirmation of amplification. Gel electrophoresis not only confirms the presence of DNA in a large quantity, but also estimates the produced amplicon size for comparison to the expected size. The fragment size expected was 100 bp, which is confirmed by Figure 4.18: Gel electrophoresis of direct PCR fragments, after amplification. The positive control, Lane 2, was performed with control gDNA and was known to work through confirmation in 5.2.1: Verification of FOR100 and REV100 primers above, and shows an amplified band at 100 bp. The direct sample lanes show the same band, confirming amplification of the correct locus.

All lanes, including the NTC, showed nonspecific amplification. This is highly undesirable in a MB assay, especially in a direct format. A direct sample is difficult to amplify on its own; nonspecific annealing will produce fragments which cannot be identified by the MBs in the assay, thus significantly lowering the sensitivity of the assay. Some fragments may complement the stem, leading to nontarget annealing and separation of the DAP, resulting in increased noise and a reduced SNR. An efficient primer does not show nonspecific annealing at the assay's operating temperature; since the gel shown in Figure 4.1 does show

nonspecific annealing, this is an inefficient primer at this temperature. The optimal annealing temperature of the MBs is 48°C, which explains the nonspecific annealing; however, a more efficient primer was necessary for this assay to continue.

5.4.2: Design of new primer set

The new primer was designed such that the fragment size shifted from 100 bp to 158 bp, and was named accordingly: REV158. This primer was first verified alongside the REV100 primer for comparison. The priming site was selected for its distance from the MB locus. The primer and MB should not overlap in an alignment. Complementary overlap can cause nontarget signal generation. The alignment is shown as Figure 3.3.

5.4.3: Comparison of new primer set

As shown in Figure 4.20, the REV100 samples both showed weak amplification bands at roughly 50 bp, whereas neither of the REV158 samples showed nonspecific amplification. The fragment size runs slightly smaller than 200 bp, confirming its approximate size. Note that this does not confirm the exact, intended sequence is amplified; this technique simply estimates the amplicon size is approximately correct, compared to what is theoretically expected. However, it can be reasonably assumed to be the correct fragment after development methods such as BLAST are taken into account. This primer set was tested virtually using NCBI's BLAST online tool to ensure specificity prior to ordering. Since specificity was confirmed computationally, and the fragment size is approximately the same

as what is expected, reasonable assurance of functionality and specificity of the primer set were demonstrated.

After it was confirmed that REV158 was a more efficient primer, it was subjected to DPCR testing. Since the then-current DPCR-MB protocol required a 48°C annealing temperature, the PCR was carried out using the verification protocol in Section **Error! Reference source not found.Error! Reference source not found.** with the sample preparation methods described in Appendix C: Summary of Sample Preparation. An image of the subsequent gel electrophoresis is shown as Figure 4.23. This image shows amplification of the correct fragment sizes in a purified control, unpurified hair, saliva, and buccal swab samples. Amplification did not occur in blood samples. The original FOR100/REV100 direct amplification verification was done with Roche's LightCycler™ 480 Probes Master as a kit, while this amplification was carried out using BioLine's SensiFAST™ Probes Master. The difference in results can be attributed to a difference in polymerase, rather than primers. Blood presents a considerable challenge to direct amplification of DNA, given the amounts and varieties of inhibitors present. As such, the polymerase in BioLine's SensiFAST™ Probes Master is likely not as robust as that contained in the Roche LightCycler™ 480 Probes Master.

Nonspecific products were noted in the verification of amplification for this gel. Nonspecific products reduce the PCR efficiency and must be eliminated as much as possible. Since a DPCR protocol contains inherent PCR inhibitors, further inhibition and reduction in efficiency is unacceptable in a MB-DPCR assay. The longer nonspecific fragments can be reduced by simply reducing the annealing/extension times as smaller PCR fragments will be preferentially amplified. Nonspecific fragments smaller than the intended amplicon will be increased, however, and as such an appropriate temperature and time combination

must be used. Additionally, no dimer formation was observed in the REV158 primer set due to the absence of nonspecific fragments.

5.4.4: Analysis of primers in real-time using melt analysis

Referencing Figure 4.21: Melt analysis of qPCR comparison of REV100 and REV158 primers after 48°C annealing, there are many peaks in the real-time melt assay of both primer sets at 48°C. These peaks are each indicative of a different fragment size produced. Since each fragment size will melt at a different temperature based on length and composition, the nonspecificity of both primers can be confirmed at this temperature. At 58°C, however, only one PCR fragment is observed in each reaction (Figure 4.22: Melt analysis of qPCR comparison of REV100 and REV158 primers after 58°C annealing). It is also clear that more product was made using the REV158 primer when compared to the REV100 primer. Since the first derivative plot shows a higher maximum for REV158, the resulting product is in higher abundance than for the reaction utilizing REV100. This is expected, as REV100 is known to form dimers, while REV158 is not. Dimer formation of the primer will result in lower overall PCR efficiency, thus producing less of the desired amplicon given the same conditions. As such, REV158 was selected as the primer for the remainder of the study. The use of a minor-groove adsorbing dye such as SYBR Green I is necessary for this assay, as its first derivative plot will change based on fragment size, whereas that of a MB assay will not.

5.4.5: Verification of direct amplification using REV158

Once selected, the primer had to be verified as working with the samples intended. This was accomplished through standard PCR utilizing a 90 second cycling profile followed by gel electrophoresis of the resulting amplicon in a 2% agarose, 0.05% ethidium bromide gel shown as Figure 4.23: Verification of amplification of direct samples using REV158 primer. This verification was done after the change in kit to BioLine's SensiFAST Probes Master. As such, the lack of amplification in the blood lane is due to this kit's inability to amplify a blood sample, not the primer itself. All other samples were amplified successfully and at the desired fragment size.

5.4.6: Assay optimization

Traditionally, PCR contains 3 steps within the cycle, each taking 30 seconds upon reaching the specified hold temperature. While optimization of PCR kits and polymerases for amplification only or with SYBR Green I has allowed the use of a 2 step method, this is not the case with MB probes. A thorough literature review did not produce any articles in which a 2 step method was used for MB assays. By using a 2 step method for MB PCR assays, time from sample preparation to results can be significantly reduced, further satisfying the goals of the project.

In order to develop a 2 step MB PCR assay, the extension time of the cycle was reduced incrementally over the course of multiple experiments. Initially, an experiment using all three steps for 30 seconds each was performed as a baseline from which to judge performance of subsequent assays. Extension times were then reduced by 5-10 seconds each experiment until the extension step was eliminated completely. For comparison, Figure 4.24 shows the initial baseline and Figure 4.27

shows the results of an experiment without an extension step. These experiments were performed using only positive and negative controls to minimize matrix effects induced by direct samples. As can be seen in Figure 4.27, purified genomic DNA at an initial concentration of 20 ng in the reaction vessel was successfully amplified with a corresponding signal. With positive results, it was determined that, under purified conditions, it is possible to use MB PCR assays in only 2 steps, thus eliminating a substantial amount of time.

Optimization continued with the denaturation step. Since this step is independent of polymerization and hybridization, the time necessary to denature the strands is determined solely by the kinetics of breaking DNA hydrogen bonds at 95°C. This is expected to be an extremely fast step, and many commercial standard and qPCR kits cite 5 seconds to be sufficient for the separation of DNA strands. Assays were designed with 5 seconds as the goal, but done in a sequential manner to check for performance drop-off. Assays with a 30, 15, 10, and 5 second denaturation time were performed without an extension step, and using a 30 second annealing step. These are shown as Figure 4.27, Figure 4.28, Figure 4.29, and Figure 4.30, respectively. There is no significant reduction in assay performance using a 5 second denaturation without extension when compared to the baseline in Figure 4.24. C_T values are not increased, and the spectra all show exponential growth in positive controls and no growth in negative controls.

Without an extension step, polymerization of nascent DNA must occur during the annealing step. It was therefore unlikely that a reduction in the annealing step would be possible, especially with using 48°C as the annealing temperature. Since the enzyme's optimal polymerization temperature is 72°C, this reduction in temperature would correspond to a reduction in enzyme efficiency. An assay was tested using a 5 second denaturation step, 15 second annealing step, and no

extension step as shown in Figure 4.31; this assay performed just as well as the other assays in spite of the expected reduction in enzyme efficiency when comparing the spectra and C_T values.

A verification run was performed using FOR100/REV158 primer set and a series dilution from 20 ng to 0.25 ng of control gDNA per reaction. The results in Figure 4.34 and Figure 4.33 show successful amplification under control, genomic conditions with sensitivity to 0.25 ng per reaction. The assay was then tested using 20 ng of control, genomic DNA in the previously optimized assay. The results of this 2 step protocol are shown as Figure 4.33 and Figure 4.34. A 2-step sensitivity test was then performed from 10 ng to 0.1 ng of control gDNA per reaction. No assay was performed using 20 ng in the sensitivity testing, as 20 ng had already been verified as working in the initial 2 step testing for this kit. The results for this assay are shown as Figure 4.35 and Figure 4.36.

Testing for amplification of direct samples was done in two phases: verification conditions and optimized conditions. Figure 4.41 and Figure 4.42 show the testing of a known homozygous T sample on the two probes under verification conditions. In Figure 4.42, representing the C allele of the SNP, amplification of signal is seen only in the positive control sample, which is a known heterozygote. By contrast, Figure 4.41, representing the T allele of the SNP, shows amplification in all samples except the NTC as expected.

Optimization of the MB concentration was the last step in optimization. As is the case with most nucleic acid sequence analysis, reagents are inherently expensive. The purpose of this project, as laid out in Chapter 1: , is to develop a method which is time- and cost-efficient for prescreening DNA samples. Thus, the lowest reagent concentration capable of eliciting a reliable, selective, and sensitive response is preferable.

5.4.7: Buffer development

Over the course of the project, false positive results occurred. This necessitated the development of an alternate sample preparation buffer. Since this buffer is added to the reaction, it is important to ensure it is compatible and will not cause undue changes in the analytical matrix. As qPCR has an inherently complex matrix, the sample preparation matrix should contain only components which are already present in the PCR matrix in order to minimize any additional matrix effects. This complex matrix affords a number of components from which to select a component already present, thus further reducing the probability of matrix interferences.

Water was initially used to incubate samples and as a solvent for DNA; water's lack of a buffering capacity was thought to be the cause of the false positives. Changes in pH can cause anomalies in the analytical matrix, leading to false positives. Most qPCR buffers contain TE as the bulk of the solution. The addition of water to the qPCR master mix, effectively diluting it, reduced the buffering capacity to the point that pH was affected. No pH measurement instruments on this scale were available, and the collection solution was changed to TE. The TE buffer, while slightly more reliable, still produced false positives. The addition of EDTA was considered, as it sequesters ions in solution, and reduced effective ion concentration would cause issues with hybridization of MBs.

The final approach was to use a 1X TE buffer, diluted with a Tris-HCl buffer at the same pH by a factor of 10. This allowed for the same buffering capacity since the pH and concentration of Tris were constant, while still diluting the EDTA. The presence of EDTA is advantageous, as cells contain enzymes which can digest DNA; these enzymes require metal cations as cofactors, however, and EDTA is able to sequester these cofactors and deactivate the enzymes capable of DNA

degradation. It is expected that the concentration of these ions in the cell, and the concentration of cells in solution, is low enough such that 1 mM EDTA in the incubation buffer is sufficient to sequester these ions without causing matrix effects in the qPCR.

As can be seen in Figure 4.38, false positive results are present using a standard 1X TE buffer. However, Figure 4.40 does not show a false positive result. A small exponential growth phase is visible, but does not rise appreciably above the background and is not indicative of a positive result for this allele. Figure 4.38 shows the results using 1X TE, while shows results using 0.1X TE. These data are produced from an experiment using a known, purified heterozygous control and a T homozygous sample, with an NTC using NFW instead of template. As such, the positive control should show signal response with respect to both probes, while the samples should show signal response with respect to the 21tSNP69 probe only, and none with respect to the 21cSNP69 probe. This modified TE buffer was used for sample incubation throughout the project after development.

5.5: Phase IV: Allelic discrimination using MB-PCR assays

5.5.1: Discrimination based on a 3-step format

After all verification was completed and the target assay developed, test samples were analyzed in conjunction with the direct PCR format. First, a known homozygous T/T sample was examined in order to screen for any issues with MB-DPCR reactions. This was done through standard 90 second cycling in order to ensure the primers and beacons had sufficient time to anneal and the polymerase sufficient time to synthesize nascent strands. As shown in Figure 4.41 and Figure 4.42, the T/T sample was successfully amplified and correctly genotyped based on

the exponential growth from the T probe's spectrum, and linear response in the C probe's spectrum. The three major sample types—hair, saliva, and buccal swabs—were tested, with each being correctly genotyped. With successful identification of a homozygous sample in a direct format under standard conditions, it was hypothesized that the assay could generate reliable profiles for all three genotypes in an accelerated, 2-step format.

5.5.2: Discrimination based on a 2-step format

The prior success led to the testing of the same homozygous T/T sample under the optimized conditions. Again, only a homozygous T/T sample was tested initially. Figure 4.43 and Figure 4.44 show successful amplification and identification of this homozygous T/T sample using a 5 second denaturation phase, 30 second annealing phase profile for 35 cycles. All three sample types were tested and correctly genotyped. This demonstrates the assay's ability to discriminate genotypes under the accelerated conditions.

The assay was then tested for the discrimination of all genotypes. Figure 4.45 through Figure 4.50 shows all samples tested with the accelerated conditions. Beginning with Figure 4.45 and Figure 4.46, the homozygous T/T samples were tested alongside all others. The successful amplification of this genotype from the previous experiment was confirmed. All samples from this genotype were correctly identified with strong signal generation for the T probe with none for the C probe, and thus unremarkable.

The heterozygous samples, Figure 4.47 and Figure 4.48, were also correctly identified. The heterozygous saliva and hair samples showed strong signal generation in both C and T probes. The buccal swab samples did not show strong signal generation; they were still distinguishable from the NTC, but would require

replicate testing in order to conclusively define the genotype. This is likely due to a low quantity of DNA recovered from the swab. If the participant does not swab well, a low cell count will be obtained resulting in low DNA quantity. If possible, the sample should be screened with the Nanodrop to ensure sufficient quantity is added to the reaction, above the LOD. Ideally, multiple swabs from the same individual could be collected.

The analysis of the homozygous C/C samples, Figure 4.49 and Figure 4.50, was less successful. The buccal swab sample performed as expected; strong signal was generated from the C probe, while no distinguishable signal was generated from the T probe. The saliva and hair samples, though, showed strong signal as a heterozygote. The most likely cause is the extremely low annealing temperature of the assay. The same individual was asked to donate additional samples for study, with the same results. Contamination is unlikely, as the NTC is performing as expected, and replicate results occurred on separate days. A new MB was designed with a higher dissociation temperature in response to this.

5.6: Development of new MB

5.6.1: Software predictions and theoretical design

Since the NTC samples performed without amplification, it was expected that the false positives—even after the development of a new buffer and use of a new primer—could be attributed to the extremely low temperature of the annealing phase in the assay. Most PCR applications use an annealing temperature of around 60°C; an annealing temperature so low as 48°C will have nonspecific annealing of primers especially, and possibly even the MBs themselves. To correct this issue, a new MB was designed for continued study. The newly designed beacon was

extended to 34 bp from 21 bp, allowing for stable hybrid formation at a significantly increased temperature of 58.0°C. The stem was designed with a dissociation temperature of 62.8°C. The stem must have a higher dissociation temperature than the probe, as it must compete with mismatch hybrid formation. The nBLAST and BioEdit alignment data suggest high specificity for the intended target. The secondary structure data suggests efficient thermodynamic competition. With this in mind, the probe and its complementary and mismatch oligonucleotides were ordered. The probe annealing temperature prediction is shown in Figure 4.51, with the full analysis of the MB shown as Figure 4.52. The BioEdit alignment data with both primers is shown as Figure 4.53. The predicted temperatures fall within the ranges necessary for the existing assay. The stem melting temperature is higher than the probe melting temperature, and the probe T_M is at 58.0°C, which is much more compatible with PCR.

5.6.2: Thermal profile and initial testing of new MB

The first step of validating a new MB is to generate its thermal profile. The *in silico* data gives a great insight as to the theoretical predictions, but those predictions are based on different generalized algorithms and averages; each oligonucleotide will differ slightly from these predictions under various conditions such as salt, BSA, and dNTP concentrations. As such, the thermal profile is necessary for each MB in the presence of each genotype is necessary. This process took place as before, during Phase 1. The assays consisted of 1.000 µM MB, 10.00 µM complementary (or mismatch) strand, in 8.00 mM Tris-HCl, 280 nM $MgCl_2$, pH 8.0. The assay was held at 95°C for two minutes to dissociate any dimers and secondary structures, cooled to 55°C for two minutes to anneal the MB to the oligonucleotides, and cooled again to 25°C. Once cooled to 25°C, the assay was

held for 30 seconds before signal acquisition, then heated 1°C and repeated until a final temperature of 90°C.

The resulting fluorescence was normalized to unity and plotted as a function of temperature, shown as Figure 4.54. Both the new and original MBs were tested for comparison. The new MB exhibited markedly increased signal generation, with hybrid formation stable and specific between 60-68°C. Specificity was determined by the difference in match vs. mismatch fluorescence at the given temperatures; since the mismatch showed almost no fluorescence, while the match showed comparatively high fluorescence, the probe is specific over this temperature range. This matches well with the thermal profile of the primer set FOR100/REV158, as shown in Figure 4.19. The thermal profile of the primer set obtained via annealing temperature gradient suggests that the ideal temperature for the assay now lies above 60°C. The primer set does not show nonspecific annealing at or above 61.7°C based on the gel shown in Figure 4.19, and the probe hybrid is stable and highly specific at this temperature. This suggests greatly increased PCR efficiency and identification when compared to the current assay.

The thermal profile only tests a synthetic oligonucleotide, and not genomic DNA. The next step in validation of the new MB was to test the MB with a known homozygous C/C and homozygous T/T sample for performance under purified genomic conditions. This was done using a standard 90 second cycling profile at 60.0°C. Initial testing was performed at a lower annealing temperature than target in order to screen for nonspecific annealing to genomic DNA; nontarget hybrid formation is more preferable at a lower temperature due to thermodynamic considerations of base-base mismatching as discussed in 2.7: Thermodynamic basis for allelic discrimination. The resulting data is shown in Figure 4.55: Initial verification of new MB, standard conditions. This data shows clear amplification

and identification under genomic, purified conditions for the intended homozygous C/C target, with no exponential signal generated for the homozygous T/T sample or NTC. This is the expected outcome, as this probe was designed for specificity to the C allele. Future assays should test this MB under direct conditions in the same format for performance; if performance of this MB is maintained under direct, 2-step conditions, the accompanying MB targeted for the T allele should be ordered and investigated as a pair for direct genotyping in real-time, continuing this project.

5.7: Limitations of methods

It is important to note that the use of MBs in a direct PCR format for allelic discrimination does have limitations. It should not be considered a replacement to current industry-standard practices. This technique, while capable of successfully genotyping individuals, does not have the level of reliability of STR profiling. Further, it has not been tested in a true multiplex format. Only one locus was examined in this study. As such, no conclusions can be made about the success of genotyping an individual sample at more than one locus simultaneously. If this technique cannot be adapted to a multiplex, it may become prohibitively expensive for a preliminary screening.

It is also important to note that the enzymes in the kits used in this project were not designed with DPCR as an end-use. While still successful in amplifying the fragments, a different enzyme may perform at a better or worse rate, thus affecting the success rate of genotyping. The reaction mixture should be optimized similar to that of a commercial DPCR kit, as the analytical matrix will be affected by the presence of PCR inhibitors, many of which have not been studied for their effects on MB-PCR assays. As not much is known about the effects of these

inhibitors on the MB conformational changes, care should be taken when interpreting results.

Finally, reliability studies were not performed in this project. No data was collected about the intra-assay, intra-day, intra-investigator, or investigator-to-investigator variability. Therefore, no empirical reliability can be attributed to this assay. This was not within the scope of the project, but is important to note.

5.8: Future directions

The project, while considered complete, does have areas in which future study should be conducted. Primarily, the next phase of the study shows promise with these results. This next phase is the incorporation of InDel-targeted MB-DPCR assays. As SNPs have been successfully genotyped, InDels should prove a much more reliable target with a higher success rate. This proof-in-concept supports further study.

Another important future direction is the study of the assay reliability. This assay, while quite frequently successful, was not studied with empirical variability in mind. Rather, this was designed as a more qualitative investigation into the viability of this technique. Nothing can currently be said about its success rate, false positives, or false negatives; also, as only one investigator performed all assays, there should be a study into the success rate between investigators.

A critical follow-up is the development of a kit which is specific to MB-DPCR assays. While many polymerases have been designed for DPCR assays, and are commercially available, most contain endo- or exo-nuclease activity as a mechanism for correcting erroneous polymerization due to PCR inhibitors and a more complex analytical matrix. These enzymes, however, are incompatible with MB chemistry, as any endo- or exo-nuclease activity will cleave the quencher,

reporter, or both from the oligonucleotide, resulting in null data. Standard PCR or qPCR kits are not designed with these inhibitors in mind, and thus do not contain agents which can help alleviate this inhibition to PCR efficiency or MB conformational changes. Thus, a new type of kit should be developed in order to correct for this deficiency. Adjusting the BSA, dNTP, and salt concentrations should be the first step in the development of a new kit. Finally, the development of a *Pyrococcus furiosus*-like (*Pfu*) polymerase without an exonuclease domain could potentially allow functionality with more difficult samples, such as blood, as *Pfu*-based enzymes are currently in use for direct PCR kits.

CHAPTER 6: CONCLUSIONS

CHAPTER 6: CONCLUSIONS

6.1: Introduction

The specific objectives of this project, per 1.5: Aim and scope of the project, were to verify primer functionality, develop a thermal profile of the MBs, verify a working MB assay, screen for genotypes needed for the project, test various PCR master mixes for compatibility, develop a working MB-DPCR assay, and optimize the MB-DPCR assay. Phase I of the project was centered around verifying the primer functionality and developing the MB profiles and assay. Phase II focused on selecting an appropriate PCR kit. Phase III was designed to determine the suitability of sample types and develop the protocol, and developed a new primer. Phase IV confirmed the viability of the assay and established the basis for future work, and developed a new beacon for further testing.

6.2: Conclusion of Phase I

Phase I was used to validate components and techniques in current literature. Thermal lysis is a widely accepted technique for destroying cell membranes and liberating genomic DNA from the nucleolus of a eukaryotic cell. Most techniques for DNA extraction utilize enzymes such as Proteinase K; however, since one of the major objectives is to use a direct sample, Proteinase K is not an option. The use of Proteinase K without subsequent extraction would destroy the polymerase after only a few cycles, resulting in no amplification. Electroporation, another commonly used technique in biochemistry and molecular biology for disrupting a cell membrane, would not be as efficient or user-friendly.

Thermal lysis was chosen as it is capable of reliably destroying cell membranes and liberating DNA.

This phase of the project was also necessary to validate the reaction components. Care was taken to ensure methodical testing of the primer sets, thermal profiles of the MBs, MBs for identification in an assay separately, MBs as a pair, and to screen for potential candidates. This phase produced the foundation for the subsequent work; without this phase, nothing would have been known to work. Phase I also provided the purified DNA needed for controls in Phase III and Phase IV.

6.3: Conclusion of Phase II

Phase II rigorously tested multiple kits for use in the project. While the original kit used in Phase I was known to work with all components, it was important to screen for the use of potential other kits as the Roche LightCycler 480 has not been tested in a direct PCR assay. The Phusion kits from ThermoScientific were an extremely attractive option as they are designed for direct PCR with high fidelity. The high fidelity of the kit is desirable, as mutations early in the cycling can result in incorrect or failed genotyping. Additionally, they are designed with direct PCR as an end-use, and minimal modification should be necessary. Unfortunately, these kits are incompatible with MBs for the same reason they are attractive: the fidelity of the polymerase is owed to its 3'-5' exonuclease domain, which destroys the MB and prevents genotyping.

The Roche LightCycler 480 was then compared to the BioLine SensiFAST Probes Master as a potential replacement. The BioLine kit was shown to have improved sensitivity over the Roche kit, and selected in its place. Upon screening for suitability of samples, however, the BioLine kit was unable to amplify a direct

sample of blood in any tested concentrations. The improved sensitivity of the BioLine kit was too important to trade for the amplification of this sample type, and this sample type was excluded from further research in this project. This is a proof-in-concept project, and blood proved to be too difficult a sample to amplify. Blood was determined to be an ill-suited sample for initial study, and the remainder of the project focused on hair, saliva, and buccal swab samples.

6.4: Conclusion of Phase 3

Phase III focused on novel method development. This phase took the majority of the project. Amplification of direct samples was first verified using both the Roche and BioLine kits. Observation of nonspecific priming led to the development of a new primer, REV158, and its validation. The new primer exhibited improved performance in comparison to REV100, and less nonspecific priming. Some nonspecific priming was still observed, given the thermal profile of the MBs and the assay design. Direct amplification of the intended three sample types studied using REV158 was successful, determined through standard PCR and gel electrophoresis.

Assay optimization was crucial. The determination of an optimal annealing temperature for the assay was determined through annealing temperature gradient of the primers and thermal profile of the MBs. The assay required the use of an extremely low 48.0°C annealing temperature due to the thermal profile of the MBs, leading to nonspecific amplification. The thermocycle of the assay was optimized stepwise and methodically through the determination of cycle threshold values. Cycling as fast as a 1 second denaturation and 15 second annealing phase for 35 cycles was achieved under purified genomic conditions, with the complete elimination of the extension step. This elimination of the extension step improves

the time to results significantly, and is not currently documented in the literature for MB assays. This is a considerable breakthrough for MB use.

False positives in direct assays were observed, and a new buffer was developed for sample preparation. The sample preparation buffer is a modified version of a standard TE buffer. The reduced EDTA concentration improved genotyping fidelity, but did not completely eliminate issues with false positives. This new buffer did, however, make a significant impact on the fidelity of genotyping and was used in Phase IV.

6.5: Conclusion of Phase IV

Phase IV lays the foundation for future work in this field. This phase focused on allelic discrimination using the novel method. Purified genomic conditions allowed for cycling as fast as 1 second denaturation, 15 seconds annealing. Given the inherent inhibition of direct samples, however, additional time was allotted to each step with the exception of extension. Verification under standard cycling was performed before accelerated cycling. Once verified, the accelerated assay was tested and also succeeded in genotyping the majority of samples. Some samples were not correctly identified even after replicate assays; the likely cause is the extremely low annealing temperature of the assay required by the current MBs. A new MB has been developed and is in testing at the time of publication. Initial data suggests improved sensitivity and specificity, but this is not conclusive and must be tested as stringently as the rest of the project.

6.6: General project conclusions

As a proof-in-concept, this project has succeeded in showing the applicability of MB assays to DPCR of forensic samples. While some samples may need to be tested more than once, the pillar of any analytical technique is replicate measurements. Even in single-hair samples, enough sample is prepared to perform at least triplicate measurements.

While the use of SNP-targeted MB assays in DPCR does have inherent intra-assay variability, this lack of variability is low enough to justify further study. As SNPs are the most difficult polymorphism to analyze—especially with hybridization probes—this was an extreme application to predict the success rate of a much simpler polymorphism. As the data clearly show, MB-DPCR 2-step assays are capable of allelic discrimination of SNPs. The speed of the assay developed is a marked advantage over current profiling methods. This has the potential to greatly speed identification of individuals in a forensic application.

This technique is not limited to forensic science, however. Biomedical applications could very well benefit from this also. The screening of known polymorphisms follows a very similar workflow to current STR profiling, with the same time and resource constraints, and opportunities for contamination. Point-of-care assays could be developed with this same workflow in mind. The reliability of this technique should not be judged alongside industry-standard methods. Current standard practices require the purification of DNA in order to limit matrix interference and ensure reliability in results. This is necessary, as these results are commonly used for legal proceedings. Any ambiguity would negate its value. This technique was not designed as a replacement for current industry-standard practices. Rather, this is intended as a supplement.

Given the current nature of the DNA backlog and the fact that labs are receiving an ever-increasing amount of samples, a preliminary test which can be performed much faster than standard methods is desirable. This preliminary test could help to prioritize the most likely samples for successful identification, potentially even giving a probability of a match. While this probability should not be used as a definitive identification, it could help to further prioritize sample analysis. Additionally, this procedure can eliminate nonviable samples in a much faster timeframe.

WORKS CITED

ABOUL-ELA, F., KOH, D., TINOCO, I. & MARTIN, F. H. Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A,C,G,T).

ABRAVAYA, K., HUFF, J., MARSHALL, R., MERCHANT, B., MULLEN, C., SCHNEIDER, G. & ROBINSON, J. 2003. Molecular beacons as diagnostic tools: Technology and applications. *Clinical Chemistry and Laboratory Medicine,* 41**,** 468-474.

AKANE, A., MATSUBARA, K., NAKAMURA, H. & TAKAHASHI, S. 1994. Identification of the Heme Compound Copurified with Deoxyribonucleic Acid (DNA) from Bloodstains, a Major Inhibitor of Polymerase Chain Reaction (PCR) Amplification. *JOURNAL OF FORENSIC SCIENCES,* 39**,** 362.

ALBERTS, B., JOHNSON, A., LEWIS, J., MORGAN, D., RAFF, M., ROBERTS, K. & WALTER, P. 2015. *Molecular Biology of the Cell,* New York, NY, Garland Science.

AMORIM, A. & PEREIRA, L. 2005. Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs. *Forensic Science International,* 150**,** 17-21.

ANTONY, T. & SUBRAMANIAM, V. 2002. Comparison of nucleic acid hybridization using molecular beacons or FRET-oligos: Sequence length and salt effects. *Biophysical Journal,* 82**,** 429a-429a.

ARBELOA, I. L. P. 1981a. Dimeric and trimeric states of the fluorescein dianion. Part 1.—Molecular structures. *Journal of the Chemical Society, Faraday Transactions 2,* 77**,** 1725-1733.

ARBELOA, I. L. P. 1981b. Dimeric and trimeric states of the fluorescein dianion. Part 2.—Effects on fluorescence characteristics. *Journal of the Chemical Society, Faraday Transactions 2,* 77**,** 1735-1742.

BERG, J. 2012. *Biochemistry,* New York :, WH Freeman.

BERLEPSCH, H. V. & BÖTTCHER, C. 2012. The Morphologies of Molecular Cyanine Dye Aggregates as Revealed by Cryogenic Transmission Electron Microscopy. *In:* KOBAYASHI, T. (ed.) *J-aggregates.* Singapore: World Scientific Publishing Company.

BERNACCHI, S. & MÉLY, Y. 2001. Exciton interaction in molecular beacons: a sensitive sensor for short range modifications of the nucleic acid structure. *Nucleic acids research,* 29**,** E62-2.

BIOLINE REAGENTS 2015. MyTaq Blood PCR Kit. London, UK: BioLine Reagents,,.

BIOSEARCH TECHNOLOGIES 2016. Fluorophore & BHQ Dye Selection Chart.

BONNET, G., TYAGI, S., LIBCHABER, A. & KRAMER, F. R. 1999. Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 6171-6.

BOONEN, T., VITS, K., HOSTE, B., HUBRECHT, F. & COPENHAGEN DENMARK 20070821-20070825 COPENHAGEN, D. 2008. The visualization and quantification of cell nuclei in telogen hair roots by fluorescence microscopy, as a pre-DNA analysis assessment. *Forensic Science International: Genetics Supplement Series,* 1**,** 16-18.

BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. & WITTWER, C. T. 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry,* 55**,** 611-622.

BUTLER, J. M. 2005. *Forensic DNA typing : biology, technology, and genetics of STR markers,* Burlington, MA, Elsevier Academic Press.

CARRACEDO, A. 2005. *Forensic DNA typing protocols,* Totowa, N.J., Humana Press.

CASPI, R., ALTMAN, T., BILLINGTON, R., DREHER, K., FOERSTER, H., FULCHER, C. A., HOLLAND, T. A., KESELER, I. M., KOTHARI, A., KUBO, A., KRUMMENACKER, M., LATENDRESSE, M., MUELLER, L. A., ONG, Q., PALEY, S., SUBHRAVETI, P., WEAVER, D. S., WEERASINGHE, D., ZHANG, P. & KARP, P. D. 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research,* 42**,** D459-D471.

CELL PROJECTS, L., ,. 2016. *Isohelix DNA Buccal Swabs* [Online]. Available: http://www.isohelix.com/products/isohelix-dna-buccal-swabs/ [Accessed 7 January 2016].

CHAKRABORTY, R. 1992. Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Human biology,* 64**,** 141-59.

CHAKRABORTY, R., STIVERS, D. N., SU, B., ZHONG, Y. & BUDOWLE, B. 1999. The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. *ELECTROPHORESIS,* 20**,** 1682-1696.

CHAMBERS, G. K. & MACAVOY, E. S. 2000. Microsatellites: consensus and controversy. *Comparative biochemistry and physiology. B: Comparative biochemistry,* 126**,** 455 - 476.

CHIEN, A., EDGAR, D. B. & TRELA, J. M. 1976. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *Journal of Bacteriology,* 127**,** 1550-1557.

CLEGG, R. M. 1996. Fluorescence Resonance Energy Transfer. *In:* WANG, X. F. & HERMAN, B. (eds.) *Fluorescence Imaging Spectroscopy and Microscopy.* New York, NY: John Wiley & Sons, Inc.

CLEGG, R. M. 2009. Chapter 1 Förster resonance energy transfer—FRET what is it, why do it, and how it's done. *Laboratory Techniques in Biochemistry and Molecular Biology.* Elsevier Science & Technology.

COMEY, C. T., KOONS, B. W., PRESLEY, K. W. & SMERICK, J. B. 1994. DNA Extraction Strategies for Amplified Fragment Length Polymorphism Analysis. *JOURNAL OF FORENSIC SCIENCES,* 39**,** 1254.

COSTA, G., DARIO, P., LUCAS, I., RIBEIRO, T., ESPINHEIRA, R. & GEADA, H. Autosomal SNPs in paternity investigation. *Forensic Science International: Genetics Supplement Series,* 1**,** 507-509.

DEMERS, D. B., CURRY, E. T., EGHOLM, M. & SOZER, A. C. 1995. Enhanced PCR amplification of VNTR locus D1S80 using peptide nucleic acid (PNA).

DIDENKO, V. V. 2006. *Fluorescent energy transfer nucleic acid probes : designs and protocols,* Totowa, N.J. :, Humana Press.

DIXON, L. A., MURRAY, C. M., ARCHER, E. J., DOBBINS, A. E., KOUMI, P. & GILL, P. 2005. Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. *Forensic Science International,* 154**,** 62-77.

ECKHART, L., BACH, J., BAN, J. & TSCHACHLER, E. 2000. Melanin Binds Reversibly to Thermostable DNA Polymerase and Inhibits Its Activity. *Biochemical and Biophysical Research Communications,* 271**,** 726-730.

EL-HAJJ, H. H., MARRAS, S. A., TYAGI, S., SHASHKINA, E., KAMBOJ, M., KIEHN, T. E., GLICKMAN, M. S., KRAMER, F. R. & ALLAND, D. 2009. Use of sloppy molecular beacon probes for identification of mycobacterial species. *Journal of clinical microbiology,* 47**,** 1190-8.

EMERSON, E. S., CONLIN, M. A., ROSENOFF, A. E., NORLAND, K. S., RODRIGUEZ, H., CHIN, D. & BIRD, G. R. 1967. The geometrical structure and absorption spectrum of a cyanine dye aggregate. *The Journal of Physical Chemistry,* 71**,** 2396-2403.

FEDERAL BUREAU OF INVESTIGATION. 2015. *CODIS and NDIS Fact Sheet* [Online]. Washington, DC: Federal Bureau of Investigation,,. Available: https://www.fbi.gov/about-us/lab/biometric-

analysis/codis/codis-and-ndis-fact-sheet [Accessed 08 January 2016].

FREDERICKS, J. 2011. *Development of New Tools for Forensic Analysis of DNA from Compromised Bone.* Ph. D., Cranfield University.

FREDERICKS, J. 2014. (unpublished work).

FREDERICKS, J. D., BENNETT, P., WILLIAMS, A. & ROGERS, K. D. 2012. FTIR spectroscopy: A new diagnostic tool to aid DNA analysis from heated bone. *Forensic Science International: Genetics,* 6**,** 375-380.

FREDERICKS, J. D., RINGROSE, T. J., DICKEN, A., WILLIAMS, A. & BENNETT, P. 2015. A potential new diagnostic tool to aid DNA analysis from heat compromised bone using colorimetry: A preliminary study. *Science & Justice,* 55**,** 124-130.

FRENCH, D., MCDOWELL, D., THOMSON, J., BROWN, T. & DEBENHAM, P. HyBeacons®: A novel DNA probe chemistry for rapid genetic analysis. International Congress Series, 2006. Elsevier, 707-709.

GILL, P. 2001. An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *International Journal of Legal Medicine,* 114**,** 204-210.

GILL, P., WERRETT, D. J., BUDOWIE, B. & GUERRIERI, R. 2004. An assessment of whether SNPs will replace STRs in national DNA databases. *SCIENCE AND JUSTICE,* 44**,** 51-54.

GOEL, G., KUMAR, A., PUNIYA, A. K., CHEN, W. & SINGH, K. 2005. Molecular beacon: a multitask probe. *Journal of Applied Microbiology,* 99**,** 435-442.

GUNDRY, C. N., VANDERSTEEN, J. G., REED, G. H., PRYOR, R. J., CHEN, J. & WITTWER, C. T. 2003. Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clinical chemistry,* 49**,** 396-406.

HUIJSMANS, R., DAMEN, J., VAN DER LINDEN, H. & HERMANS, M. 2007. Single Nucleotide Polymorphism Profiling Assay to Confirm

the Identity of Human Tissues. American Society for Investigative Pathology.

HVIID, T. V., MADSEN, H. O. & MORLING, N. 1992. HLA-DPB1 typing with polymerase chain reaction and restriction fragment length polymorphism technique in Danes. Wiley-Blackwell Munksgaard.

INTERNATIONAL HUMAN GENOME SEQUENCING, C. 2001. Initial sequencing and analysis of the human genome. *NATURE - LONDON-*, 860-921.

JEFFREYS, A. J., WILSON, V. & THEIN, S. L. 1985. Hypervariable 'minisatellite' regions in human DNA. *Nature*.

JELLEY, E. E. 1936. Spectral Absorption and Fluorescence of Dyes in the Molecular State. *Nature,* 138**,** 1009-1010.

JELLEY, E. E. 1937. Molecular, Nematic and Crystal States of I : I′-Diethyl-Ψ-Cyanine Chloride. *Nature,* 139**,** 631-631.

JOHANSSON, M. K. 2006. Choosing Reporter-Quencher Pairs for Efficient Quenching Through Formation of Intramolecular Dimers. *In:* DIDENKO, V. V. (ed.) *Fluorescent Energy Transfer Nucleic Acid Probes.* Totowa, NJ: Humana Press.

JOHANSSON, M. K., FIDDER, H., DICK, D. & COOK, R. M. 2002. Intramolecular Dimers: A New Strategy to Fluorescence Quenching in Dual-Labeled Oligonucleotide Probes. *Journal of the American Chemical Society,* 124**,** 6950-6956.

KASHA, M. 1963. Energy Transfer Mechanisms and the Molecular Exciton Model for Molecular Aggregates. *Radiation Research,* 20**,** 55-70.

KASHA, M. 1991. Energy transfer, charge transfer, and proton transfer in molecular composite systems. *Basic life sciences,* 58**,** 231-51; discussion 251-5.

KHAIRUTDINOV, R. F. & SERPONE, N. 1997. Photophysics of Cyanine Dyes: Subnanosecond Relaxation Dynamics in Monomers, Dimers, and H- and J-Aggregates in Solution. *The Journal of Physical Chemistry B,* 101**,** 2602-2610.

KIDD, K. K., PAKSTIS, A. J., SPEED, W. C., GRIGORENKO, E. L., KAJUNA, S. L. B., KAROMA, N. J., KUNGULILO, S., KIM, J.-J., LU, R.-B., ODUNSI, A., OKONOFUA, F., PARNAS, J., SCHULZ, L. O., ZHUKOVA, O. V. & KIDD, J. R. 2006. Developing a SNP panel for forensic identification of individuals. *Forensic Science International,* 164**,** 20-32.

KOSTRIKIS, L. G., TYAGI, S., MHLANGA, M. M., HO, D. D. & KRAMER, F. R. 1998a. Molecular beacons - Spectral genotyping of human alleles. *Science,* 279**,** 1228-1229.

KOSTRIKIS, L. G., TYAGI, S., MHLANGA, M. M., HO, D. D. & KRAMER, F. R. 1998b. Spectral genotyping of human alleles. *Science,* 279.

LAKOWICZ, J. R. 2013. *Principles of fluorescence spectroscopy*, Springer Science & Business Media.

LEE, H. C. & LADD, C. 2001. Preservation and collection of biological evidence. *Croatian medical journal,* 42**,** 225-8.

LI, Q. 2002. A new class of homogeneous nucleic acid probes based on specific displacement hybridization. *Nucleic Acids Research,* 30**,** 5e-5.

LI, Y. S., ZHOU, X. Y. & YE, D. Y. 2008. Molecular beacons: An optimal multifunctional biological probe. *Biochemical and Biophysical Research Communications,* 373**,** 457-461.

LIGTENBERG, A. J. M. & VEERMAN, E. C. I. 2014. Saliva : secretion and functions. Basel :: Karger.

LIVAK, K. J., FLOOD, S. J., MARMARO, J., GIUSTI, W. & DEETZ, K. 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR methods and applications,* 4**,** 357-62.

MACKAY, I. M., ARDEN, K. E. & NITSCHE, A. 2002. Real-time PCR in virology. *Nucleic Acids Research,* 30**,** 1292-1305.

MANDELKERN, M., ELIAS, J. G., EDEN, D. & CROTHERS, D. M. 1981. The dimensions of DNA in solution. *Journal of Molecular Biology,* 152**,** 153-161.

MARRAS, S. A., KRAMER, F. R. & TYAGI, S. 2003. Genotyping SNPs with molecular beacons. *Methods in molecular biology (Clifton, N.J.),* 212**,** 111-28.

MARRAS, S. A. E., KRAMER, F. R. & TYAGI, S. 1999. Multiplex detection of single-nucleotide variations using molecular beacons. *Genetic Analysis-Biomolecular Engineering,* 14**,** 151-156.

MARRAS, S. A. E., KRAMER, F. R. & TYAGI, S. 2002. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Research,* 30**,** e122.

MHLANGA, M. M. & MALMBERG, L. 2001. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods,* 25**,** 463-471.

MORRISON, T. B., WEIS, J. J. & WITTWER, C. T. 1998. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *BioTechniques,* 24**,** 954-8, 960, 962.

MÜLLER, K., KLEIN, R., MILTNER, E. & WIEGAND, P. 2007. Improved STR typing of telogen hair root and hair shaft DNA. *ELECTROPHORESIS,* 28**,** 2835-2842.

MULLIS, K. B., ERLICH, H. A., ARNHEIM, N., HORN, G. T., SAIKI, R. K. & SCHARF, S. J. 1987. *Process for amplifying, detecting, and/or-cloning nucleic acid sequences*. United States patent application US 06/828,144. Jul 28, 1987.

NATIONAL LIBRARY OF MEDICINE. 2011a. *Microsatellite Repeats* [Online]. Available: https://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Microsatellite+Repeats&field=entry [Accessed 08 January 2016].

NATIONAL LIBRARY OF MEDICINE. 2011b. *Minisatellite Repeats* [Online]. Available: https://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Minisatellite [Accessed 08 January 2016].

NAZARENKO, I. A., BHATNAGAR, S. K. & HOHMAN, R. J. 1997. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Research,* 25**,** 2516-2521.

NELSON, M., CHASE, R. & DEPALMA, L. 2013. Making Sense of DNA Backlogs, 2012 - Myths vs. Reality. *In:* OFFICE OF JUSTICE PROGRAMS (ed.). Washington, D.C.

OFFICE OF JUSTICE PROGRAMS. 2012. *DNA Evidence: Basics of Analyzing* [Online]. Available: http://www.nij.gov/topics/forensics/evidence/dna/basics/pages/analyzing.aspx [Accessed 08 January 2016].

OHHARA, M., KUROSU, Y. & ESUMI, M. 1994. Direct Pcr of Whole-Blood and Hair Shafts by Microwave Treatment. *Biotechniques,* 17**,** 726-728.

ORTIZ, E., ESTRADA, G. & LIZARDI, P. M. 1998. PNA molecular beacons for rapid detection of PCR amplicons. *Molecular and Cellular Probes,* 12**,** 219-226.

OTTENS, R., TAYLOR, D., ABARNO, D. & LINACRE, A. 2013a. Successful direct amplification of nuclear markers from a single hair follicle. *Forensic Science Medicine and Pathology,* 9**,** 238-243.

OTTENS, R., TAYLOR, D., ABARNO, D., LINACRE, A. & PROGRESS IN FORENSIC GENETICS 15 PROGRESS IN FORENSIC, G. 2013b. Optimising direct PCR from anagen hair samples. *Forensic Science International: Genetics Supplement Series,* 4**,** e109-e110.

PARK, S. J., KIM, J. Y., YANG, Y. G. & LEE, S. H. 2008. Direct STR amplification from whole blood and blood- or saliva-spotted FTA (R) without DNA purification. *Journal of Forensic Sciences,* 53**,** 335-341.

PEREIRA, F., CARNEIRO, J., MATTHIESEN, R., VAN ASCH, B., PINTO, N., GUSMÃO, L. & AMORIM, A. 2010. Identification of species by multiplex analysis of variable-length sequences. *Nucleic Acids Research,* 38**,** e203-e203.

PROMEGA CORPORATION 2015. PowerPlex 16 System. Madison, WI: Promega Corporation,,.

PUBLIC HEALTH RESEARCH INSTITUTE. 2016. *Molecular Beacon Design* [Online]. Available: http://molecular-beacons.org/MB_SC_design.html [Accessed 09 March 2016 2016].

PUERS, C., HAMMOND, H. A., JIN, L., CASKEY, C. T. & SCHUMM, J. W. 1993. Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01[AATG]n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *American journal of human genetics,* 53**,** 953-8.

REYNOLDS, R., SENSABAUGH, G. & BLAKE, E. 1991. Analysis of genetic markers in forensic DNA samples using the polymerase chain reaction. *Analytical chemistry,* 63**,** 2-15.

ROBERTSON, K., MCNEVIN, D. & ROBERTSON, J. 2007. STR genotyping of exogenous hair shaft DNA. *Australian Journal of Forensic Sciences,* 39**,** 107-122.

ROOT, D. D., VACCARO, C., ZHANG, Z. L. & CASTRO, M. 2004. Detection of single nucleotide variations by a hybridization proximity assay based on molecular beacons and luminescence resonance energy transfer. *Biopolymers,* 75**,** 60-70.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNApolymerase. *Science*.

SCHNEIDER, P. M. 2007. Scientific standards for studies in forensic genetics. *Forensic Science International,* 165**,** 238-243.

SCHOLES, G. D. & GHIGGINO, K. P. 1994. Electronic Interactions and Interchromophore Excitation Transfer. *The Journal of Physical Chemistry,* 98**,** 4580-4590.

SCIENTIFIC WORKING GROUP ON MATERIALS ANALYSIS (SWGMAT) 2005. Forensic Human Hair Examination Guidelines. *In:* FEDERAL BUREAU OF INVESTIGATION (ed.). United States Government.

SINGER, M. F. 1982. *Highly repeated sequences in mammalian genomes*.

SOLINAS, A., BROWN, L. J., MCKEEN, C., MELLOR, J. M., NICOL, J., THELWELL, N. & BROWN, T. 2001. Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic acids research,* 29**,** E96.

SULLIVAN, K. M., MANNUCCI, A., KIMPTON, C. P. & GILL, P. 1993. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *BioTechniques,* 15**,** 636-8, 640-1.

TAUTZ, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic acids research,* 17**,** 6463-71.

TAUTZ, D. 1993. Notes on the definition and nomenclature of tandemly repetitive DNA sequences. *In:* PENA, S. D. J., CHAKRABORTY, R., EPPLEN, J. T. & JEFFREYS, A. (eds.) *DNA Fingerprinting: State of the Science.* Birkhäuser Basel.

THERMO-FISHER 2015a. Phire Animal Tissue Direct PCR Kit. Waltham, MA: Thermo-Fisher,,.

THERMO-FISHER 2015b. Phusion Blood Direct PCR KIt. Waltham, MA: Thermo-Fisher,,.

THERMO-FISHER 2015c. Phusion Human Specimen Direct PCR Kit. Waltham, MA: Thermo-Fisher,,.

TSOURKAS, A., BEHLKE, M. A., ROSE, S. D. & BAO, G. 2003. Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Research,* 31**,** 1319-1330.

TULLY, G., SULLIVAN, K. M. & GILL, P. 1993. Analysis of 6 VNTR loci by 'multiplex' PCR and automated fluorescent detection. *Human Genetics,* 92**,** 554-562.

TYAGI, S., BRATU, D. P. & KRAMER, F. R. 1998. Multicolor molecular beacons for allele discrimination. *Nature biotechnology,* 16**,** 49-53.

TYAGI, S. & KRAMER, F. R. 1996. Molecular beacons: Probes that fluoresce upon hybridization. *Nature Biotechnology,* 14**,** 303-308.

TYAGI, S. & KRAMER, F. R. 2012. Molecular Beacons in Diagnostics. *F1000 Medicine Reports,* 4**,** 10.

URQUHART, A., KIMPTON, C. & GILL, P. 1993. Sequence variability of the tetranucleotide repeat of the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus. *Human Genetics,* 92**,** 637-638.

URQUHART, A., KIMPTON, C. P., DOWNES, T. J. & GILL, P. 1994. Variation in Short Tandem Repeat sequences —a survey of twelve microsatellite loci for use as forensic identification markers. *International Journal of Legal Medicine,* 107**,** 13-20.

VACHA, M., SAEKI, M., ISOBE, O., HASHIZUME, K.-I. & TANI, T. 2001. Mapping the orientation of exciton transition dipoles along individual nanostructures of molecular J-aggregates. *Journal of Chemical Physics,* 115.

VON WURMB-SCHWARK, N., SCHWARK, T., HARBECK, M. & OEHMICHEN, M. 2004. A simple duplex-PCR to evaluate the DNA quality of anthropological and forensic samples prior short tandem repeat typing. *Legal Medicine,* 6**,** 80-88.

WANG, K. M., TANG, Z. W., YANG, C. Y. J., KIM, Y. M., FANG, X. H., LI, W., WU, Y. R., MEDLEY, C. D., CAO, Z. H., LI, J., COLON, P., LIN, H. & TAN, W. H. 2009. Molecular Engineering of DNA: Molecular Beacons. *Angewandte Chemie-International Edition,* 48**,** 856-870.

WANG, M., SILVA, G. L. & ARMITAGE, B. A. 2000. DNA-Templated Formation of a Helical Cyanine Dye J-Aggregate. *Journal of the American Chemical Society,* 122**,** 9977-9986.

WEST, W. & PEARCE, S. 1965. The Dimeric State of Cyanine Dyes. *The Journal of Physical Chemistry,* 69**,** 1894-1903.

WHITCOMBE, D., THEAKER, J., GUY, S. P., BROWN, T. & LITTLE, S. 1999. Detection of PCR products using self-probing amplicons and fluorescence. *Nature biotechnology,* 17**,** 804-7.

WILSON, D. S. & KEEFE, A. D. 2001. Random Mutagenesis by PCR. *Current Protocols in Molecular Biology.* John Wiley & Sons, Inc.

WILSON, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology,* 63**,** 3741-51.

WILSON, M. R., POLANSKEY, D., BUTLER, J., DIZINNO, J. A., REPLOGLE, J. & BUDOWLE, B. 1995. Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *BioTechniques,* 18**,** 662-9.

WU, P. & BRAND, L. 1994. Resonance energy transfer: methods and applications. *Analytical biochemistry,* 218**,** 1-13.

WÜRTHNER, F., KAISER, T. E. & SAHA-MÖLLER, C. R. 2011. J-Aggregates: From Serendipitous Discovery to Supramolecular Engineering of Functional Dye Materials. *Angewandte Chemie International Edition,* 50**,** 3376-3410.

ZHOU, Y. H., ZHANG, X. P. & EBRIGHT, R. H. Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase.

ZIĘTKIEWICZ, E., WITT, M., DACA, P., ŻEBRACKA-GALA, J., GONIEWICZ, M., JARZĄB, B. & WITT, M. 2012. Current genetic methodologies in the identification of disaster victims and in forensic analysis. *Journal of Applied Genetics,* 53**,** 41-60.

ZIPPER, H., BRUNNER, H., BERNHAGEN, J. & VITZTHUM, F. 2004. Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Research,* 32**,** e103-e103.

APPENDIX A: FULL MIQE DATA

Appendix A: Full MIQE Data

Experimental Design:

- Definition of experimental and control groups:
  - Experimental groups: unpurified DNA samples collected from hair with epithelial sheath, saliva, buccal swabs, or capillary blood obtained through diabetic lancets. All genotypes included in experimental group.
  - Control groups: All genotypes of purified DNA using a BioLine Isolate II Genomic Extraction Kit.
- Number within each group:
  - One sample in each group.
- Assay carried out by core or investigator's laboratory?
  - Assay carried out by investigator's laboratory.

Sample:

- Description:
  - Human hair, saliva, buccal swab, and blood.
- Volume/mass of sample processed:
  - Hair: 1 hair in 10 µL modified TE buffer. 1.000 µL sample size.
  - Saliva: 20 µL, undiluted. 1.000 µL sample size.
  - Buccal swab: 1 swab in 10 µL modified TE buffer. 1.000 µL sample size.
  - Blood: 2 µL in 8 µL modified TE buffer. 1.000 µL sample size.
- Microdissection or macrodissection:
  - No dissection used.
- Processing procedure:
  - Samples were collected by using forceps to extract multiple hairs; through a saline rinse followed by saliva expulsion; buccal swab; or use of a diabetic lancet on a disinfected palm followed by accurate pipetting of a

blood aliquot. Samples were prepared by suspending in modified TE buffer and lysed at 95°C for 5 minutes. Samples were centrifuged at 2000G for 10 seconds.

- If frozen, how and how quickly?
  - Samples not frozen; prepared and analyzed immediately after extraction.
- If fixed, with what and how quickly?
  - Samples not fixed.
- Sample storage conditions and duration (especially for FFPE samples)
  - Samples not stored; samples discarded after use.

Nucleic Acid Extraction:

- Procedure and/or instrumentation
  - Control groups isolated using BioLine Isolate II Genomic Prep Kit. Experimental groups not extracted.
- Name of kit and details of any modifications
  - BioLine Isolate II Genomic Prep Kit. No modifications.
- Source of additional reagents used
  - None.
- Details of DNase or RNase treatment
  - None.
- Contamination assessment (DNA or RNA)
  - Control groups: no contamination. Samples performed as expected.
  - Experimental groups: no contamination with other DNA. Sample impure with respect to cellular components, as it is a direct, unpurified sample.
- Nucleic acid quantification instrument and method
  - ThermoScientific Nanodrop 2000, as directed.
- Inhibition testing (C_q dilutions, spike, or other)

- Not performed; direct samples, inhibition expected.

qPCR Target Information:

- Sequence accession number
  - AC005010.2
- Location of amplicon
  - 37955-38112
- In silico specificity screen (BLAST, and so on)
  - Discussed in depth in **Error! Reference source not found.Error! Reference source not found.** and **Error! Reference source not found.Error! Reference source not found.**.
- Sequence alignment
  - Figure 3.2: Alignment using BioEdit, including FOR100, REV100, C probe, and T probe and Figure 3.3: Alignment using BioEdit, including FOR100, REV158, C probe, and T probe.
- Secondary structure analysis of amplicon
  - None observed using Oligo7.

qPCR Oligonucleotides:

- Primer sequences:
  - FOR100: ATCAAGTGCTTTCTGTTGACATTTG
  - REV100: GGCTACTCCCTCATAATGTAATGC
  - REV158: GCATAACCCAAAGAGTAACAAAATCAA
- Probe sequences:
  - Original C probe:
    5’[HEX]CTTGCGTCTAACTCAACTGCATTACATCGCAAGG[BHQ1]
    3’
  - Original T probe:

5’[6FAM]CTTGCGTCTAACTCAATTGCATTACATCGCAAGG[BHQ1]3’

- New C probe: 5’[HEX]CCCGCGGATATTATTTCTAACTCAACTGCATTACATTATGACCGCGGG[BHQ1]3’

- Manufacturer of oligonucleotides
  - Sigma Life Science
- Purification method
  - Primers: Standard desalting
  - Probes: HPLC

qPCR Protocol:

- Reaction volume and amount of DNA
  - 20 µL reaction volume; variable amount of DNA.
- Buffer/kit identity and manufacturer
  - BioLine SensiFAST Probes Master No ROX.
- Additives
  - None.
- Complete thermocycle parameters
  - Discussed in detail in 4.4.6: Assay Optimization.
- Reaction setup (manual/robotic)
  - Manual.
- Manufacturer or qPCR instrument
  - Qiagen Rotor Gene Q 5-plex with HRM

qPCR Optimization:

- Evidence of optimization (from gradients)

  - Figure 4.19: Gel electrophoresis image of REV158 primer annealing temperature gradient.
- Specificity (gel, sequence, melt, or digest)
  - Figure 4.18: Gel electrophoresis of direct PCR fragments, after amplification.
- Evidence for LOD
  - Figure 4.33: Initial sensitivity test, T probe, after optimization and Figure 4.34: Initial sensitivity test, C probe, after optimization.

Data analysis:

- qPCR analysis program (source, version)
  - Rotor Gene Q Series Software (Qiagen, version 2.3.1)
- Method of C_q determination
  - Manual. Threshold set above pre-exponential background.
- Description of normalization method
  - See 3.9: Normalization methods

APPENDIX B: COMPARISON OF SELECTED PROFILING METHODS

Appendix B: Comparison of Selected Profiling Methods

Method	Advantages	Disadvantages	References
VNTR	• Uses RFLP or PCR • RFLP has less worry of contamination.	• PCR risks allelic dropout due to preferential amplification of smaller alleles • RFLP can take a number of days • Only two copies per cell; potential issues with degraded	(Demers et al., 1995, Tully et al., 1993)
STR	• Autosomal • Multiallelic; higher power of discrimination per allele • Multiplex; a number of STRs can be amplified in one reaction	• Only two copies per cell; potential issues with degraded DNA • Potential artifacts, rendering complex results Potential for contamination; small	(Butler, 2005)
SNP	• Autosomal • Potential for higher success with degraded DNA • No artifacts associated • Uses qPCR or microchip assays Faster results than other methods	• Typically biallelic, limiting power of discrimination greatly • Mixture interpretation is challenging • Microchip assays are expensive • Multiplexing limited to	(Butler, 2005, Chakraborty, 1992, Gill et al., 2004, Ziętkiewicz et al., 2012)

APPENDIX C: SUMMARY OF SAMPLE PREPARATION

Appendix C: Summary of Sample Preparation

	Hair	Saliva	Buccal Swab	Blood
Pre-collection preparation	N/A	30 second rinse, 0.9% $^{w}/_{v}$ saline	30 second rinse, 0.9% $^{w}/_{v}$ saline	Isopropanol swab at site (fingertip)
Collection Equipment	Sterilized dissection forceps	Sterile 1.5 mL microfuge tube	Puritan Sterile Foam-Tipped Applicator	20-gauge hypodermic needle
Amount Collected	5-10 hairs	0.5-1.0 mL	1 swab	2.000 µL
Amount Prepared	1 hair follicle with 1-2 mm hair shaft	100.0 µL	1 swab	2.000 µL
Volume of Buffer	10.00 µL	N/A	20.00 µL	8.00 µL
Thermal Lysis	5 minutes, 98°C	5 minutes, 98°C	5 minutes, 98°C	5 minutes, 98°C
Post-lytic treatment	N/A	Centrifugation, 10 seconds at 2000 G	N/A	Centrifugation, 10 seconds at 2000 G
Volume in Reaction	1.000 µL	1.000 µL	1.000 µL	1.000 µL

APPENDIX D: IRB APPROVAL

Graduate Education and Research
Division of Sponsored Programs
Institutional Review Board

EASTERN KENTUCKY UNIVERSITY
Serving Kentuckians Since 1906

Jones 414, Coates CPO 20
521 Lancaster Avenue
Richmond, Kentucky 40475-3102
(859) 622-3636; Fax (859) 622-6610
http://www.sponsoredprograms.eku.edu

NOTICE OF IRB APPROVAL

Protocol Number: 16-033

Institutional Review Board IRB00002836, DHHS FWA00003332

Review Type: ☐Full ☒Expedited

Approval Type: ☒New ☐Extension of Time ☐Revision ☐Continuing Review

Principal Investigator: **Dr. Jamie Fredericks**

Project Title: **The Application of Molecular Beacons as a Genotyping Tool for DNA Analysis from Forensic Samples**

Approval Date: **9/8/2015** Expiration Date: **12/15/16**

Approved by: **Dr. Jonathan Gore, IRB Member**

This document confirms that the Institutional Review Board (IRB) has approved the above referenced research project as outlined in the application submitted for IRB review with an immediate effective date.

Principal Investigator Responsibilities: It is the responsibility of the principal investigator to ensure that all investigators and staff associated with this study meet the training requirements for conducting research involving human subjects, follow the approved protocol, use only the approved forms, keep appropriate research records, and comply with applicable University policies and state and federal regulations.

Consent Forms: All subjects must receive a copy of the consent form as approved with the EKU IRB approval stamp. Copies of the signed consent forms must be kept on file unless a waiver has been granted by the IRB.

Adverse Events: Any adverse or unexpected events that occur in conjunction with this study must be reported to the IRB within ten calendar days of the occurrence.

Research Records: Accurate and detailed research records must be maintained for a minimum of three years following the completion of the research and are subject to audit.

Changes to Approved Research Protocol: If changes to the approved research protocol become necessary, a description of those changes must be submitted for IRB review and approval prior to implementation. Some changes may be approved by expedited review while others may require full IRB review. Changes include, but are not limited to, those involving study personnel, consent forms, subjects, and procedures.

Annual IRB Continuing Review: This approval is valid through the expiration date noted above and is subject to continuing IRB review on an annual basis for as long as the study is active. It is the responsibility of the principal investigator to submit the annual continuing review request and receive approval prior to the anniversary date of the approval. Continuing reviews may be used to continue a project for up to three years from the original approval date, after which time a new application must be filed for IRB review and approval.

Final Report: Within 30 days from the expiration of the project, a final report must be filed with the IRB. A copy of the research results or an abstract from a resulting publication or presentation must be attached. If copies of significant new findings are provided to the research subjects, a copy must be also be provided to the IRB with the final report.

Other Provisions of Approval, if applicable: None

Please contact Sponsored Programs at 859-622-3636 or send email to tiffany.hamblin@eku.edu or lisa.royalty@eku.edu with questions about this approval or reporting requirements.