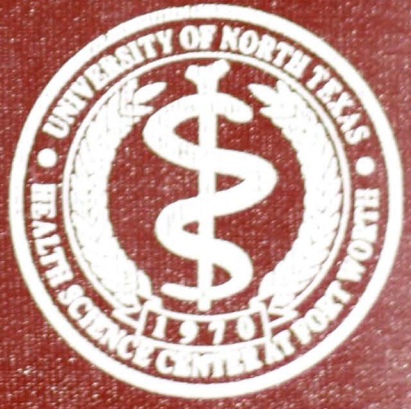
UNIVERSITY OF NORTH TEXAS
HEALTH SCIENCE CENTER AT FORT WORTH
1970

Campbell, Rowan Stewart, Automatable Virtual Array Screening System For Rapid Analysis of Mitochondrial DNA Polymorphism. Doctor of Philosophy (Biomedical Sciences, May, 2002, 156 pp., 11 tables, 48 illustrations, bibliography, 96 titles.

The goal of this research project was to develop alternative methods to traditional forensic mtDNA sequence analysis. Conventional forensic mtDNA analysis requires the direct sequencing of Hypervariable Region I and Hypervariable Region II in both the forward and reverse directions. This method is time consuming, labor intensive and expensive. Two methods for determining mtDNA haplotypes through the direct interrogation of Single Nucleotide Polymorphisms within HVI and HVII have been developed. A Sequence Specific Oligonucleotide Hybridization assay was developed on the Luminex 100 ™ flow cytometer, as well as a Single Base Extension assay developed for the ABI Prism® 310 Genetic Analyzer. The SNP typing of mtDNA sequences can provide a significant benefit in many forensic and human identification cases. The reassociation of mass disaster remains, mass grave analysis, and the screening of large numbers of crime scene samples are examples of their potential application. Their inclusion as a standard screening tool would be highly beneficial since more extensive DNA analysis would be reserved for those samples that possess the greatest evidentiary value.

In a blind study of 50 samples, the Sequence Specific Oligonucleotide Hybridization assay incorrectly identified the mtDNA haplotypes in 7 samples, whereas the Single Base Extension assay correctly identified each of the SNP positions interrogated. The SNaPshot™

primer extension assay was approximately 20 –25 times more sensitive, than the standard sequencing approach. This would suggest that this system could be a viable alternative to sequence analysis when samples are limited, as well as being more robust in detection and typing of heteroplasmic sites. A statistical evaluation of the SNP panels revealed that the genetic diversity estimated for the 50 Southwestern Hispanic samples tested was 0.9624 for the primer extension array and 0.9559 for the hybridization-based array. The probability of two randomly selected individuals from a population group having the same mtDNA haplotype was 0.0568 for the Single Base Extension assay and 0.0632 for the Sequence Specific Oligonucleotide Hybridization assay. A forensic mtDNA SNP array consisting of the positions evaluated in this study could provide a reasonable alternative to the full sequencing of the HVI and HVII regions.

AUTOMATABLE VIRTUAL ARRAY SCREENING SYSTEM

FOR RAPID ANALYSIS OF

MITOCHONDRIAL DNA POLYMORPHISM

Rowan S. Campbell, B.A., B.S.

APPROVED:

Major Professor

Committee Member

Committee Member

Committee Member

Committee Member

University Member

Chair, Department of Biomedical Sciences

Dean, Graduate School of Biomedical Sciences

AUTOMATABLE VIRTUAL ARRAY SCREENING SYSTEM FOR RAPID ANALYSIS OF MITOCHONDRIAL DNA POLYMORPHISM

DISSERTATION

Presented to the Graduate Council of the
Graduate School of Biomedical Sciences
University of North Texas
Health Science Center at Fort Worth
in Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

Rowan Stewart Campbell, B.A., B.S.

Fort Worth, Texas

May, 2002

ACKNOWLEDGMENTS

I would first like to thank Dr. Arthur J. Eisenberg for all his help in the preparation of this thesis. A special word of praise should go to him for his long lived patience, even at the height of my procrastination. I am very grateful to Drs. Bruce Budowle, John Planz, Joseph Warren and John David Tune. They were always there when I needed to know something. Their advice and especially their friendship was greatly appreciated.

This thesis is in large part due to the friendship, inspiration, and help of all the members of the DNA Identity Laboratory who have put up with me for the last seven years. I hope all my friends know that their help and friendship is greatly felt, and I thank them all from the bottom of my heart.

Finally I would like to dedicate this thesis to my family and my wife. Without their love and help I could never have completed it. They were always there when I needed them and gave me more than I could ever hoped or dreamed for.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABBREVIATIONS

A	adenine
ABI	Applied Biosystems Inc
ATP	adenosine triphosphate
AFDIL	Armed Forces DNA Identification Laboratory
bp	base pairs
C	cytosine
CE	capillary electrophoresis
CCD	charge coupled device
Cr	control region
ddATP	dideoxy adenosine triphosphate
ddCTP	dideoxy cytosine triphosphate
ddGTP	dideoxy guanine triphosphate
ddNTP	dideoxyribonucleotide triphosphate
ddTTP	dideoxy thymine triphosphate
DNA	deoxyribonucleic acid
DP	discriminating power
FBI	Federal Bureau of Investigation
G	guanine

LIF	laser induced fluoresce
MERRF	myoclonic epilepsy and ragged-red fiber
MFI	mean fluorescent intensities
mtDNA	mitochondrial deoxyribonucleic acid
nucDNA	deoxyribonucleic acid
N_x	Negative Control
OD	optically dense
PCR	polymerase chain reaction
rCRS	revised Cambridge Reference Sequence
RFLP	restriction fragment length polymorphism
rfu	relative fluorescent units
RNA	ribonucleic acid
SNP	single nucleotide polymorphism
SSO	sequence specific oligonucleotide
STR	short tandem repeats
SWGDAM	Scientific Working Group for DNA Analysis Methodologies
T	thymine
Taq	Thermus aquaticus
TMAC	tetramethylammonium chloride
U	unit
VNTR	variable number of tandem repeats

INTRODUCTION

Since 1985, deoxyribonucleic acid (DNA) typing of biological samples has become one of the most powerful tools for positively identifying an individual. Today DNA testing is commonly used for immigration, paternity, missing persons, mass disaster, military identification and criminal investigations. In forensic crime laboratories, DNA analysis has essentially replaced all other forms of serological-based tests. The use of DNA for identification provides a power of discrimination, many orders of magnitude greater than any other typing methods previously available (Weedn and Hicks, 1997).

Background of DNA Typing

In 1944 Oswald Avery identified a cellular component responsible for the transmission of heritable traits (Avery et al., 1944). James Watson and Francis Crick in 1953 were able to identify the structure of the DNA molecule as a double helix (Watson and Crick, 1953). Edward Southern in 1975 developed a protocol which enabled electrophoretically separated DNA fragments to be transferred efficiently from an agarose gel to the surface of a solid support or membrane (Southern, 1975). This technique, which has become known as Southern transfer or Southern blotting method, provided one of the first tools for elucidating the structure and function of the human genome. In 1980, David Botstein was the first to utilized variations found within the DNA of individuals as genetic markers (polymorphisms) for the construction of

a human genome map (Botstein et al., 1980). The first genetic variations detected were based upon length polymorphisms resulting from the creation or elimination of restriction enzyme recognition sites. The technique used to identify these polymorphisms became known as the restriction fragment length polymorphism (RFLP) method. The most polymorphic loci identified were those shown to contain length variation of a tandomly repeating unit. Because of their discriminatory power (DP), variable number tandem repeat (VNTR) loci proved ideal candidates for use in forensic DNA testing (Wyman and White, 1980; Jeffreys et al., 1985).

In 1982, Actagen Corporation, was formed with the idea of utilizing VNTR polymorphisms for the purpose of human identification. In 1985 Actagen Corporation changed its name to LifeCodes Corporation, which developed RFLP technology utilizing a series of DNA probes to identify genetic polymorphism at unique chromosomal locations. Concurrently in 1984, Alec Jeffreys pioneered a methodology that became commonly known as "DNA fingerprinting". The multilocus DNA probes developed by Alec Jeffreys enabled simultaneous identification of polymorphisms at a large number of genetic loci and the resultant pattern resembled that of a bar code (Jeffreys et al., 1985). In 1986, the polymerase chain reaction (PCR) was invented by Kary Mullis, who was awarded a portion of the Nobel Prize in chemistry for his work (Mullis et al., 1986). The PCR process provided the forensic community a means of analyzing minute and degraded quantities of DNA that could not be typed by conventional RFLP methods. Similar to loci used for RFLP analysis, PCR is used to amplify short tandem repeats (STR) loci that are 2 to 7 base pairs in length. RFLP and PCR technologies have been the cornerstones of forensic DNA typing.

Application of DNA Typing

The first use of DNA typing methods were for verification of familial relationships in both immigration and paternity cases (Jeffreys et al., 1986). Families have been reunited, both in the United States and abroad, based on the identification of a child or sibling by DNA typing. Paternity testing that previously relied on protein-based analysis, is now almost exclusively performed by DNA typing. In the year 2000, more than 300,000 paternity tests were done in the United States by DNA typing methods (AABB, 2000).

DNA typing provides valuable clues for the identification of kinship. The identity of biological remains may be verified by typing putative parents, offspring, or other family members and comparing their DNA profiles to that of the discovered remains. Similarly, a child abducted at a very young age might be reunited with his or her biological parents based on the outcome of comparative DNA testing.

Following a mass disaster, DNA typing has proven to be an essential tool for identification. The 1993 Branch Davidian standoff in Waco, Texas required that badly charred remains of numerous bodies be identified using PCR DNA typing methods (Scruggs et al., 1993). Tissue samples exposed to saltwater following an airplane crash such as TWA flight 800 (Altman, 2001) and Swiss Air flight 111 (Altman, 2001) were successfully typed by DNA analysis. Identification of individuals found in the 9/11 World Trade Center wreckage who could not be recognized by traditional methods (visual, dental, tattoo's, etc.) are being identified through the use of DNA typing techniques (Jenkins, 2001). Similarly, the Armed Forces DNA Identification Laboratory (AFDIL) has identified the remains of numerous soldiers who died during the Vietnam War using PCR DNA typing methods (Holland et al., 1995). One of these individuals included the soldier who was laid to rest in the tomb of the Vietnam Unknown

Soldier. The remains were exhumed and positively identified fourteen years post mortem as Air Force 1st Lt. Michael J. Blassie (Daoudi et al., 1998). Cognizant of the usefulness of DNA, the military has established a repository in which samples are routinely collected and stored from all members of the armed forces (Butler and Levin, 1998). DNA extracted from samples stored in this repository can serve as a reference for the positive identification of remains.

DNA in Forensics

The first forensic application of DNA typing took place in England in 1986 utilizing the "DNA fingerprinting" technology developed by Alec Jeffreys (Jeffreys, Wilson et al., 1985). During a three year period two young girls were found raped and murdered in Leicester, England. A seventeen-year-old male was arrested and confessed to both murders. DNA fingerprint patterns from semen recovered from both victims however, did not match the pattern from a blood sample obtained from the police's primary suspect. Based upon DNA evidence, the original suspect was released. After an extensive police investigation, the British Home Office was able to identify Colin Pitchfork as the individual responsible for the murders of the two young girls. His DNA was analyzed and the patterns found were identical to that of the sperm samples recovered from the victims (Wambaugh, 1989).

Since DNA testing was first introduced, it has been used in thousands of criminal cases. Often it is the definitive evidence that aids in the conviction of the guilty parties. DNA typing of biological evidence is the most powerful method available for the identification of an assailant in a crime. Evidence containing DNA recovered from a crime scene has been used to file arrest warrants for unnamed suspects before the statue of limitation for a particular crime expires. The identification of a suspect by only his or her DNA profile has been considered sufficient, since

warrants can be brought against unnamed individuals based on other identifying characteristics (Lonsway, 2000). These "warrants", based only on a DNA profile, have been issued in several states. FBI statistics show that in rape cases approximately 33% of the time, DNA testing excludes the police's primary suspect (Erickson, 1991). The exclusion of a suspect by DNA is absolute. Wrongly convicted men continue to be freed from prison after old biological evidence has been re-tested using DNA typing techniques (Butterfield, 1996).

One of the more creative uses of DNA was in the conviction of one of the men accused in the first World Trade Center Bombing. On March 3rd, 1993, a typewritten letter was received at the New York Times. The individual claimed responsibility for the bombing of the World Trade Center in the name of Allah. Minute amounts of DNA recovered from the saliva used to lick the envelope were analyzed by PCR. The STR profiles were shown to match that of one the defendants, Nidel Ayyad, a chemist working for the Allied Signal Corporation in New Jersey (Williams, 1998).

Cellular DNA

The majority of genetic material comprising the human genome resides in the nucleus of each cell. An additional source of DNA can be found within the mitochondria, an organelle which is often referred to as the powerhouse of the cell due to the oxidative phosphorylation pathway it contains (Alberts et al., 1989).

Nuclear DNA (NucDNA) of autosomal cells are diploid, containing two copies of each genetic locus, one of which is inherited via the egg from the mother and the other copy is inherited from the father via sperm. Chromosome shuffling and recombination of DNA loci between homologous chromosomes during meiosis, results in unique variations found among

individuals (Alberts, Bray et al., 1989). In contrast, mitochondrial DNA (mtDNA) is considered haploid with the DNA only inherited maternally. Within the mitochondrial genome, there is no recombination; an individual contains the same mtDNA sequence as other members of his/her maternal lineage (Giles et al., 1980).

Within the nucleus of each diploid cell there are two genomic equivalents of DNA (Alberts, Bray et al., 1989). However, within a single mitochondrion there may be 8-10 copies of mtDNA (Alberts, Bray et al., 1989). A mature oocyte is estimated to contain thousands of mitochondria and more than 100,000 copies of mtDNA (Piko and Matsumoto, 1976; Michaels et al., 1982). Depending upon the tissue type, cells can contain between 200 to 1700 copies of the mtDNA genome (Bogenhagen and Clayton, 1980; Robin and Wong, 1988). However, the total amount of genetic information in mtDNA is much less than that of nucDNA. Human nucDNA contains about three billion base pairs (bp) that code for around thirty thousand genes, is linear, and organized into chromosomes (Venter, 2001). Human mtDNA contains approximately 16,559 bp, that codes for 37 genes (Wallace, 1992), is circular and more closely resembles a prokaryotic genome than a eukaryotic chromosome (Alberts, Bray et al., 1989). The function and structure of mtDNA is very different from prokaryotic DNA from which it is theorized to have evolved (Gray, 1992). Mitochondria are believed to have descended from a free living prokaryote that formed a symbiotic relationship with a eukaryotic (nucleus-containing) cell in the distant past (Gray, 1992). This symbiotic relationship with the host cell enabled mitochondria to lose some of the redundant resources necessary for independent existence and share the resources of its host (Scheffler, 1999).

MtDNA codes for 13 enzyme subunits, 2 ribosomal RNAs and 22 transfer RNAs (Wallace, 1992). The mitochondrial genome also contains an 1100 bp region of non-coding

DNA referred to as the displacement loop (D-loop) or control region (Cr). The origin of mtDNA replication occurs within this region (**Figure 1**) (Alberts, Bray et al., 1989). The control region of the mitochondrial genome mutates at a rate 5 to 10 times that of nucDNA. This elevated mutation rate results from a lack of a polymerase mismatch repair system and relaxed selective pressure (Brown et al., 1979). It has been estimated that mtDNA varies 1% to 2.3% between unrelated individuals (Greenberg et al., 1983). This is the equivalent of 1 or 2 nucleotides out of every 100 bp. The D-loop or control region contains two hypervariable regions designated Hypervariable I (HVI) and Hypervariable II (HVII) that are well suited to forensic analysis due to their high sequence variability. The designation of the location of these regions is based upon a numbering system following the first sequenced mitochondrial genome (Anderson et al., 1981). The current reference standard (i.e. revised Cambridge Reference Sequence (rCRS)) used for mtDNA analysis is based upon minor revisions to the original Anderson sequence (Andrews et al., 1999). The two regions commonly used in forensic analysis, HVI and HVII, typically contain nucleotides 16024 through 16365 and 73 through 340, respectively. When reporting the sequence analysis of the control region, rather than including the entire 650 bp of data, only differences (i.e. base substitutions, additions or deletions) compared to the rCRS are recorded as the mtDNA haplotype.

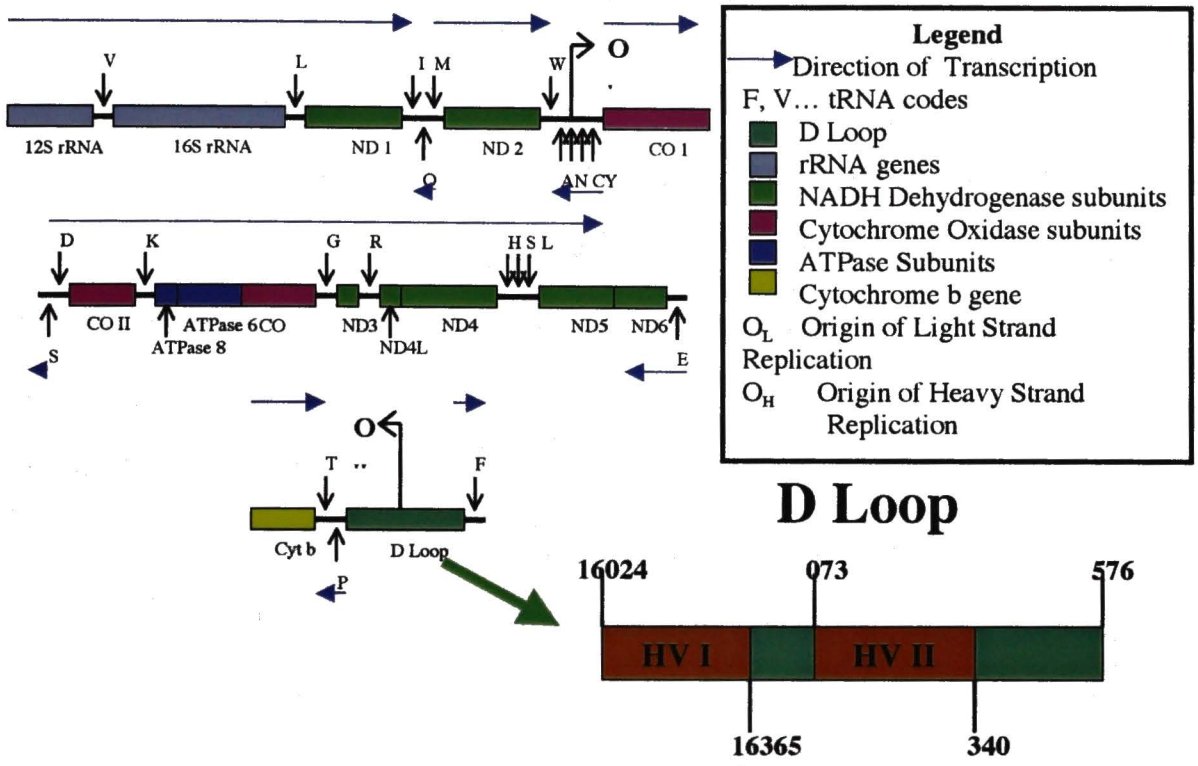

Figure 1. Map of human mitochondrial genome including expanded diagram of non-coding region. Listed are the genes for 12S and 16S ribosomal RNAs, subunits of NADH, Cytochome *c* oxidase complex, cytochrome *b*, ATP synthase and 22 tRNA's. Numbering system follows that of the standard reference sequence (Anderson, Bankier et al., 1981; Planz, 1999).

Forensic Application of mtDNA Analysis

Both nucDNA and mtDNA are routinely used for the evaluation of forensic evidentiary samples. Since each human cell may contain several thousand of copies of the mtDNA genome, there is a greater chance of recovering mtDNA from degraded biological samples. These include highly decomposed corpses, (Sullivan et al., 1992) human skeletal remains (Holland et al., 1993), hairs that do not have sheath material attached, e.g. hair shafts (Wilson et al., 1995) and teeth (Ginther et al., 1992). In order to identify the source of evidentiary material, a known reference sample is required. Since mtDNA is inherited maternally, any individual along a maternal lineage can be used as a reference sample. This is advantageous when only distant relatives are available as a reference sample source.

Non-forensic Application of mtDNA Analysis

A number of serious human diseases caused by mutations in gene-coding regions of the mtDNA molecule have been identified (Momoi, 1993). Common syndromes associated with mitochondrial tRNA or protein-coding mutations are myoclonic epilepsy and ragged-red fiber (MERRF), mitochondrial encephalomyopathy, and lactic acidosis. In addition, molecular anthropologists have been using mtDNA sequence analysis to study genetic variation in human populations throughout the world. MtDNA analysis has provided insight into immigration patterns and geographic origins of ancient populations (Cann et al., 1987). Recently, mtDNA was extracted and sequenced from a Neanderthal skeleton found in Europe. The sequence data indicate that Neanderthals are not ancestors of modern man and must have taken a completely separate evolutionary path than that of the modern humans, which exist today. It would appear that Neanderthals never shared an evolutionary path with homo *sapiens* and became extinct (Krings et al., 1997). The same sequencing technologies used in medical and anthropological studies are used in forensic identification.

Current mtDNA Analysis Protocols

DNA can be recovered from many biological samples including hair, bone, teeth, blood, feces and urine. PCR is used to amplify the two hypervariable regions (HVI and HVII) of the mitochondrial genome when insufficient nucDNA is present. Amplification of mtDNA is accomplished using specific primers sets that create overlapping amplicons covering the regions of interest. Due to the extreme sensitivity of the mtDNA analysis process, precautions are required to eliminate the introduction of exogenous DNA during both the extraction and

amplification steps. These precautions include dedicated laboratory space, equipment, reagents and materials. Separate areas within the lab are used for DNA extraction, PCR amplification setup, and post amplification analysis. Laboratory personnel are required to wear gloves, masks, and lab coats. Each of these items must be removed prior to moving from one area of the lab to another. Additional procedures for minimizing potential contamination typically include ultraviolet irradiation of space, equipment, and supplies, the use of aerosol-resistant barrier pipette tips and autoclaving of tubes and reagent stocks (Wilson et al., 1995; Wilson, Polanskey et al., 1995).

Following PCR amplification, sequencing reactions are performed. The Sanger dideoxynucleotide sequencing method (Sanger et al., 1977) is commonly utilized for mtDNA analysis. This method incorporates nucleotides that contain the sugar dideoxyribose in place of deoxyribose. The nucleotide dideoxyribose lacks the essential 3'-OH group, which is required for the attachment of the next nucleotide in chain elongation. Therefore, the incorporation of a dideoxynucleotide immediately terminates further DNA extension. Each dideoxynucleotide terminator is labeled with a different fluorescent dye (C, dTAMRA; A, dR6G; T, dROX; G, dR110). The fluorophore-labeled DNA fragments created by the cyclic amplification process are separated based upon size using capillary or gel electrophoresis. Migration of the DNA fragments past a laser results in the excitation of the fluorophore. Each fluorophore attached to a dideoxynucleotide, emits light a unique wavelength. A CCD camera detects the wavelength of light emitted by the fluorophore, which is then translated by a computer into a base sequence. Instrumentation such as the Applied Biosystems Inc., (ABI) 310 capillary electrophoresis system enables automation of the process. HVI and HVII are sequenced in both the forward and reverse

directions. This redundancy provides confirmation of the mtDNA nucleotide sequence and improves the accuracy of sequencing.

The standard protocol in most forensic labs is to process the reference and evidentiary samples separately. The purpose is to ensure that the evidentiary material does not get contaminated by the DNA from the reference samples. The sequences from the evidentiary and reference samples are then compared to determine if they are concordant. In the event the sequences match, the Scientific Working Group on DNA Analysis and Methodology (SWGDAM) mtDNA database, which is maintained by the FBI, is searched to estimate the relative frequency of the haplotype. The number of times this haplotype has been observed in the database is reported. A upper 95% confidence interval is frequently calculated as a percentage of individuals within the population who are estimated to be a particular haplotye based on the size of the database examined.

Limitation of mtDNA Analysis

One of the most significant concerns in mtDNA analysis is the contamination of evidentiary material with extraneous sources of mtDNA. The contamination can occur either during initial collection and handling of evidentiary samples or through introduction of cells or amplified product within the lab. Stringent precautions are necessary to insure that this contamination does not occur.

Sequence analysis of the control region can be difficult as a result of homopolymeric stretches of cytosine residues that reside in the HVI and HVII regions. There is a common polymorphism at position 16,189 in which a thymine to cytosine transition commonly occurs. The result is a long stretch of 10-13 cytosine residues, which promotes destabilization of the

sequencing chemistry. This makes the interpretation of any sequence beyond this point difficult without further testing (Bendall and Sykes, 1995; Butler and Levin, 1998). Sequencing from the opposite direction and use of sequence primers flanking the hompolymeric tracts are often required to obtain clean sequencing data.

Since all individuals in a maternal lineage share the same mtDNA sequence, mtDNA cannot be considered unique to any individual. The frequency of occurrence of a particular haplotype is determined by counting the number of times that sequence is seen in a database. Currently, there are approximately 4100 individual human mitochondrial HVI and HVII sequences in the FBI's database (Fischer et al., 2000). In Caucasians the haplotype (263G, 315.1C) appears in approximately 7% of Caucasians. However, in general, the haplotypes observed in most populations around the world, with the exception of a few populations of anthropological interest, are rare, and relatively few types are present with frequencies greater than 1%. Almost two thirds of newly typed samples contain novel sequences, previously unseen.

Heteroplasmy

Heteroplasmy is defined as the existence of more than one population of mtDNA genomes within an individual. When PCR amplified mtDNA is sequenced, individuals typically harbor a single mitochondrial DNA type that is distinguishable from that of other maternal lineages. Thus most individuals are considered homoplasmic. However, due to genetic bottlenecks between generations (Howell, 1992) and depending upon the sensitivity of the method used for detection, heteroplasmy may be observed. Heteroplasmy can exist within the mtDNA population as two different types; either as length heteroplasmy or site heteroplasmy.

Length heteroplasmy is where subpopulations of mtDNA genomes exhibit either an insertion or deletion of a base or bases Within HVI, most people have a T at position 16,189 that interrupts a run of C's on each side. When a "C" is substituted for the T at 16,189 this creates an unbroken run of C's that is apparently replicated with low fidelity by the mtDNA polymerase system. This result is a population of mtDNA genomes that differ in length. The electropherogram traces down stream of the heteroplasmy are often unintelligible. If encountered, additional PCR amplification or sequencing reactions are required. Length heteroplasmy also occurs in HVII C-stretch positions 303-315, but in this case there is often an identifiable predominant length variant, even when heteroplasmy is present (Parson et al., 1998). Site heteroplasmy is defined as the appearance of two or more bases at one position in an otherwise unmixed sequence. Site heteroplasmy can be confirmed with an independent PCR and sequencing of both light and heavy stands to rule out sequencing artifacts. Heteroplasmy has the potential to complicate or strengthen forensic identity testing and must be taken into account. Fortunately, the frequency of heteroplasmy in the population (when analyzed by DNA sequence analysis is relative low (i.e. 2-8% of the population) (Monnat and Loeb, 1985). The first documented instance of heteroplasmy within the human CR occurred during the identification of the skeletal remains of Tsar Nicholas II (Gill et al., 1994; Ivanov et al., 1996).

Evaluation of Current mtDNA Testing Procedures

MtDNA sequence analysis is the most rigorous and time-consuming of all DNA forensic testing. Based on informal statistics available from forensic laboratories performing mtDNA sequencing, the rate of throughput is approximately 1-4 cases/analyst/month (Planz, 2002). The reasons for this include: 1) small/degraded samples requiring numerous PCR reactions to obtain

sufficient DNA template for sequencing, 2) exhaustive procedures to control for contamination, and 3) sequence analysis of both forward and reverse strands of DNA from HVI and HVII regions. The sequence analysis can be further complicated if heteroplasmy is detected.

Despite the inherent cost and difficulties in performing mtDNA sequence analysis, mtDNA typing can provide probative evidence in cases where nucDNA typing is not possible. MtDNA forensic testing is utilized primarily in situations where nucDNA typing is not an option, or as a fallback method in the event that nuclear typing has been attempted and was unsuccessful. In these cases, mtDNA typing can provide additional information about the potential source of a biological sample heretofore unavailable. As of 2002, dozens of cases have been tried in US courtrooms using mtDNA sequence analysis to augment more traditional forms of evidence, and several post-conviction exonerations have been obtained (Conners et al., 1996).

As the need to incorporate mtDNA analyses into forensic casework has increased, methods to increase the ease and throughput of mtDNA typing have been developed. The use of methods that analyze intra-specific mtDNA variation has seen increased attention in recent years. Most notable of these systems was the development of a sequence specific oligonucleotide (SSO) probe hybridization assay which interrogated nine regions of the mtDNA control region with 23 probes (Stoneking et al., 1991). The application of this method was extensively used to characterize the extent of population heterogeneity expressed in Asians, Europeans and sub-Saharan Africans (Melton and Stoneking, 1996; Melton et al., 1997; Melton et al., 1997). Gel-based minisequencing assays were also developed to expedite mtDNA sequence analysis for crime laboratory and databasing operations (Sullivan et al., 1995; Tully et al., 1996). The minisequencing format interrogates nucleotide substitutions at 13 polymorphic positions in the HVI and HVII regions of the mtDNA control region. Random match probabilities from this

system were reported as 0.054 and 0.026 for British Caucasians and Afro-Caribbean's, respectively. Comparable results were reported for these populations using an SSO assay (Stoneking, Hedgecock et al., 1991). A reverse dot-blot typing kit has been adapted from the SSO system (Reynolds et al., 1996). These SSO systems were recently used to identify missing individuals discovered in mass graves situated throughout Croatia and Southern Bosnia (Gabriel et al., 2001). More recently, genetic bit analysis (Nikiforov et al., 1994) and time-of-flight mass spectrophotometry (Butler, 1999) have been used to directly assay several of the single nucleotide polymorphisms (SNPs) responsible for the sequence heterogeneity detected in human populations. The utility and discrimination power of a mtDNA SNP array was demonstrated by analyzing individuals from six populations using 10 polymorphic sites (Planz et al., 1997). Planz et al. (Planz, Fish et al., 1997), found a 10-SNP panel was capable of providing approximately 85-90 percent power of discrimination in most of the population sets screened. A major advantage to the direct SNP assay approach is the ability to unambiguously determine the base composition of sites exhibiting heteroplasmy.

Presently, the Federal Bureau of Investigation (FBI) lists seven laboratories in the US as being online for mtDNA sequence analysis for forensic testing (Isenberg and Moorem, 1999). In the past 3 years, the number of samples submitted for mtDNA analysis has radically increased. Previously the majority of mtDNA cases consisted of limited numbers of samples, such as an unknown and a reference sample. Most cases submitted now include numerous reference samples in addition to the evidentiary material. If heteroplasmy is detected in an evidentiary sample, the number of reference samples that may be needed to be processed can potentially increase to 10-15 samples per individual case (Planz et al., 1999; FBI DNA Analysis Unit II, 2000). The increased number of samples for mtDNA typing has in many instances

overwhelmed the capabilities of the testing laboratories. The increased number of samples analyzed per case also results in a significant increase in the average cost per case. Testing methodologies are needed that can provide significant cost, operational and informational advantages over current practices.

An Automatable Virtual Array Screening System for Rapid Analysis of Mitochondrial DNA Polymorphisms

The goal for the proposed research is three-fold: 1) to develop and evaluate novel mtDNA-typing assays utilizing equipment, that is readily available in crime labs; 2) provide a more efficient methodology for mtDNA typing, especially in cases where numerous samples must be screened; 3) determine heteroplasmic positions without sequencing. Two distinct systems were developed in which each could interrogate specific mtDNA SNP positions in lieu of offering direct sequencing.

The first was a hybridization-based Sequence Specific Oligonucleotide (SSO) method utilizing Luminex LabMAP™ system developed by Luminex Corporation, Austin, Tx. The second system was the direct interrogation of SNP loci via primer extension utilizing the ABI Prism® Snapshot™ dideoxyribonucleotide triphosphate (ddNTP) primer extension kit developed by Applied Biosystems, Foster City, Ca. Following development and refinement of these two systems, a blind study was carried out utilizing samples donated by the FBI.

The LabMAP™ system utilizes polystyrene microspheres that are internally dyed with two spectrally distinct fluorophores. Using precise ratios of these fluorophores, an array of colored beads is created consisting of a set of 100 different microspheres each with a specific spectral address. The surface of the microsphere contains a reactive group to which a probe can be attached. A capture oligonucleotide probe complementary specific mtDNA SNP can be

attached to each microsphere. A different oligonucleotide probe is attached to each spectrally distinct bead. Since microsphere sets can be distinguished by their spectral addresses, they can be multiplexed in a single reaction vessel thus enabling multiplexing, of up to 100 different analytes. A third fluorophore, for example Bodipy TMR, coupled to the PCR amplicon, identifies the hybridization reaction between the probe and the amplified product that has occurred on the microsphere surface (**Figure 2**).

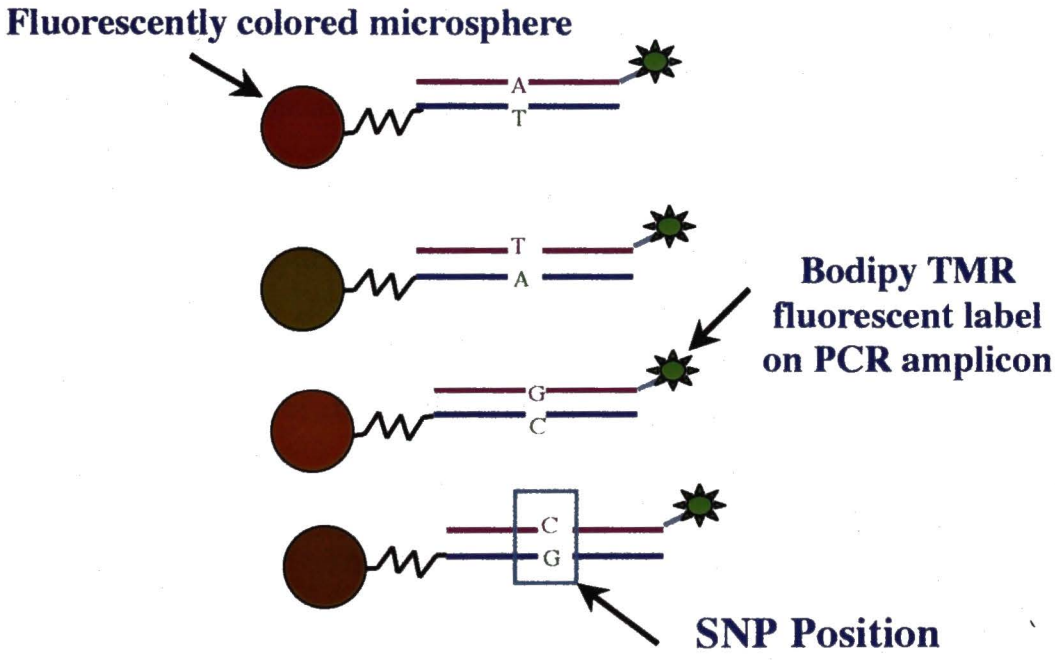

Figure 2. Hybridization assay using Luminex™ color coded microspheres and a Bodipy TMR fluorescent labeled PCR amplicon. Each color shade of the microsphere represents a distinct SSO probe, which interrogates a specific SNP position.

Microspheres are interrogated individually in a rapidly flowing fluid stream as they pass by two separate lasers in the Luminex 100™ Analyzer. High-speed digital signal processing classifies the microsphere based on its spectral address and quantifies the reaction on the surface in a few seconds per sample.

To evaluate the utility of this platform for mtDNA SNP screening, a panel of complementary capture oligonucleotide probes were obtained and coupled to microspheres to interrogate 31 different SNP positions utilizing a SSO hybridization assay. Haplotype data from different population sample sets were used to determine a panel of SNP positions that could provide the highest level of discrimination for identity testing purposes (Stoneking, Hedgecock et al., 1991; Melton and Stoneking, 1996; Tully, Sullivan et al., 1996; Planz, Fish et al., 1997). Ideally when these SNP position are interrogated, the SNP haplotype obtained would provide a power of discrimination approaching that of sequencing. Also to be addressed was the potential of the SNP panels to differentiate heteroplasmy.

The second method to be developed and tested was the direct interrogation of a SNP locus via primer extension. In this format, an oligonucleotide primer is designed to be complementary to the sequence immediately adjacent to the SNP of interest. In the presence of a polymerase and an excess of all four fluorescently labeled ddNTP, the primer is extended by a single base. The nucleotide present at the SNP side is determined by a single base addition of a fluorescently-labeled ddNTP. The incorporated ddNTP then terminates at the SNP site **(Figure 3)**.

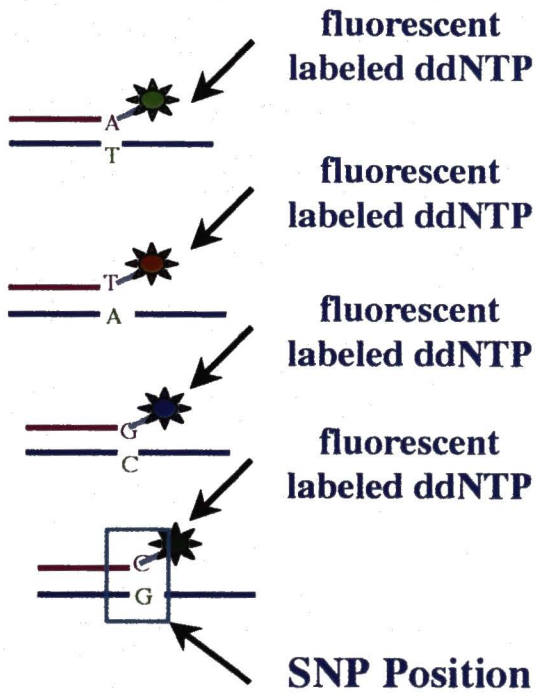

Figure 3. Single Base Extension. Each fluorescently label ddNTP identifies the termination of a cycle sequence reaction.

With the use of capillary electrophoresis (CE), high-resolution electrophoretic separation of the extended primer panel can be achieved in a semi-automated fashion. Correlating the primer length with the migration time during CE, one is able to determine which SNP is being interrogated. With the laser-induced detection of the fluorescently-terminated primer products the extended base can be determined. Thus, the rapid generation of a number of SNPs for a mtDNA haplotype is possible. In this study, typing mtDNA haplotypes were accomplished by capillary electrophoresis using the ABI Prism® 310 Genetic Analyzer (Applied Biosystems).

In conclusion genomic DNA typing has developed into one of the most robust methods of identification since the discovery of the human fingerprint. In addition to genomic DNA, mtDNA, a second human genome is found in the cell in multiple copies and can serve as an alternate source for DNA typing. While techniques for nuclear DNA are well defined, alternate

methods are needed to facilitate mtDNA typing. The goal of the proposed research was to develop and evaluate mtDNA-typing assays utilizing both a hybridization-based SSO method on the Luminex 100™ detection platform and direct interrogation of SNP loci via primer extension using the ABI Prism® 310 Genetic Analyzer. Multiplexed arrays of mtDNA oligonucleotides were developed to identify a panel of SNP sites. The typing system was validated through comparison with actual sequence data generated from samples prepared for inclusion in the FBI mtDNA database as well as testing in simulated and non-probative forensic samples.

MATERIALS AND METHODS

Samples

Blood samples obtained from individuals within the DNA Identity Lab (UNTHSC) were utilized for the development of mtDNA typing protocols. The HVI and HVII mtDNA sequences from each volunteer were previously determined utilizing the Sanger dideoxyribonuclotide method (Sanger, Nicklen et al., 1977). To compare the various mtDNA typing systems, anonymous blood samples were provided by Dr. Bruce Budowle of the FBI's Laboratory Division. These blood samples were provided as dried stains on Whatman FTA™ Paper. The HVI and HVII sequences for each of these samples were known, but were not made available until the SNP analysis had been completed.

Extraction of Mitochondrial DNA From Whole Blood

Mitochondrial DNA was isolated from whole blood samples utilizing a procedure adopted by the UNTHSC Paternity Lab for the isolation of genomic DNA (Grimberg et al., 1989). The blood samples were vortexed immediately prior to use in order to resuspend cells. A 0.5 mL aliquot of whole blood was removed and placed into a 1.5 mL microcentrifuge tube. A 1.0 mL aliquot of ice cold (4°C) Cell Lysis Buffer (320mM sucrose, 10 mM Tris-HCL, 5mM $MgCl_2$, 1%(v/v) Triton X-100) was added to each sample. The microcentrifuge tubes were capped and vortexed for 10 to 15 seconds. The tubes were then centrifuged at 1,400 x g for 5

minutes to pellet the white cell nuclei and other cellular organelles. The supernatant was discarded into a biological hazard waste container. Care was taken not to disturb the pellet. The residual fluid was blotted onto an absorbent tissue and discarded appropriately. Another 1.0 mL aliquot of Cell Lysis Buffer was added. The pellets were resuspended and the above process repeated for a total of two washes.

A 300 μL aliquot of a solution containing Protein Lysis Buffer (10 mM Tris-HCL, 10 mM EDTA, 10mM NaCl, 0.5% SDS pH 7.4) and 300 mg/mL Proteinase K (*Tricharium album*) was added to each microcentrifuge tube. In order to break up the pellets, the Protein Lysis Buffer solution was pipetted up and down several times. The samples were vortexed for 15 seconds and placed in a 55°C heat block for a minimum of one hour. During the 55°C incubation, each tube was vortexed for 15 seconds, at 15-minute intervals to ensure that the pellet was resuspended.

At the end of the 55°C incubation, the tubes were removed from the heat block and vortexed for 30 seconds. The tubes were then centrifuged for 2 minutes at 15,000 x g to pellet any cellular debris. The supernatants were carefully poured into a new 1.5 mL microcentrifuge tube. Proteins were precipitated by the addition of 150 μL 7.5 M LiCL to the supernatant, briefly vortexed, and placed on ice for 10 minutes. The tubes were then centrifuged for 15 minutes at 15,000 x g. Degraded proteins are effectively precipitated, leaving DNA in solution. The supernatant is carefully transferred to a new tube to which 1 mL of room temperature 100% EtOH is added. The tubes are gently mixed by inversion followed by centrifugation for 10 minutes at 15,000 x g. The supernatant was decanted and disposed of appropriately. A 1 mL aliquot of 70% EtOH was added to the precipitated DNA, and briefly vortexed. The tubes were centrifuged for 5 minutes at 15,000 x g. The supernatant was decanted and the residual alcohol

was blotted dry. The microcentrifuge tubes were placed in a Speed Vac to remove any residual EtOH. The DNA pellet was resuspended in 750 µL of TE^{-4} (10 mM Tris-HCL, 0.1 mM EDTA, pH 8.0). The DNA was vortexed and incubated at 55°C for 5 minutes. The DNA samples were then quantitated using a yield gel.

Extraction of DNA from FTA™ Bloodstain Cards

For each sample analyzed, two MicroAmp® tubes (200 µL capacity) were labeled with HV1 or HV2. From each FTA™ bloodstain card, four 1.2 mm paper punches were removed using a Harris Punch (Fitzco Corp.) Two of the 1.2 mm FTA punches from each individual were placed into each tube labeled HV1 and HV2. A 100 µL aliquot of FTA™ Purification solution (10mM Tris-HCL, 10mM NaCl, 1mM EDTA, 0.5% Triton X-100) was added to each tube containing the FTA™ punches. The tubes were capped and inverted repetitively for one minute. The tubes were then incubated at room temperature for three minutes. The tubes were then inverted for another minute. The FTA™ Purification solution was carefully removed from the tubes using a disposable pipette. It is essential that the two punches remain at the bottom of the tube when removing the FTA™ Purification solution. A second 100 µL aliquot of FTA™ Purification solution was added to the two FTA™ Paper punches. The tubes were capped and inverted repetitively for a total of one minute. The tubes were then incubated at room temperature for three minutes and inverted for another minute. The FTA™ Purification solution was carefully removed from the tubes using a disposable pipette. The two punches were washed for a third and final time with 100 µL of FTA™ Purification solution. During the third wash, the tubes were incubated for five minutes at room temperature. The tubes were not inverted during the third wash in order to avoid the fragmentation of the FTA™ Paper punches. The

FTA™ Purification solution was removed and 100 µL of TE^{-4} was added to the punches. The tubes were inverted several times and incubated at room temperature for 5 minutes. The TE^{-4} was removed with a disposable pipette and a second 100 µL aliquot of TE^{-4} was added to the punches. The punches were incubated at room temperature for five minutes. After the TE^{-4} was removed, and the tubes were placed in a 37°C incubator for two hours with the caps open to ensure adequate drying of the punches. The tubes caps were capped and immediately vortexed to ensure that the punches were not bound to the side of the MicroAmp® tubes (Eisenberg, 2000).

DNA Quantitation

The concentration of the DNA extracted from the whole blood samples was determined by gel electrophoresis on a 1% agarose yield gel. The gels were run in Tris-Acetate-EDTA (40 mM Tris, 20 mM Acetate, 1 mM EDTA pH 8.0) buffer containing eithidium bromide (0.5 mg/L)(Maniatis et al., 1989). In order to determine the approximate concentration of the DNA in each sample, quantitation standards (100 ng, 75 ng, 50 ng, 25 ng, and 12.5 ng of genomic human DNA (Invitrogen Corp, MD) were loaded onto the gel. Electrophoresis was carried out at 100 volts for 10 minutes and the DNA was visualized under a UV-light (302 nm). The yield gel also provided a qualitative assessment of the integrity of the isolated DNA.

PCR Amplification for Single Base Extension Assay

The mtDNA PCR amplicons analyzed by the SNaPshot™ ddNTP single base extension assay were obtained using the primer sets specified by the FBI DNA Unit II (FBI DNA Analysis Unit II, 2001) (**Table 1**). The amplification of HVI is typically divided into two separate reactions in which the primer pair A1-B2 is used for HVIA and A2-B1 is used for HVIB. For

the development of the single base extension assay, the amplification of HVI was accomplished in a single reaction using the primer pair A1-B1. The amplification of HVII is also typically divided into two separate reactions using the primer pair C1-D2 for HVIIA and C2-D1 for HVIIB. For the development of the single base extension assay, the amplification of HVII was done in a single reaction using the primer pair C1-D1.

Table 1. Amplification Primers for Single Base Extension Assay

Hypervariable Region I

A1	(L15997)	5'-CACCATTAGCACCCAAAGCT-3'
B2:	(H16236)	5'-CTTTGGAGTTGCAGTTGATG-3'
A2:	(L16159)	5'-TACTTGACCACCTGTAGTAC-3'
B1:	(H16391)	5'-GAGGATGGTGGTCAAGGGAC-3'

Hypervariable Region II

C1:	(L048)	5'-CTCACGGGAGCTCTCCATGC-3'
D2:	(H285)	5'-GGGGTTTGGTGGAAATTTTTTTG-3'
C2:	(L172)	5'-ATTATTTATCGCACCTACGT-3'
D1:	(H408)	5'-CTGTTAAAAGTGCATACCGCCA-3'

The amplification mix consisted of 1X PCR Buffer (10 mM Tris-HCl pH 9.0 at 20°C, 50 mM KCl, Applied Biosystems), 160 μg/mL BSA, 1.5 mM $MgCL_2$ (Applied Biosystems), 200 μM dNTP's (Applied Biosystems), 5 units Platinum *Taq* Polymerase (Invitrogen Inc.), and 0.4 μM of the appropriate HVI or HVII primer pair. Approximately 5 ng of total human DNA or

two of the 1.2 mm FTA™ Paper punches were added to each amplification reaction. DNA samples were amplified in a 50 µl reaction using the GeneAmp® 9700 Thermal Cycler (Perkin Elmer, Applied Biosystems). The thermocycling parameters were those recommended by the FBI DNA Unit II (FBI DNA Analysis Unit II, 2001). The tubes were heated to 95°C for 1 minute, followed by 36 cycles of 95°C denaturation for 10 seconds, 61°C anneal for 30 seconds, and 72°C extension for 30 seconds. This was linked to a final 20 minute extension at 70°C, followed by a 4°C hold.

The amplified products were then quantitated using the PicoGreen® dsDNA quantitation reagent. Any sample that did not contain at least 4 ng/µL of amplified product was diluted to approximately 2 X 10^{-14} g/µL. One µl of the diluted sample was re-amplified using the parameters previously described. The amplified products were then quantitated with PicoGreen®.

Post-Amplification Analytical Gels

The presence of PCR product was verified using a 3% agarose gel stained with SYBR Gold (Molecular Probes). In order to visualize both the double stranded mtDNA amplicons, and the single stranded DNA resulting from the asymmetric PCR process, a method of pre-staining the amplified product was adopted. The SYBR Gold fluorescent nucleic acid stain was mixed with a 5X Bromopheonol blue loading dye (65% Sucrose, 10 mM Tris-HCL, pH 7.5, 10 mM EDTA, 0.3% w/v Bromopheonol blue) and added to the PCR product prior to analysis. The SYBR Gold was diluted to a ratio of 1:200 in 0.5X TBE. The Bromophenol blue/SYBR Gold mix was prepared as follows; 30 µL 0.5X TBE, 20 µL 1:200 Syber Gold dilution and 50 µL 5X Bromophenol blue loading dye. A 100 bp Ladder (Invitrogen, Inc.) was loaded on the agarose

gels to verify the size of the PCR products. The size marker was prepared as follows; 150 μL 0.5X TBE, 40 μL Bromophenol blue/SYBR Gold mix, and 10 μL of the 100 bp Ladder. A 5 μL aliquot of the size marker was loaded on the gel. A 1 μL aliquot of the Bromophenol blue/SYBER Gold mix was added to 4 μL of PCR product, and loaded on a 3% agarose gel. The gels were run at 4 V/cm for three hours in 0.5X TBE buffer (0.45 mM Tris 0.45 mM Boric Acid, pH 8.3, 1 mM EDTA) (Maniatis, Fritsch et al., 1989). The PCR product was visualized using the Molecular Dynamics Fluorimager. The scanner was calibrated at 900 V using a 530 nm(D) wavelength filter. Background correction for contrast and brightness was accomplished using the Image Quant.™ 5.0 software provided with the instrument.

PicoGreen® Quantitation of mtDNA Amplicons

In order to determine the amount of PCR product, of HVI and HVII, PicoGreen® dsDNA quantitation reagent {Molecular Probes, 1996 #64} was utilized. A series of known DNA quantitation standards were prepared through the dilution of a 100 μg/mL stock of Bacteriophage Lambda DNA. Lambda DNA was diluted in TE (10mM Tris-HCL, 10 mM EDTA, pH 8.0) to a final concentration of 80 ng/μL, 40 ng/ μL, 20 ng/μL, 10 ng/μL, 5 ng/μL, 2.5 ng/μL, 1.25 ng/μL, 0.625 ng/μL, and 0 ng/μL. The assay was carried out in a 96 well flat bottomed, translucent microtiter plate (Whatman/ Polyfiltronics). To each of the wells, either a 5 μL sample of the DNA standards or a 5 μL sample of the amplified product was added. After the standards and unknown samples were added to the appropriate wells, a solution containing 94 μL of TE buffer and 1 μL of PicoGreen® reagent was added. The samples were thoroughly mixed and allowed to incubate for 5 minutes in the dark. The microtiter plate was then scanned in a Molecular Dynamics Fluorimager. The Molecular Dynamics Fluorimager was calibrated at a power setting

of 500V using a 530nm (D) wavelength filter. The densitometric analysis of the microtiter plate was performed using the ImageQuant 5.0 software package supplied with the instrument (Molecular Dynamics). The optical density (OD) values obtained for the DNA standards and the unknown samples were exported into a Microsoft Excel worksheet. The OD values from the DNA standards were used to generate a calibration curve. The calibration curve was then used to calculate the amount of amplified dsDNA product in each amplified sample.

Post-Amplification mtDNA Product Cleanup

Following amplification of the primers and unincorporated dNTPs are removed utilizing an exonuclease that fragments single stranded DNA and a phosphatase that cleaves off the phosphate group from deoxynucletoides. After PicoGreen® quantitation, the amplicons concentrated was adjusted to 10 ng/μL. A 10 μL aliquot (100 ng)of PCR product was placed into a 0.2 mL MicroAmp® tubes. To each sample either, 2 μL Shrimp Alkaline Phosphatase SAP (1 Unit/μL) (USB Corp) and 0.2 μL E.coli Exonuclease I Exo I (10 Units/μL) or 4 μL ExoSAP-IT™ (0.5 Unit/μL Shrimp Alkaline Phosphatase and 0.5Unit/μL E.coli Exonuclease I) was added. The tubes were placed in a GeneAmp® 9700 thermocycler (Applied Biosystems) and incubation at 37° C for 90 minutes followed by 80°C for 15 minutes to inactivate the enzymes. The samples were then stored at 4° C.

PCR Amplification for Sequence Specific Oligonucleotide Hybridization Assay

The Sequence Specific Oligonucleotide Hybridization assay required the input of a labeled single stranded mtDNA product. The labeled singled stranded mtDNA product was generated by asymmetric PCR, which incorporated a primer containing a Bodipy TMR

fluorescent tag. The HVI and HVII amplicons generated from the samples analyzed in the single base extension assay were used as the mtDNA template for the asymmetric PCR reactions. The HVI and HVII amplicons for each sample were diluted to a concentration of approximately 0.02 pg/μL. The primers used for the re-amplification of the diluted HVI and HVII amplicons were identical to the primers used for single base extension assay with the exception that the primers complementary to the heavy chain were fluorescently labeled with Bodipy TMR. The primers used to generate the labeled mtDNA amplicons for the sequence specific oligonucleotide hybridization assay are shown in **Table 2.**

Table 2. Amplification Primers for Sequence Specific Oligonucleotide Hybridization Assay

Hypervariable Region I

A1:	((L15997)	5'-CACCATTAGCACCCAAAGCT-3'
B2:	(H16236)	Bodipy TMR-5'-CTTTGGAGTTGCAGTTGATG-3'
A2:	(L16159)	5'-TACTTGACCACCTGTAGTAC-3'
B1:	(H16391)	Bodipy TMR-5'-GAGGATGGTGGTCAAGGGAC-3'

Hypervariable Region II

C1:	(L048)	5'-CTCACGGGAGCTCTCCATGC-3'
D2:	(H285)	Bodipy TMR-5'-GGGGTTTGGTGGAAATTTTTTTG-3'
C2:	(L172)	5'-ATTATTTATCGCACCTACGT-3'
D1:	(H408)	Bodipy TMR-5'-CTGTTAAAAGTGCATACCGCCA-3'

For the amplification of HVI, primer subsets A1-B2 and B1-A2 were used, and for amplification of HVII, primer subsets C1-D2 and D1-C2 were used. A total of four fluorescently tagged amplicons were generated per individual. The amplification mix for the sequence specific oligonucleotide hybridization assay consisted of 1X PCR Buffer (10 mM Tris-HCl pH 9.0 at 20°C, 50 mM KCl, Applied Biosystems), 160 μg/mL BSA, 1.5 mM $MgCL_2$ (Applied Biosystems), 200 μM dNTP's (Applied Biosystems), 5 units Platinum *Taq* Polymerase (Invitrogen Inc.), and 0.2 μM of the labeled primers, and 0.02 μM of the unlabeled primers, and 0.04 pg of the HVI or HVII amplicons. DNA samples were re-amplified in a 50 μL reaction using the GeneAmp® 9700 Thermal Cycler (Perkin Elmer, Applied Biosystems). The thermocycling parameters were 94°C for 4 minutes, followed by 40 cycles of 94°C denaturation for 30 seconds, 61°C anneal for 30 seconds, and 72°C extension for 30 seconds. This was linked to a 4 minute extension at 70°C, followed by a 4°C indefinite hold.

All PCR reactions were prepared in a laminar flow hood, which contained a germicidal UV light (polychromatic wavelength of 280-320 nm). The germicidal UV light on for a minimum of two hours prior to use. All micro amp tubes used for PCR were exposed to UV light (305 nm) for approximately 20 minutes prior to use. All reagents used for PCR were kept at 4°C after thawing.

Luminex LabMAP™ System

The Sequence Specific Oligonucleotide Hybridization assay developed on the Luminex LabMAP™ System can be described as a "reverse bead blot". Luminex Corp. developed a method to incorporate two spectrally distinct fluorophores into polystyrene microspheres (**Figure**

4). Using precise ratios of these fluorophores, Luminex Corp. has produced an array of 100 different microspheres.

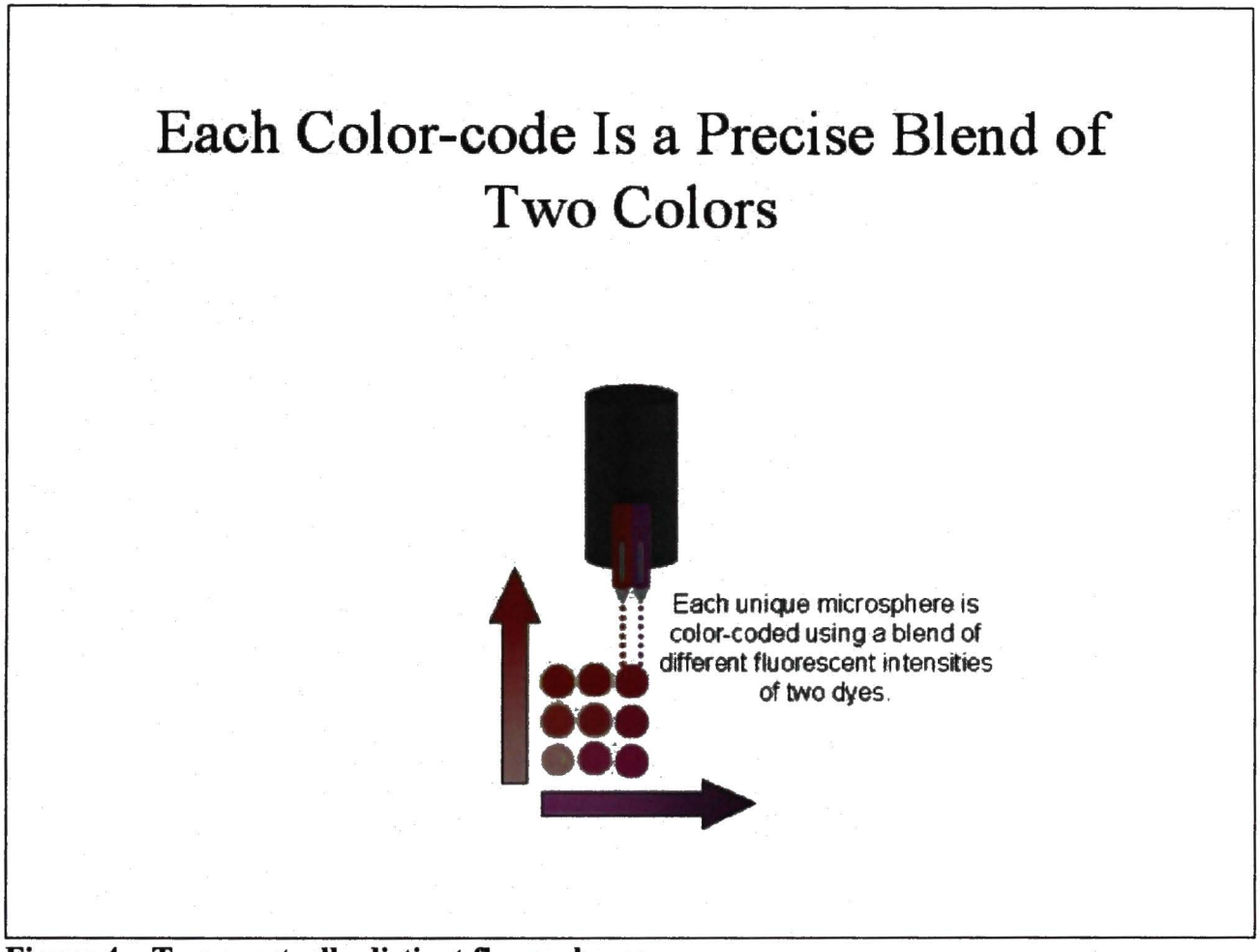

Figure 4. Two spectrally distinct fluorophores.
From http://www.luminexcorp.com/tech/QuickCourse2.htm

The fluorescent dyes contained within each microsphere provide the classification information necessary to distinguish the bead sets. As a result, the microsphere sets can be combined allowing up to 100 different capture oligonucleotides to be analyzed simultaneously in a single reaction. A capture oligonucleotide probe complementary to a specific mtDNA region is coupled to a unique microsphere set. The microspheres with the attached capture probes are then hybridized with Bodipy TMR fluorescent labeled single stranded PCR product. Hybridization takes place between a unique capture probe and its complement within a mtDNA

amplicon. The hybridization mixture is then injected into the Luminex 100™ flow cytometer. Microfluidics align the microspheres in single file where a pair of lasers illuminates the fluorophores on the inside and on the surface of the beads (**Figure 5**).

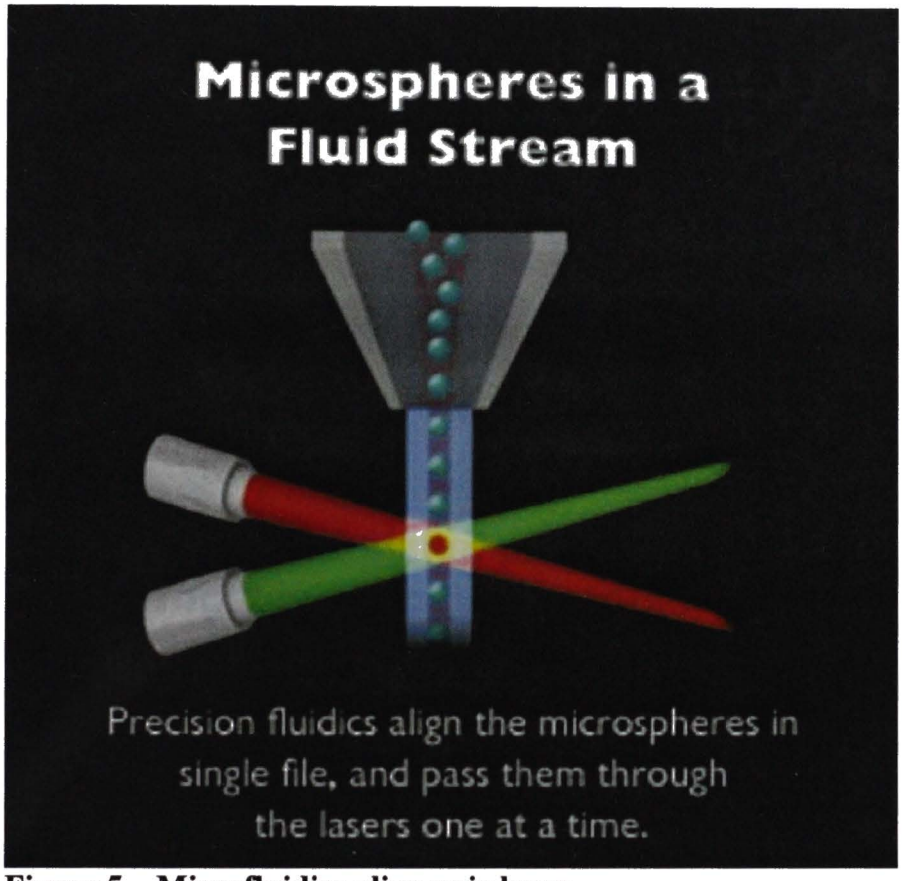

Figure 5. Microfluidics align pair laser.
From http://www.luminexcorp.com/tech/QuickCourse2.htm

The optics within the Luminex 100™ capture the color signal, and a digital signal processor classifies the microspheres based on its spectral address and quantifies the amount of labeled PCR product captured (**Figure 6**). The net mean fluorescence intensity (MFI) is reported for each microsphere set. The MFI is a measurement of the amount of labeled PCR product hybridized to the microsphere bound capture probe. Thousands of microspheres are interrogated per second resulting in an analysis system capable of analyzing up to 100 different mtDNA SNPs in a single reaction tube in a few seconds.

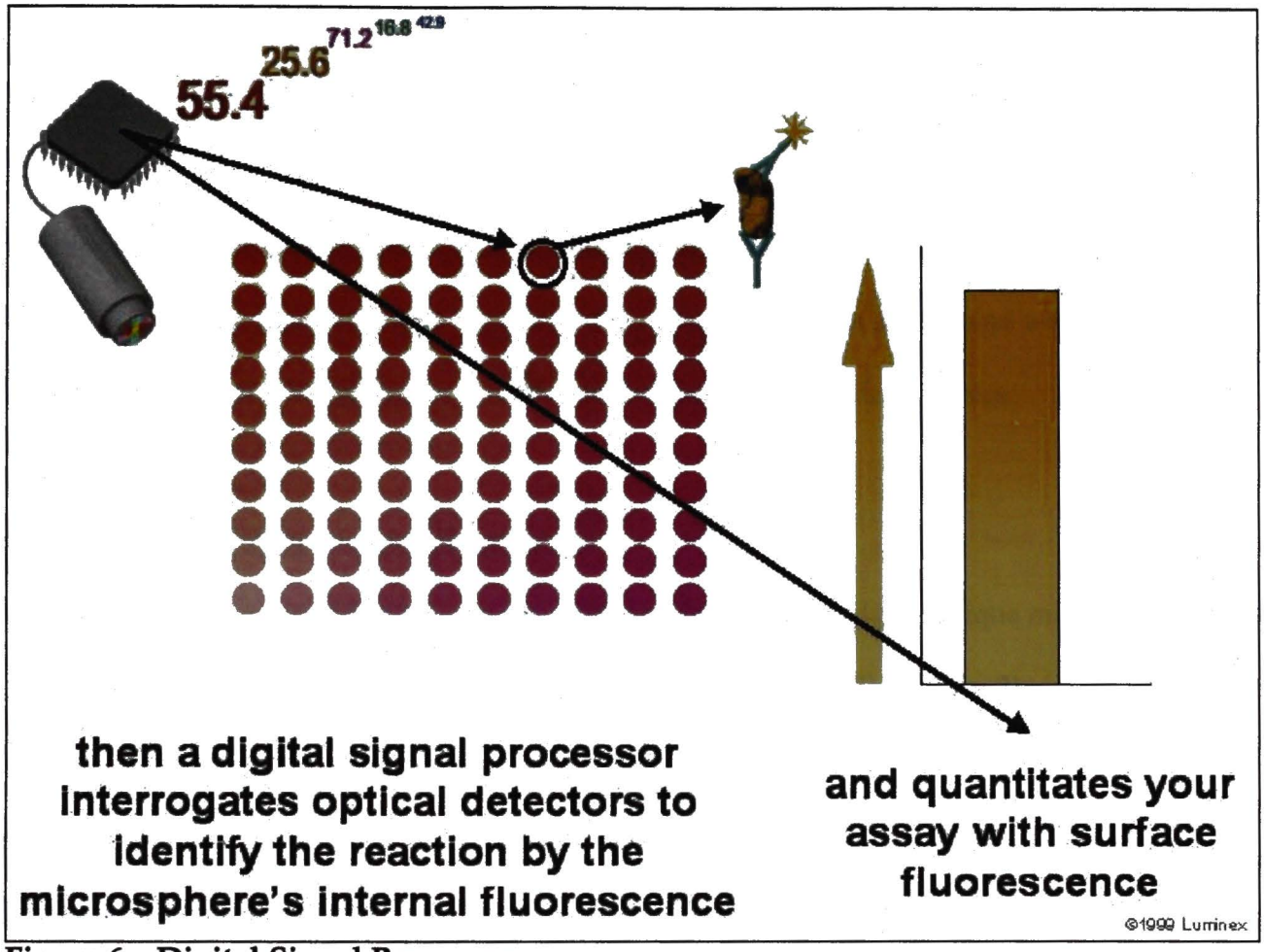

Figure 6. Digital Signal Processor.
From http://www.luminexcorp.com/tech/QuickCourse2.htm

Oligonucleotide Capture Probes

A panel of oligonucleotide capture probes was developed based on polymorphism frequencies among 2426 individuals included within an early version of the FBI's mtDNA database. The database included Caucasian, Hispanic, African American, and Asian individuals. The following criteria were used to design the oligonucleotide capture probes: 1. SNPs were identified that occurred in ten percent or more of the individuals in the database; 2. The most frequent polymorphisms within ten base pairs of the SNP position of interest were identified; 3. Hybridization probes were redesigned to compensate for the polymorphisms flanking the SNP. It was essential to compensate for these secondary polymorphisms, in order

to reduce the frequency of null alleles and to increase and the discriminatory power of the multiplex. A panel of 45 oligonucleotide capture probes (**Table 3**) was initially developed to interrogate 31 SNP positions. Since hybridization is dependent upon the entire sequence within the capture probe, each oligonucleotide actually defines a specific sequence motif rather than a unique SNP. The oligonucleotide capture probes were 20 bases in length, and were synthesized with a 5' amine Uni-Link group for coupling to the carboxylated microspheres.

Oligonucleotide Conjugation to Microspheres

Each oligonucleotide capture probe was covalently attached to a unique microsphere set using a one-step carbodiimide coupling method (Fulton et al., 1997) (**Figure 7**).
Each coupling reaction contained 1 μM of amino-substituted oligonucleotide and 1 x 10^8 microspheres/mL in 0.1 M MES (2[N-Morpholino]ethanesulfonic acid) buffer, pH 4.5. EDC (1-ethyl-3-[3dimethylaminopropyl]carbodiinide hydrochloride)was added at 10 mg/mL and the reaction was incubated for 30 minutes at room temperature in the dark. The EDC solution must be made fresh each time just prior to use. A second aliquot of EDC was added and incubated for an additional 30 minutes. The coupled microspheres were washed one time in Phosphate Buffered Saline (PBS: 150 mM $NaCl_2$, 10mM NaH_2PO_4 pH 7.4) containing Tween® 20 (0.02 % v/v), and pelleted by centrifugation at 10,000 x g for 1 minute. The microsphere pellet was washed two times in PBS containing Sodium Dodecyl Sulfate (SDS, 0.1% w/v). The microspheres were pelleted by centrifugation at 10,000 x g for 1 minute. The microsphere pellet was then resuspended in 400 μL of 0.1 M MES buffer (pH 4.5) and counted using a hemocytometer. The set of 45 coupled microspheres were combined so that each bead was present at a final concentration of 400 beads per μL. The multiplex bead set was stored in a 0.1

M MES buffer (pH 4.5) at 4°C. The coupled beads were wrapped in aluminum foil to minimize exposure to light.

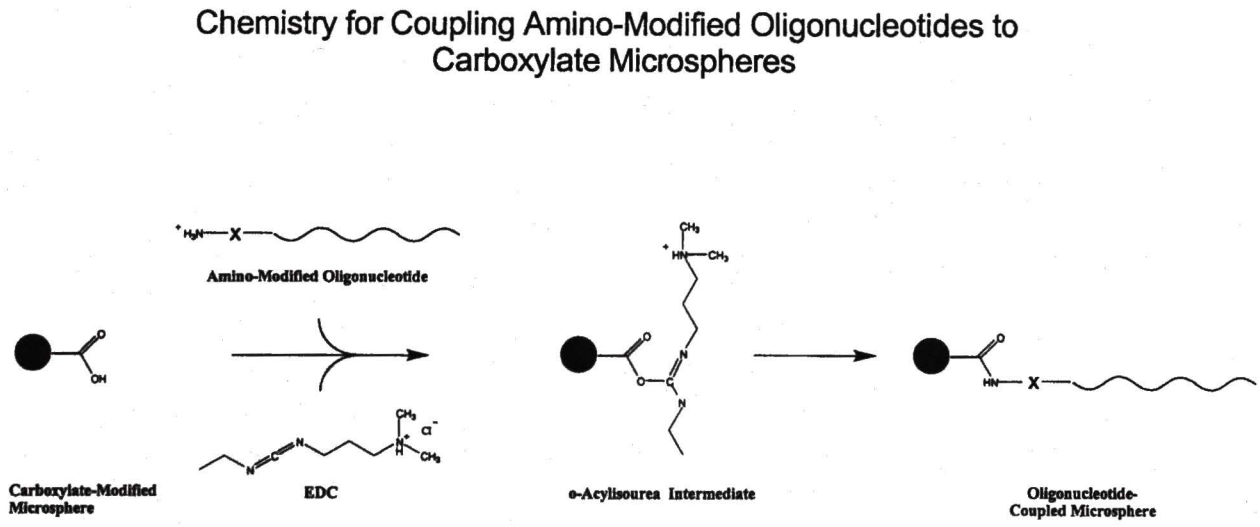

Figure 7. Schematic for the Coupling of Oligonucleotide Capture Probes to Carboxylated Microspheres. (Luminex Corporation, 2000)

Table 3. Oligonucleotide Capture Probes

BEAD NUMBER	HVI SNP MOTIF	SEQUENCE
106	Anderson	$^{+}H_3N$--X--CCCATGAAATAACAGAGAAT-3'
107	16126C	$^{+}H_3N$--X--CCCATGAAA**G**AACAGAGAAT-3'
108	16129A	$^{+}H_3N$--X--CCCATG**T**AATAACAGAGAAT-3'
109	16124C	$^{+}H_3N$--X--CCCATGAAATAA**G**AGAGAAT-3'
118	Anderson	$^{+}H_3N$--X--GATAGTTGAGGGTTGATTGC-3'
119	16223T	$^{+}H_3N$--X--GATAGTTGA**A**GGTTGATTGC-3'
120	16217C	$^{+}H_3N$--X--GATAGTTGAGGGTTG**G**TTGC-3'
121	16224C	$^{+}H_3N$--X--GATAGTTG**G**GGGTTGATTGC-3'

BEAD NUMBER	HVI SNP MOTIF	SEQUENCE
133	Anderson	$^{+}H_3N$--X--TGTACTGTTAAGGGTGGGTA-3
134	16292T,16295T	$^{+}H_3N$--X--TGTACTGTTAAGAGTAGGTA-3
135	16294T	$^{+}H_3N$--X--TGTACTGTTAAGGATGGGTA-3
136	16294T, 16296T	$^{+}H_3N$--X--TGTACTGTTAAAGATGGGTA-3
137	16294T, 16296T, 16304C	$^{+}H_3N$--X--TGTGCTGTTAAAGATGGGTA-3
138	16298C	$^{+}H_3N$--X--TGTACTGTTGAGGGTGGGTA-3
139	16304C	$^{+}H_3N$--X--TGTGCTGTTAAGGGTGGGTA-3
150	Anderson	$^{+}H_3N$--X--AAATGGCTTTATGTACTATG-3'
152	16309G	$^{+}H_3N$--X--AAATGGCTTTATGTACCATG-3'
153	16311C	$^{+}H_3N$--X--AAATGGCTTTATGTGCTATG-3'
154	16311C, 16320T	$^{+}H_3N$--X--AAATGACTTTATGTGCTATG-3'
155	16319T	$^{+}H_3N$--X--AAATGGTTTTATGTACTATG-3'
156	16320T	$^{+}H_3N$--X--AAATGACTTTATGTACTATG-3'
172	Anderson	$^{+}H_3N$--X--ATCCATGGGGACGAGAAGGG-3'
173	16362C	$^{+}H_3N$--X--ATCCATGGGGGCGAGAAGGG-3'

BEAD NUMBER	HVII SNP MOTIF	SEQUENCE
175	Anderson	$^{+}H_3N$--X--ATCGCGTGCATACCCCCCAG-3'
177	73G	$^{+}H_3N$--X--ATCGCGTGCA**C**ACCCCCCAG-3'
179	Anderson	$^{+}H_3N$--X--TCCAGCGTCTCGCAATGCTA-3'
181	93G	$^{+}H_3N$--X--TCCAGCGTC**C**CGCAATGCTA-3'
124	Anderson	$^{+}H_3N$--X--AATAATAGGATGAGGCAGGA-3'
126	146C	$^{+}H_3N$--X--AATAATAGGATG**G**GGCAGGA-3'
128	152C	$^{+}H_3N$--X--AATAAT**G**GGATGAGGCAGGA-3'
129	146C,152C	$^{+}H_3N$--X--AATAAT**G**GGATG**G**GGCAGGA-3'
130	143A	$^{+}H_3N$--X--AATAATAGGATGAGG**T**AGGA-3'
132	150T	$^{+}H_3N$--X--AATAATAG**A**ATGAGGCAGGA-3'
142	143A,146C,152C	$^{+}H_3N$--X--AATAAT**G**GGATG**G**GG**T**AGGA-3'
143	146C,150T,152C	$^{+}H_3N$--X--AATAAT**G**G**A**ATG**G**GGCAGGA-3'
144	146C,153G	$^{+}H_3N$--X--AATAA**C**AGGATG**G**GGCAGGA-3'
148	150T,152C	$^{+}H_3N$--X--AATAAT**G**G**A**ATGAGGCAGGA-3'
159	151T,152C	$^{+}H_3N$--X--AATAAT**GA**GATGAGGCAGGA-3'
192	Anderson	$^{+}H_3N$--X--ACTTTAGTAAGTATGTTCGC-3'
193	195C	$^{+}H_3N$--X--ACTTTAGTA**G**GTATGTTCGC-3'
194	199C	$^{+}H_3N$--X--ACTTT**G**GTAAGTATGTTCGC-3'
196	189G, 200G	$^{+}H_3N$--X--ACTT**C**AGTAAGTATG**C**TCGC-3'

BEAD NUMBER	HVII SNP MOTIF	SEQUENCE
200	195C, 198T	$^{+}H_3N$--X--ACTTTAATAGGTATGCTCGC-3'
163	Anderson	$^{+}H_3N$--X--GTGTGGAAAGTGGCTGTGCA-3'
164	263G	$^{+}H_3N$--X--GTGTGGAAAGCGGCTGTGCA-3'

Sequence Specific Oligonucleotide Hybridization Assay

All the reagents used for the hybridization assay were preheated to 55°C at least one hour prior to use. The Luminex 100™ instrument contains an X-Y Platform (**Figure 8**) that is designed for a 96 well microtiter plate. The X-Y Platform; 1) permits the automatic loading of samples from a 96 well plate; 2) has a built in heater to maintain the temperature of the samples within the 96 well plate. The X-Y Platform was preheated to 55°C at least one hour prior to the analysis. Hybridization reactions between the bead mix (45 sets of microspheres with the coupled oligonucleotide capture probes) and the labeled single-stranded HVI and HVII amplicons were done in a 96 well PCR Thermowell Plate (Costar). For each sample, 30.5 µL of a Blocking Solution (150 mM $NaCl_2$, 10mM NaH_2PO_4 pH 7.4, Tween® 20 (0.02 % v/v), Bovine Serum Albumin (1 mg/mL), Sodium Azide (0.05% w/v)) was added to the wells. A 5 µL aliquot of either labeled HVI or HVII amplicon was then added to the appropriate wells. As a negative hybridization control, 5 µL of TE buffer was added to the appropriate well. The tube containing the bead mix was vortexed for 2 minutes, sonicated for 2 minutes, and then vortexed again for 45 seconds. For each sample, a 2 µL aliquot of the bead mix was added to 12.5 µL of Hybridization Solution (3M Tetramethylammonium chloride, 50 mM Tris-HCl, pH 8.0, 4 mM EDTA, pH 8.0, 0.1% Sarkosyl) in a separate tube. The tube containing the bead mix and the

hybridization solution was vortexed for 45 seconds and incubated at 55°C for 5 minutes. A 14.5 µL aliquot of the then combined bead mix and hybridization solution was then added to each well. The plate was sealed and incubated at 55°C in a Heat block for 45 minutes. The 55°C Heat block had been mounted on a rocker so that the contents of the 96 well plate were slowly agitated. The 96 well PCR plate was removed from the heat block and placed on the preheated X-Y Platform of the Luminex 100™ instrument. Each sample was automatically loaded and read within the Luminex 100™ instrument. A total of 6,000 events were counted for each hybridization reaction.

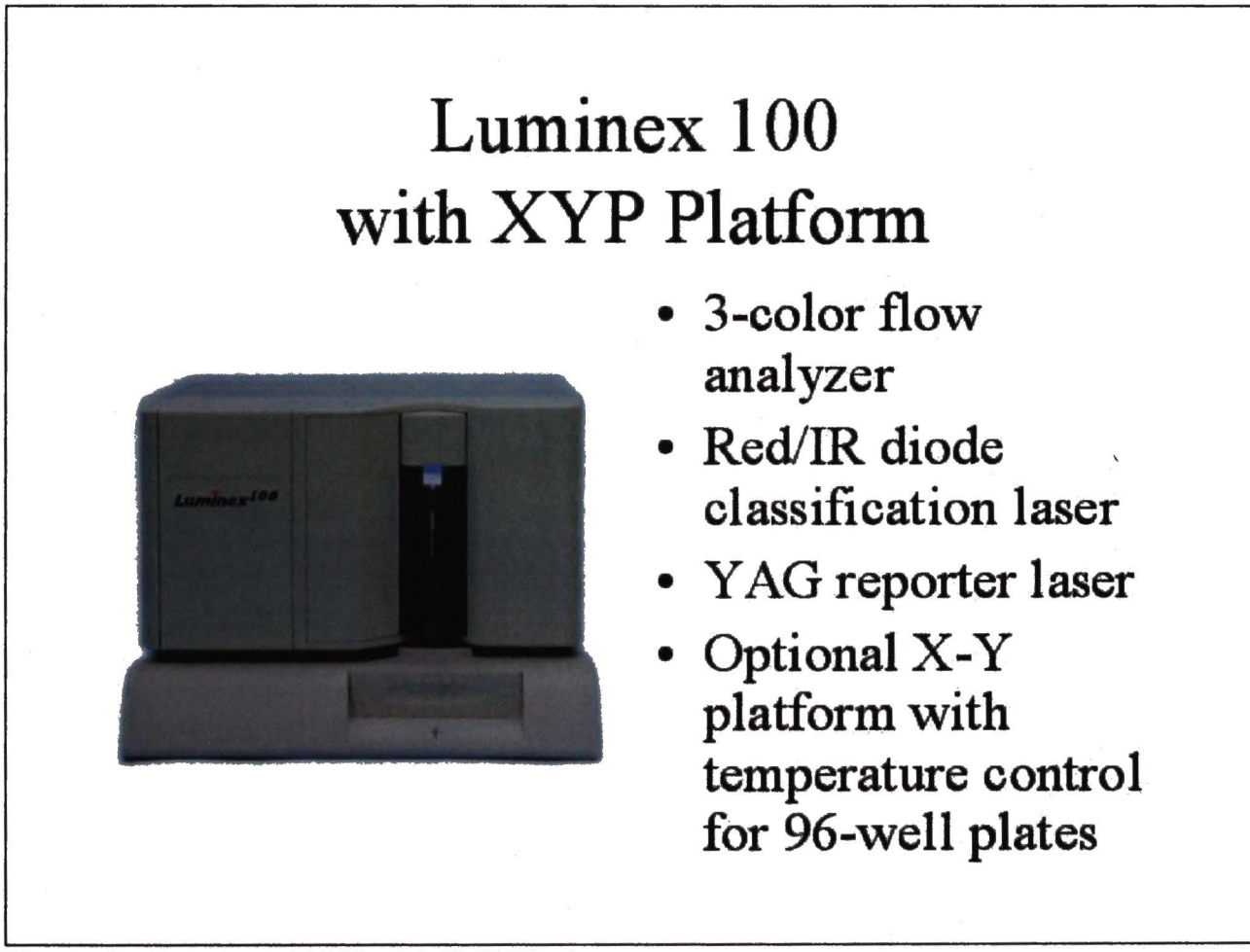

Figure 8. Luminex 100™ with X-Y Platform.
From http://www.luminexcorp.com/tech/QuickCourse2.htm

SNaPshot™ Single Base Extension Assay

The Single Base Extension (SBE) Assay was developed utilizing the ABI PRISM® SNaPshot™ Multiplex System. The assay is based on the dideoxy single base extension of an unlabeled mtDNA oligonucleotide primer (or primers). Each primer binds to a complementary template in the presence of fluorescently labeled ddNTPs and AmpliTaq® DNA Polymerase, FS. The polymerase extends the primer by one nucleotide, adding a single ddNTP to its 3′ end. The incorporation of the dideoxynucleotide prevents the further extension of the primer (**Figure 9**). The fluorescent dyes assigned to the individual ddNTPs are shown in **Table 4.**

Table 4. Fluorescent Dyes Assigned to ddNTPs in SNaPshot™ Kit

ddNTP	Dye Label	Color of Analyzed Data
A	dR6G	Green
C	dTAMRA	Black
G	dR110	Blue
T(U)	dROX	Red

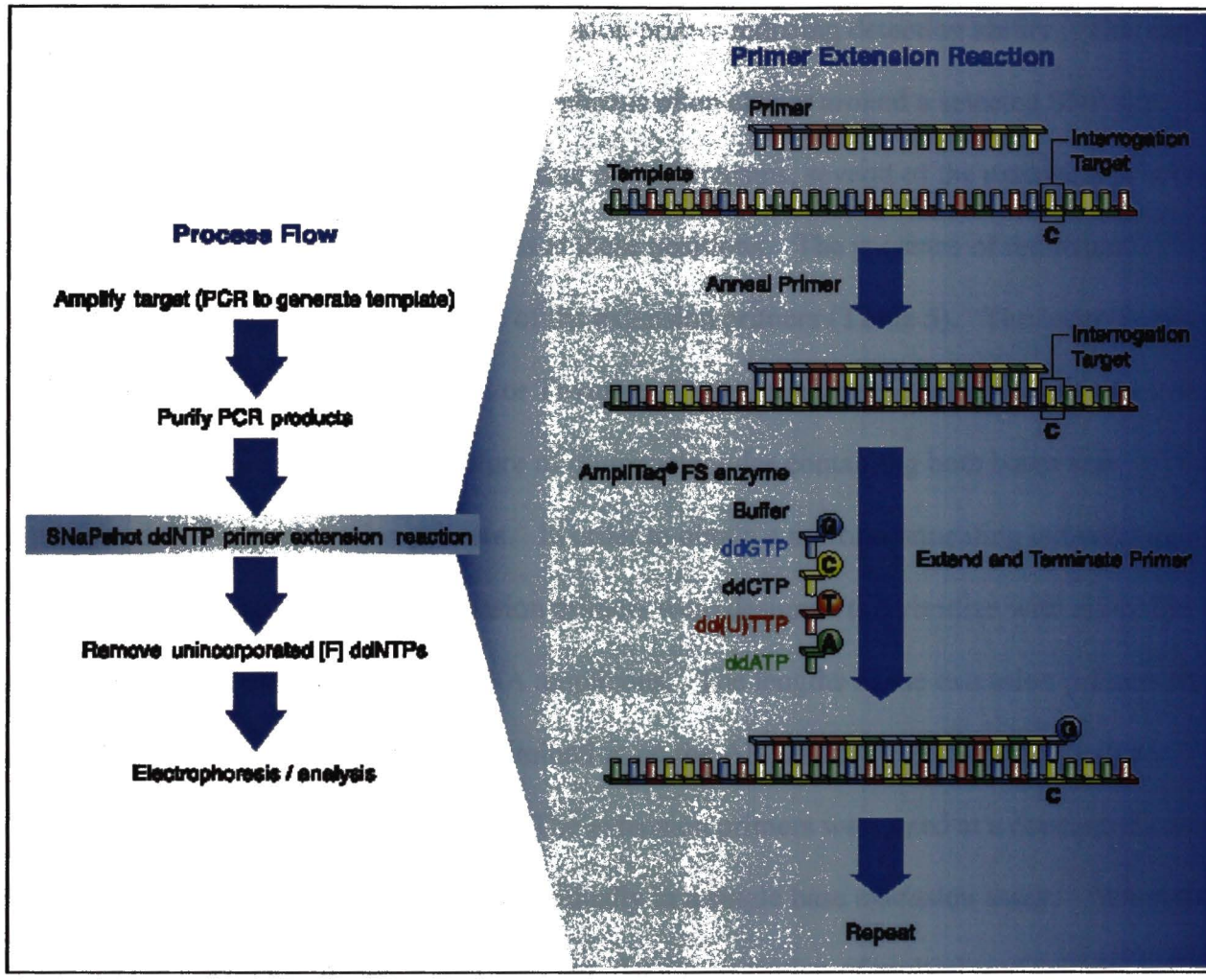

Figure 9. Outline of SNaPshot™ Primer Extension Assay.
From http://www.appliedbiosystems.com/products/productdetail.cfm?prod_id=502

SNaPshot™ Single Base Extension Primers

An initial panel of 39 oligonucleotide extension primers was developed to interrogate 34 SNP positions (**Table 5**). Two of the original extension primers were eliminated from the final testing panel. Three additional SNP positions were examined by the Single Base Extension assay as compared to those analyzed in the Sequence Specific Oligonucleotide Hybridization Assay. The extension primers were designed to minimize the presence of null typing calls, which often result from reduced or altered hybridization dynamics between the PCR amplicon and the primer. Base pair mismatches within the first six nucleotides proceeding (5') to the SNP

position can cause destabilization of an extension primer reducing detection ability. This can easily occur, since numerous other polymorphisms often cluster around a selected SNP site. To compensate for the presence of these flanking polymorphisms, several of the extension primers were designed with redundant nucleotides at these positions. The presence of redundant nucleotides can be seen in the sequence of the extension primers (**Table 5).** The letter Y designates the nucleotide Cytosine (C) or Thymine (T), and the letter R designates the nucleotide Adenine (A) or Guanine (G). A mixture of oligonucleotides containing both bases was synthesized for those extension primers. In order to obtain a uniform annealing temperature and avoid multiple priming sites, the extension primers were designed to hybridize with either the forward or reverse strands of the mtDNA amplicons. The lengths of the extension primers were also manipulated to allow the best discrimination on the ABI Prism® 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). The extension primers were used at a concentration of 0.2 μM. Each primer was first tested individually in a single base extension assay. When the primers were combined to enable multiplexing, the concentration of multiplexed primers varied due to differences in binding efficiencies. Adjustments in the final primer concentrations were done to equilibrate the relative fluorescent units (RFUs) on the ABI Prism® 310 Genetic Analyzer. Each HVI and HVII extension primer was designed to have a different electrophoretic mobility. The extension products were separated by capillary electrophoresis and visualized on the ABI Prism® 310 Genetic Analyzer.

Table 5. Primers For Single Base Extension

HVI Primers For Single Base Extension

Position	Length	Oligonucleotide Sequence (bp)	
16069.1	18	ACCACCCACGTATTGACT	Hyb. to rev
16093	34	AGTATTGACTCACCCATCAACAACCGCTATGTAT	Hyb. to rev
16124	25	CATTACTGCCAGCCACCATGAATAT	Hyb. to rev
16126	20	GCCAGCCACCATGAATATYG	Hyb. to rev
16129	23	AGGTGGTCAAGTATTTATGGTAC	Hyb. to for
16172	21	CTGTAGTACATAAAAACCCAA	Hyb. to rev
16217	23	ATGCATACAAGCAAGTACAGCAA	Hyb. to rev
16223	25	TTACAAGCAAGTACAGCAAYCAACC	Hyb. to rev
16224	21	GCAGTTCATGTGTGATAGTTG	Hyb. to for
16292	18	GATACCAACAAACCTACC	Hyb. to rev
16294	21	GGATACCAACAAACCTACCYA	Hyb. to rev
16295	19	TACCAACAAACCTACCYAY	Hyb. to rev
16304	23	GTAAATGACTTTATGTACYATGT	Hyb. to for
16309	26	TATGTACGGTAAATGRCTTTATGTRC	Hyb. to for
16311	25	CTATGTACGGTAAATGRCTTTATGT	Hyb. to for
16320	21	ATGTGCTATGTATGGTAAATG	Hyb. to for
16320.1	28	GACTGTAATGTGCTATGTACGRTRAATG	Hyb. to rev
16325	30	GGGATTTGACTGTAATGTGCTATGTACGRT	Hyb. to rev
16327	27	GGATTTGACTGTAATGTGCTATGTACG	Hyb. to for
16362	19	ACAGTCAAATCCCTTCTCG	Hyb. to rev
16362.1	33	GTACATAGCACATTACAGTCAAATCCCTTCTCG	Hyb. to rev

HVII Primers For Single Base Extension

Position	Length	Oligonucleotide Sequence (bp)	
73	29	CATGCATTTGGTATTTTCGTCTGGGGGGT	Hyb. to rev
93	16	CTCCGTCTCCAGCGTC	Hyb. to for
93.1	37	CAGATACTGCGACATAGTGTGCTCCGTCTCCAGCGTC	Hyb. to for
143	18	TATCTGTCTTTGATTCCT	Hyb. to rev
146	25	CGTAGGTGCGATAAATAATARRATG	Hyb. to for
146.1	26	ACGTAGGTGCGATAAATAATARGATG	Hyb. to for
150	27	AGTATCTGTCTTTGATTCCTRCCYCAT	Hyb. to rev
151	21	ACGTAGGTGCGATAAATAAYR	Hyb. to for
152	20	ACGTAGGTGCGATAAATAAY	Hyb. to for
153	19	ACGTAGGTGCGATAAATAA	Hyb. to for
189	22	CTACGTTCAATATTACAGGCGA	Hyb. to rev
195	26	ACGTTCAATATTACAGGCGAYCATAC	Hyb. to rev
198	29	ACGTTCAATATTACAGGCGAACATACYTA	Hyb. to rev
199	22	CATTAATTAATTAACACACTTY	Hyb. to for
200	21	CATTAATTAATTAACACACTT	Hyb. to for
247	28	ATGCTTGTAGGACATAATAATAACAATT	Hyb. to rev
263	23	AACAATTGAATGCCTGCACAGCC	Hyb. to rev

SNaPshot™ Single Base Extension protocol

The SNaPshot™ ddNTP Primer Extension Kit was used according to the manufacture's recommendations. For each sample a 5 μL aliquot of the SNaPshot Ready Reaction Premix (AmpliTaq DNA Polymerase, FS; Fluorescently labeled ddNTP's; Reaction buffer) was added to a 0.2 mL MicroAmp® tube. A 1 μL aliquot of HVI or HVII extension primer(s) (single or multiplex cocktail) (0.2 μM each primer) was added. Approximately 0.15 pmol of the HVI or HVII template DNA was added in a total of 3 μL. The total volume for the Primer Extenstion reaction was 10 μL.

The MicroAmp® tubes were placed in a GeneAmp® 9700 thermocycler. The thermocycling parameters were 96°C for 1 second, followed by 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 30 seconds. This was linked to a 4°C hold. Following the primer extension reaction, 1 μL of Shrimp Alkaline Phosphatase (SAP) (1 Unit/μL) and 1 μL of 10X PCR Buffer (PE) were added to each MicroAmp® tube. The tubes were then placed in the GeneAmp® 9700 thermocycler and incubated at 37° C for 90 minutes followed by 80°C for 15 minutes to inactivate the enzymes. The samples were then stored at 4°C.

Electrophoresis and Data Analysis on the ABI Prism® 310 Genetic Analyzer

The identification of the nucleotide sequence at the SNP site of interest was achieved in a semi-automated fashion utilizing the ABI Prism® 310 Genetic Analyzer. (**Figure 10**) The labeled primer extension products were detected by laser induced fluoresce (LIF) as the oligonucleotide passed through a detection window near the anodal end of the capillary. Fluorescent signals from 525 to 680 nm were captures by a charged-coupled device (CCD) detection system. The sensitivity of this instrument is 1000-fold greater than that of UV

detection. A matrix file was created for the SNaPshot™ extension products. This was accomplished by running the ABI dRhodamine Matrix Standard Set DS-02 that utilizes the fluorescent dyes, dR6G, dTAMRA, dR110, and dROX.

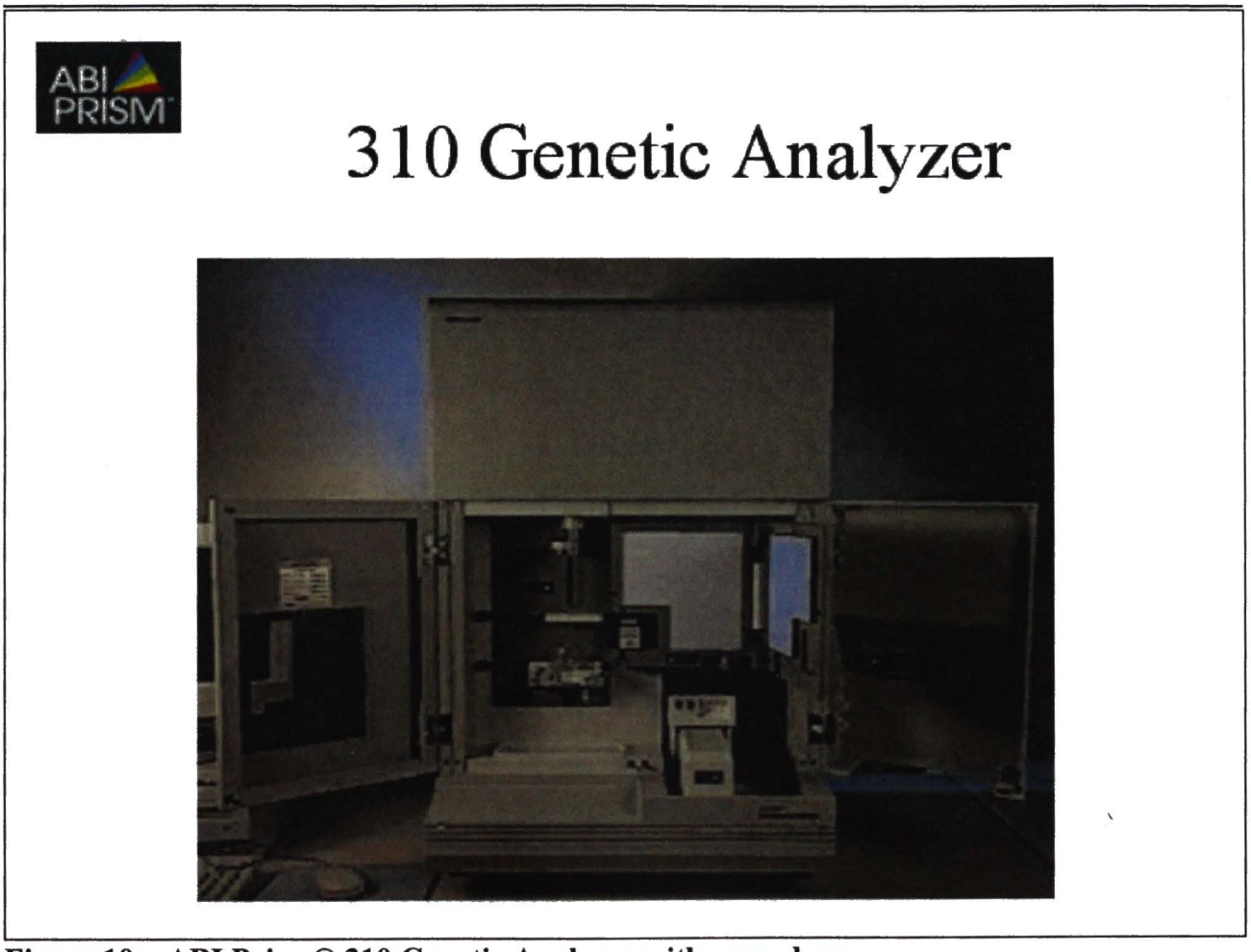

Figure 10. ABI Prism® 310 Genetic Analyzer with open doors.
From http://www.appliedbiosystems.com/products/productdetail.cfm

For each sample, 9 µL of HI-DI Formamide (Applied Biosystems) was added to a new 0.2 mL MicroAmp® tube. A 1 µL aliquot of the HVI and HVII SNaPshot™ extension products was added to each tube. The samples were denatured for 5 minutes at 95°C in a GeneAmp® 9700 thermocycler . The tubes were centrifuged at 3,000 x g for 30 seconds to bring the liquid

to the bottom of the tubes. The samples were kept at 4°C until they were loaded onto the instrument.

The run parameters for the ABI Prism® 310 Genetic Analyzer were set using the GeneScan E Run Module(GS STR POP-4 (1mL) E5 module)(Applied Biosystems, 1999). The parameters are as follows; 30 cm capillary, POP-4 Polymer, Injection time:10 seconds, Electrophoresis voltage:15kV, Collection time: 12 minutes, EP voltage: 15kV, Heat plate temperature: 60°C, Syringe pump time: 150 seconds, Preinjection EP: 120 seconds.
The spectral analysis of the primer extension products was performed using ABI GeneScan Analysis software version 3.1.2.

Cycle Sequencing

The samples used for the development of the SNP assays were cycle sequenced using the ABI Prism™ dRhodamine Terminator Cycle Sequencing Reaction Kit. The dRhodamine terminators are useful for templates with long homopolymer (>25 bases) stretches or templates with GT-rich motifs. The mtDNA PCR amplicons were obtained using the primer sets specified by the FBI DNA Unit II (FBI DNA Analysis Unit II, 2001) (**Table 1**). The amplification of HVI was divided into two separate reactions in which the primer pair A1-B2 was used for HVIA and A2-B1 was used for HVIB. The amplification of HVII was also divided into two separate reactions using the primer pair C1-D2 for HVIIA and C2-D1 for HVIIB. The dNTPs and primers were removed from the HVI and HVII amplification products using a Microcon™-100 filter unit. After filtration the amount of purified product was determined using the PicoGreen® dsDNA quantitation reagent.

Cycle sequencing was performed in a 20 μL volume which included 8.0 μL of dRhodamine Terminator Ready Reaction Mix, 1.6 μmol/L of the appropriate forward or reverse primer, and 20-35 ng of mtDNA template. The Ready Reaction Mix contains the dRhodamine dye terminators, deoxynucleoside triphosphates, AmpliTaq DNA Polymerase, FS, r*Tth* pyrophosphatase, magnesium chloride, and buffer premixed into a single tube. The dNTP mix includes dITP in place of dGTP to minimize band compressions. The primers used for the cycle sequencing reactions were identical to those used in PCR amplification. Cycle sequencing was performed in a GeneAmp® 9600 Thermal Cycler (Perkin Elmer, Applied Biosystems)as follows: a 1 minute denaturation at 96°C, followed by 25 cycles of 96°C for 15 seconds, 50°C for 5 seconds, and 60°C for 1 minute. This was linked to a 10 minute hold at 15°C, followed by a 4°C hold. Prior to electrophoresis, the unincorporated dye terminators were removed from the sequencing products using the Centri-Sep spin columns from Princeton Separation. The samples were loaded onto the center of the gel matrix without touching the matrix or the sides of the Centri-Sep column. The Centri-Sep column was centrifuged at 1,300 X g for 2 minutes. The purified sequencing fragments were dried in a Speed-Vac for approximately 25 minutes.

Each sample was resuspended in 20 μl of Template Suppression Reagent (Applied Biosystems). The samples were denatured for 4 minutes at 95°C in a GeneAmp® 9600 Thermal Cycler and placed on wet ice for a minimum of 3 minutes or until ready to use. The samples were analyzed on the ABI PRISM® 310 Genetic Analyzer using the Seq POP6 (1 mL) E run module. The sequencing reaction sample tubes were placed in an autosampler tray that holds either 48 or 96 samples. The samples were applied to a 61cm capillary by a 30 second electrokinetic injection at 2.0 kV and run for 80 minutes. The sample forms a tight band in the capillary during the injection. When the DNA fragments reach the detector window in the

capillary, a laser excites the fluorescent dye labels. Emitted fluorescence from the dyes was collected once per second by a cooled, charge-coupled device (CCD) camera and stored as digital signals. The raw data was processed by the Applied Biosystems Data Analysis software (version 2.1), the sequence electropherograms for each sample were visually evaluated to verify base calling and overall sequence quality using the ABI/Prism MT Navigator (Version 1.0.2b3). The heavy strand sequences (reverse strand), were reverse complimented prior to alignment to allow for direct comparison between the two strands sequenced.

Statistical Analysis

For each individual, the combination of sequence variants observed across all SNPs interrogated is referred to as the mtDNA haplotype. An unbiased estimate of the genetic diversity (equivalent to heterozygosity) was determined from the samples on the basis of mtDNA types, is (Equation 1)

$$h = n(1-\Sigma x^2)/ (n-1) \qquad (1)$$

Where n is the sample size and x is the estimated frequency of each mtDNA haplotype(Tajima, 1989). The probability of two randomly selected individuals from a population having identical mtDNA types was calculated using equation 2.

$$P_m = \Sigma x^2 \qquad (2)$$

The probability of a random match between two unrelated individuals (P_m) was calculated as the sum of squares of the haplotype frequencies (Tully, Sullivan et al., 1996).

RESULTS

Two mtDNA typing assays were developed and validated. The assays were based on SNP detection within HVI and HVII of the control region. The formats were 1) a competitive hybridization-based SSO method using color coded microspheres and the Luminex 100™ detection platform and 2) primer extension with subsequent electrophoretic separation on the ABI Prism® 310 Genetic Analyzer. These two systems offer simplified approaches to the currently used, more laborious Sanger sequencing method for mtDNA typing. While the discrimination power (DP) is not as high as that obtained by direct sequencing, the DP of the SNP assays is sufficiently informative for a screening method. In addition, the presence of heteroplasmy could be readily detected at interrogated sites. The assays used were rapid, cost effective, and applicable to high throughput automatable processing of multiple mtDNA samples. Lastly, multiplex assays enable typing of multiple SNP sites simultaneously across a large portion of the genome; whereas sequencing the same size region to obtain informative sites would be challenging.

Amplifications for Competitive SSO Hybridization

Amplification strategies should be considered when amplifying human mtDNA for forensic applications. The strategy selected is generally determined by the amount and/or integrity of the sample to be typed. Generally, the amplification of smaller fragments is a robust approach, because forensics samples are often somewhat degraded (unpublished results).

However, larger amplicons may be obtained where the amount of template DNA is not limited or degraded as with reference samples. For the SSO hybridization assay developed for analysis by the Luminex 100™ detection platform, small sized fragments were desired. Thus for mtDNA amplification of HVI, overlapping labeled Bodipy primer subsets A1-B2 and B1-A2 were used. For HVII, overlapping Bodipy TMR labeled primer subsets C1-D2 and D1-C2 were used. A total of four mtDNA amplicons were generated per individual sample.

The SSO hybridization assay requires that the labeled amplicons anneal with their complementary oligonucleotide coupled microspheres. The labeled amplicons therefore must be single stranded to hybridize with their complementary capture oligonucleotide coupled microsphere. However, the complementary sequences of the amplified mtDNA compete with the capture oligonucleotide probe during hybridization. Such competition reduces the amount of target mtDNA available for the capture oligonucleotide hybridization. By increasing the concentration of the target strand relative to its complement, the overall competition between the complementary strand and the oligonucleotide capture probe is greatly reduced. Excess single stranded labeled amplicons were obtained through asymmetric amplification of the mtDNA template. This was accomplished by decreasing the concentrations of the unlabeled primer. During PCR, the unlabeled primer becomes limiting and is consumed prior to the end of the amplification thus, only the labeled primer will be available in the latter cycles. The result is an excess of the labeled target single stranded mtDNA amplicons. **Figure 11** is an example of the results obtained using asymmetric amplification . The gel consists of a 100 bp ladder on each end with N1-N4 identifying the negative controls and A-D identifying four different individuals. All amplification negative controls consisted of all reagents included in the amplification premix minus the template DNA. Amplicon 1 was generated using primer sets A1-B2; amplicon 2 was

generated with primer sets B1-A2; amplicon 3 was generated with primer sets C1-D2; and amplicon 4 was generated using primer sets D1-C2. Two bands are seen per amplification. The upper band is the single stranded product, where as the lower band is double stranded 280 bp amplicon.

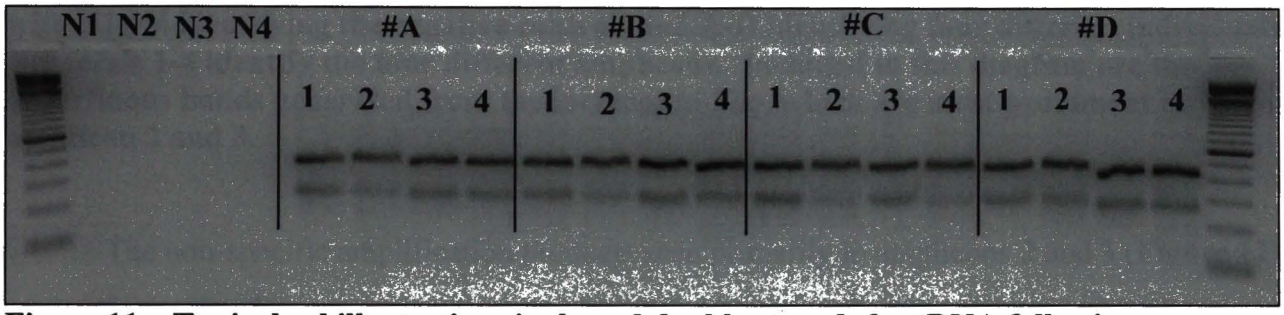

Figure 11. Typical gel illustrating single and double stranded mtDNA following asymmetric amplification. The gel consists of a 100 bp ladder on each end with N1-N4 identifying the negative controls and A-D identifying four different individuals. Numerals 1-4 identify the four different amplicons.

Initially, asymmetric amplification was performed using blood stains on FTA™ Paper. Amplification of the mtDNA template from this material was poor with sporadic results (data not shown). In an attempt to improve the yield and consistency of the amplification process, experiments were performed in which the amount of *Taq* polymerase was increased sequentially 1 U to 5 U *Taq* in 1 unit intervals. The results were evaluated using analytical gels (data not shown). At 5 U of *Taq* per reaction, the amount of single stranded amplified product obtained from FTA™ Paper punches was more consistent and significantly increase. However the increase in *Taq* resulted in a decrease in the stringency of the reaction due to the non-specific binding of the labeled primers **(Figure 12).**

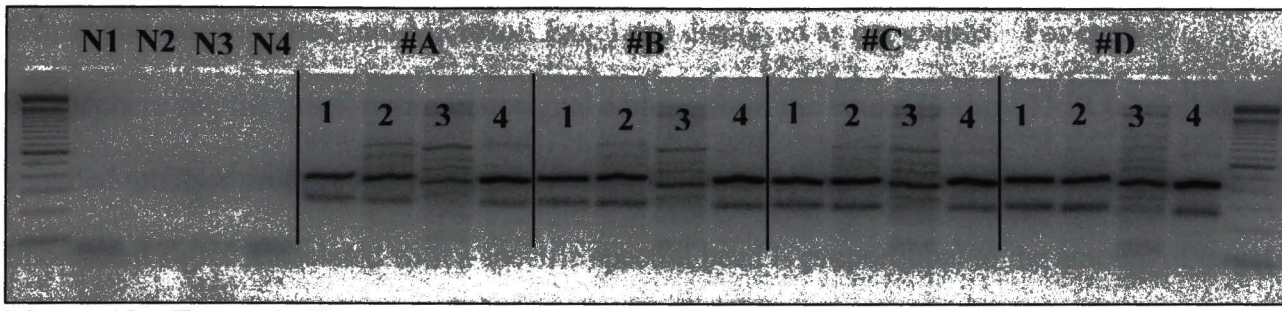

Figure 12. Example illustrating single and double stranded mtDNA following asymmetric amplification with 5 U *Taq* per reaction. The gel consists of a 100 bp ladder on each end with N1-N4 identifying the negative controls and A-D identifying four different individuals. Numerals 1-4 identify the four different amplicons. Included in this diagram are the superfluous bands generated from the increase in *Taq*, which are most prominent in amplicon 2 and 3.

The non-specific amplification products seen primarily in amplicons 2 and 3 (**Figure 12**) resulted in a decrease in the yield of single stranded amplicons. To reduce non-specific primer binding, the stringency of the amplification reaction was modified in a series of experiments that sequentially increasing the annealing from 55°C to 63°C in 1° increments (data not shown). At 61°C, the presence of non-specific products was reduced. If the annealing temperature was increased above 61° C, the yield of all amplicons was notably reduced.

To evaluate the specificity of the SSO Hybridization Assay and to establish the initial assay parameters, three separate bead set pairs were used as test controls [(163-263A, 164-263G); (172-16362T, 173-16362C); (179-93A, 181-93G)]. The bead set pairs were selected because their oligonucleotide capture probe sequences differ by only a single base at the SNP site. In preliminary studies these bead sets appeared to work effectively. Individuals in which their HVI and HVII control regions had been previously sequenced using the Sanger method were amplified and the PCR products hybridized with the oligonucleotide capture probes.

Following hybridization, the beads (unbound and with duplexes) were analyzed with the Luminex 100™ flow cytometer. The analysis time on the instrument was approximately 56 seconds per sample. The results were recorded in an ASCII comma delimited spreadsheet

format that was imported into Microsoft Excel and displayed as bar graphs. For each microsphere set, the green fluorescent signal coming from the bound Bodipy TMR labeled signal stranded mtDNA product was reported as Mean Fluorescent Intensities (MFI). The amount of fluorescent signal associated with a specific spectral bead is proportional to the amount of target mtDNA sequences that have hybridized to the capture oligonucleotide probes (minus background correction). Examples of data obtained from these bead sets are shown in **Figure 13.**

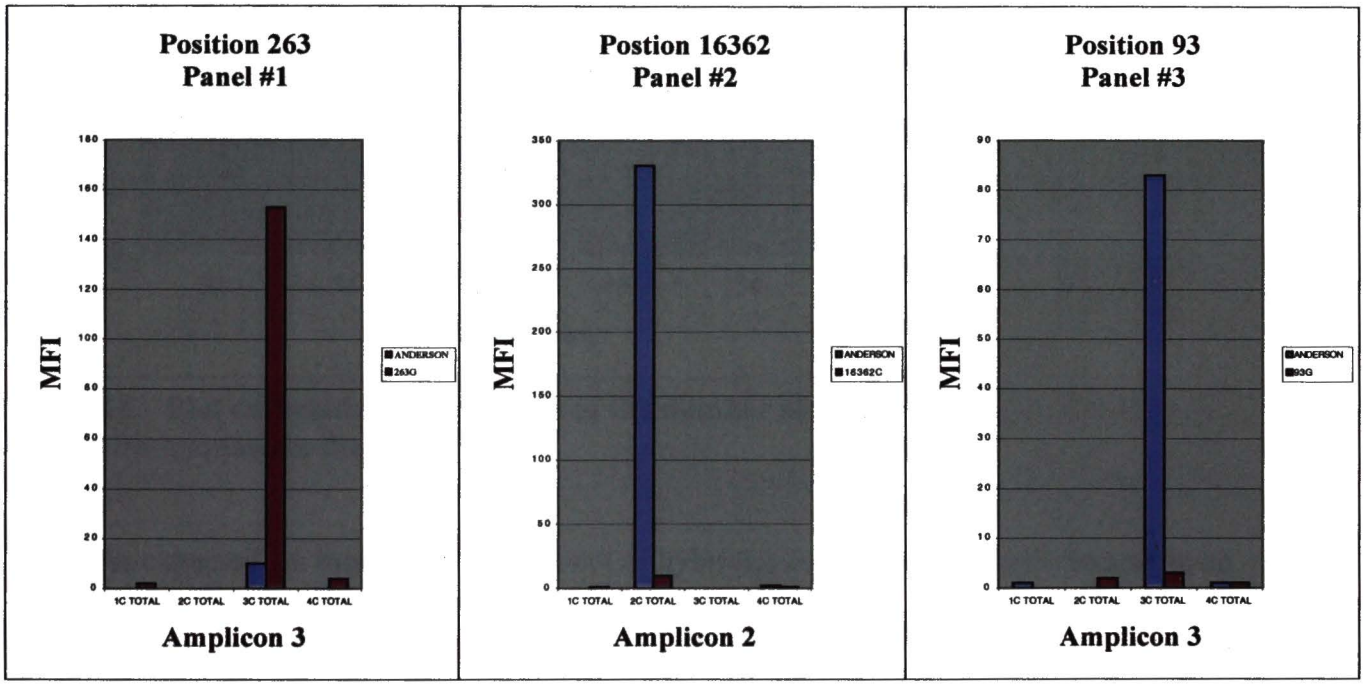

Figure 13. Example of data generated by the Luminex 100™ flow cytometer from bead sets [(163-263A, 164-263G); (172-16362T, 173-16362C); (179-93A, 181-93G)]

Interpretation of the data is by comparison of the observed peaks. If the assay is performing reliably, only one predominate signal should be observed. The highest signal should come from the capture oligonucleotide microsphere that successfully hybridized to its complementary target mtDNA sequence **(Figure 13).**

To increase the amount of single stranded mtDNA available for hybridization, the number of amplification cycles was increased from 30 to 50 in increments of 2 cycles. The same three test bead sets were used for this evaluation. The MFI values from all assays were averaged and plotted versus cycle number for each bead (**Figure 14**).

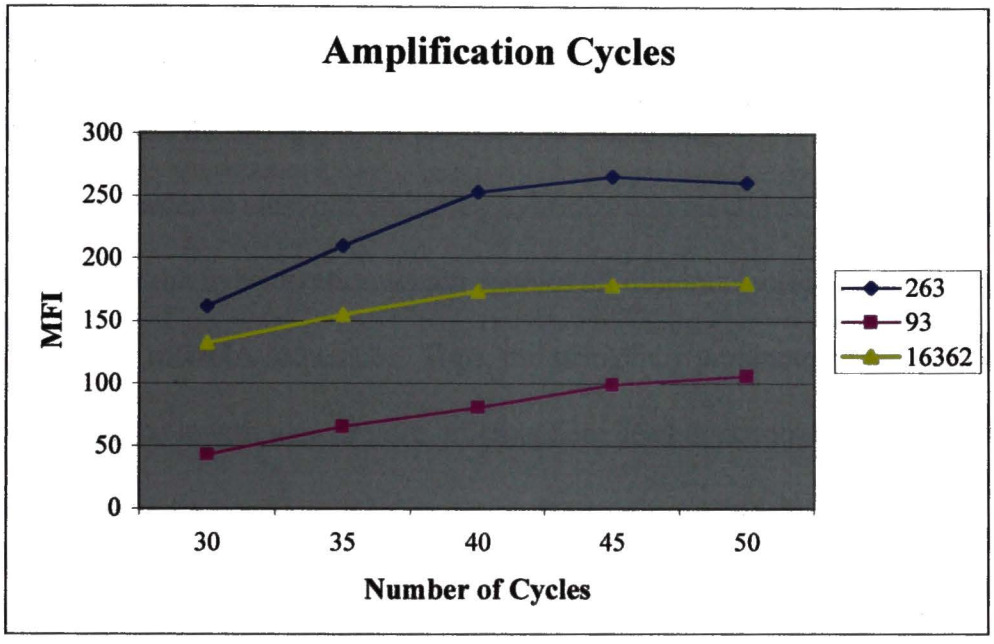

Figure 14. Plot correlating the increase in the number of amplification cycles with increase in the MFI value.

As expected, an increase in the amount of hybridization signal is directly related to an increasing number of amplification cycles. However, at 45 cycles or greater, the signal leveled off and began to decrease in some test panels. The loss of signal was due to an increase in non-specific amplification products (observed by analytical agarose gels; data not shown) that consume reagents that otherwise would be incorporated into the desired amplicon. Thus, 40 cycles were chosen for asymmetric PCR in order to generate the required labeled single stranded mtDNA product.

SSO Hybridization

Hybridization assays to detect sequence specific variations (or SNPs) is a technique that has been widely used in research, medial diagnostics, and forensics DNA analysis. The success of these reactions is dependent upon conditions that promote hybridization of oligonucleotide probes with their specific complements amid a myriad of unrelated sequences. The ability for complementary strands to successfully hybridize is dependent on the individual nucleotide sequence and the stringency of the reaction conditions.

In order to carryout multiplex hybridization reactions, a single reaction condition needs to be used for the hybridization which enables all oligonucleotide probes to hybridize in parallel to their target mtDNA sequence. Thus, the stringency parameters that include the hybridization buffer, probe length temperature, microsphere bead concentration and incubation time were evaluated.

To reduce the effects of base composition on duplex association and disassociation rates, tetramethylammonium chloride (TMAC) was added to the hybridization buffer (Wood et al., 1985; Jacobs et al., 1988). The stringency of the hybridization reaction was then controlled strictly as a function of the length of the oligonucleotide probes and the temperature. All hybridization experiments were performed with Hybridization Solution (Invitrogen) containing TMAC at 2.25 Mol/L.

Thirty one oligonucleotide probes each twenty bp in length were used in one multiplex hybridization. Capture oligonucleotides of 20 bp were used since they display a high degree of hybridization specificity (Wallace et al., 1979) with short hybridization times. The parameters for hybridization were investigated using the same three bead sets previously described; (163-263A, 164-263G); (172-16362T, 173-16362C);(179-93A, 181-93G).

Hybridization reactions were performed for 45 minutes each at different temperatures ranging from 50°C to 60°C with increments of 1°C. Data from the different hybridization temperatures were collected using the Luminex 100™ detection platform for all three bead sets. Averages of the MFI values from all individuals analyzed were plotted (**Figure 15**).

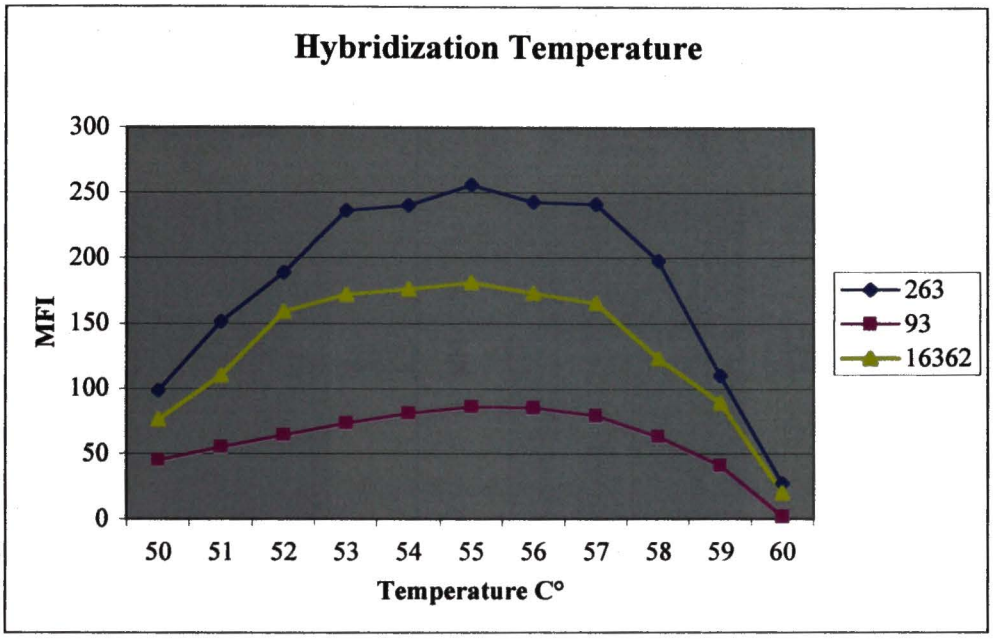

Figure 15. Plot illustrating the different hybridization temperatures.

A bell shaped curve was observed. A 55° C hybridization temperature resulted in the highest MFI value, and was therefore used throughout the rest of the validation study. At hybridization temperatures less than 55°C, the signal was reduced and non-specific binding of the alternative capture probe was observed (**Figure 16**). At the temperatures above 55°C, little to no hybridization was observed. At higher annealing temperature the duplex formation was destabilized.

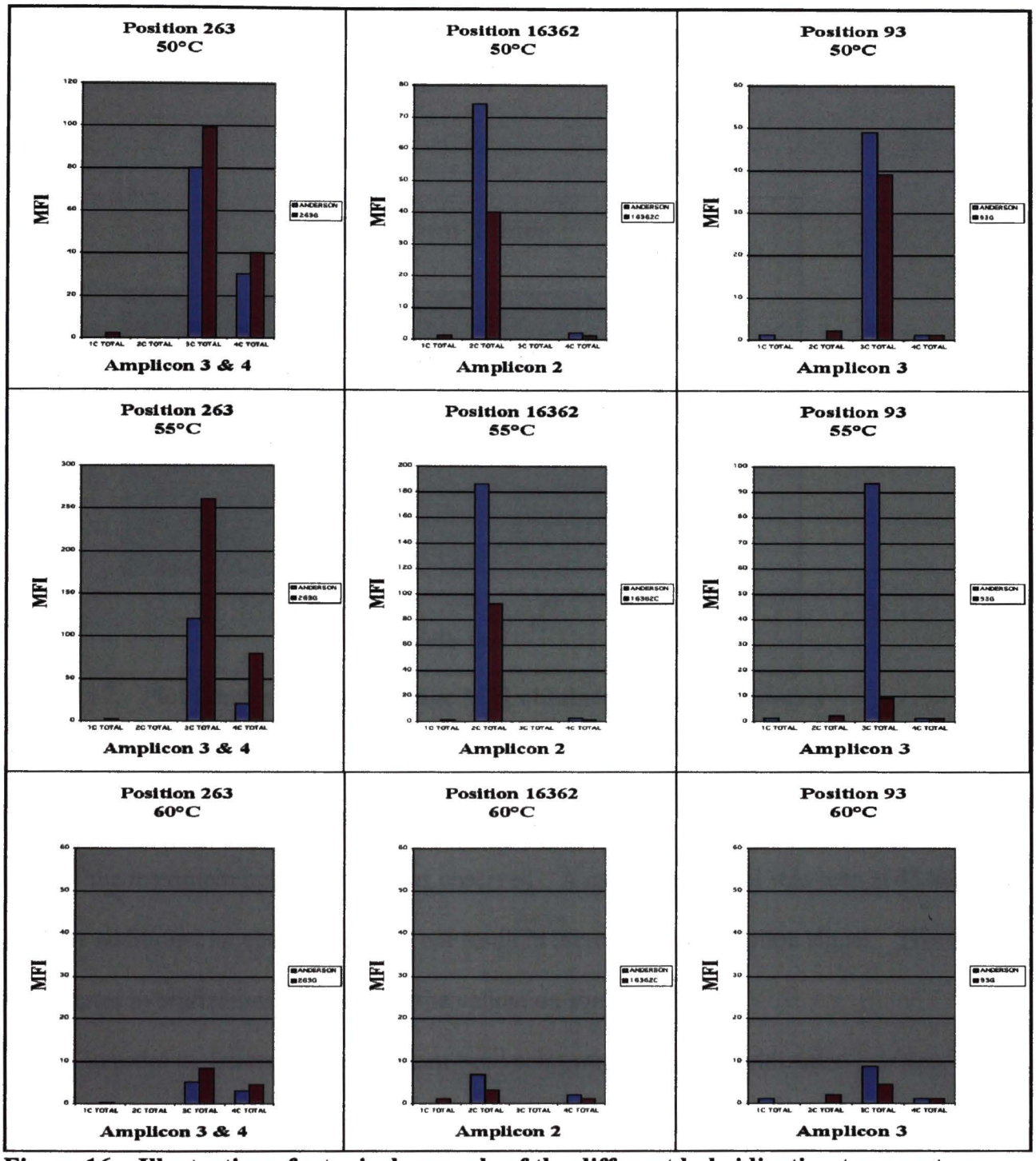

Figure 16. Illustration of a typical example of the different hybridization temperatures and the bar graphs generated from the different bead sets.

To determine the optimal hybridization period, experiments were performed that sequentially increase the incubation period. Starting at 5 minutes the hybridization periods were

increase five minute increments up to 90 minutes. The values obtained from all experiments were averaged and plotted **(Figure 17)**.

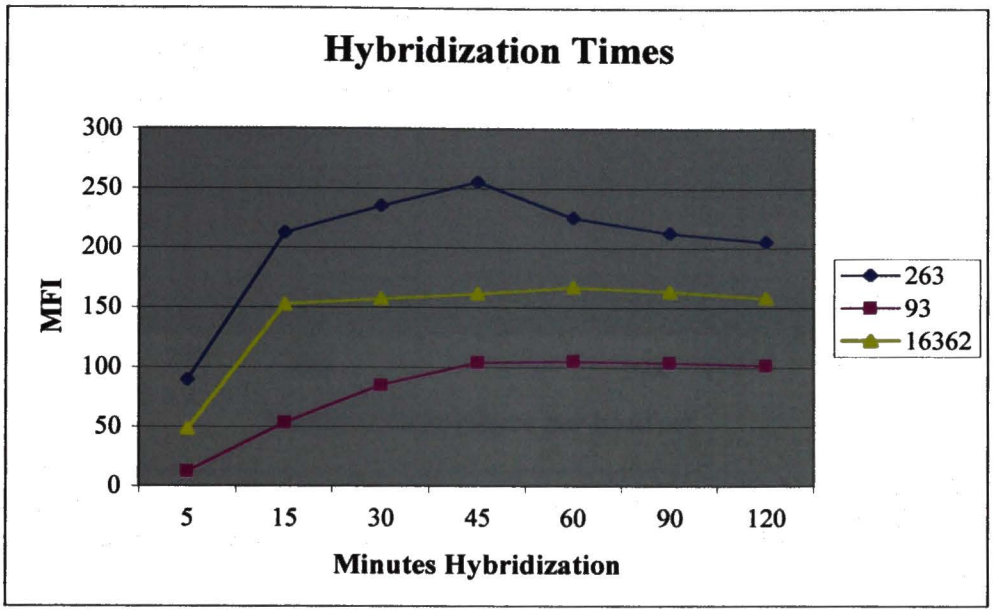

Figure 17. Plot illustrating the average MFI obtained from experiments using an increasing hybridization period.

For the first 5 minutes little or no hybridization occurred. At 30 minutes approximately 85% of the maximum hybridization was observed. A maximum signal was seen at 45 minutes and any additional incubation time did not result in increased hybridization signal. Therefore a 45 minutes hybridization was used for the validation studies.

In order to determine the optimum bead concentration in a hybridization reaction, the number of input beads was varied. The bead mix was synthesized at a concentration of 400 of each microsphere set per μL. The range of input beads was varied from 800 to 1600 beads per spectral set. At 1600 beads per reaction, the hybridization signal was increased **(Figure 18)**.

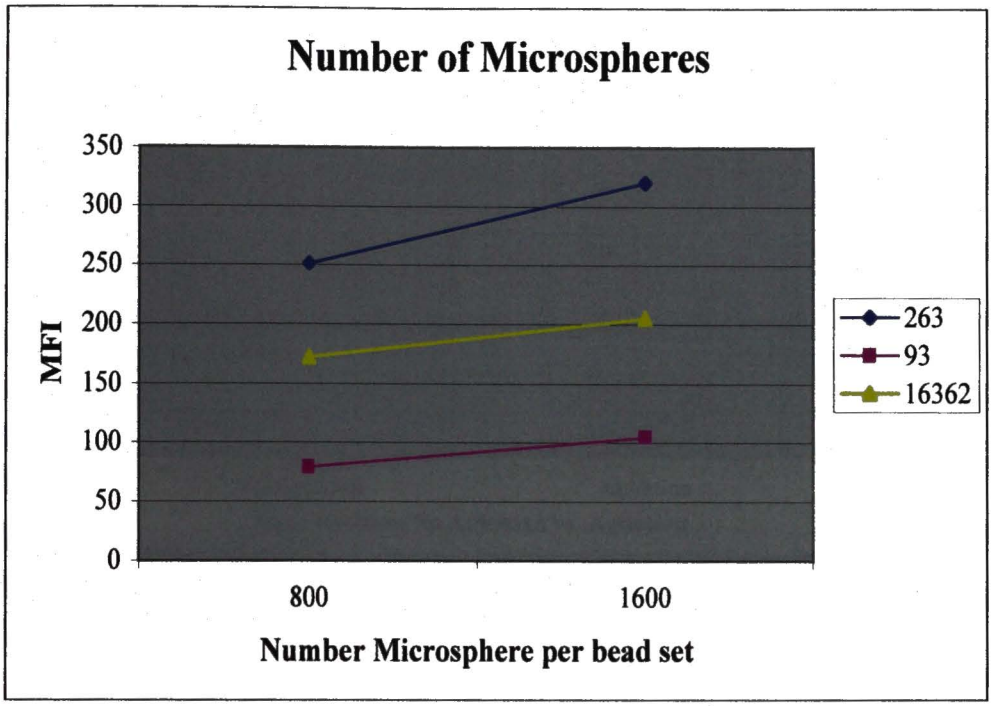

Figure 18. Illustration of the average MFI value obtained using different numbers of microspheres for each hybridization.

However, because of cost, the number of beads was set to 800 per assay. Although the MFI was reduced, acceptable signals were obtained. Thus, 2 μL of beads (400 of each bead per μL) reaction was used in further testing.

The beads have a tendency to clump and form aggregates which reduces the surface area available for hybridization. Aggregates may not always be detected in the flow cytometer and can affect the accurate counting of the microspheres. In order to ensure maximum signal and minimize aggregation, the beads were vortexed and sonicated prior to hybridization. In addition the tubes containing the hybridization reaction were agitated during the 45-minute incubation. This was possible by mounting the hybridization heat block on a shaker. Experiments showed that agitation increased the MFI values by approximately 10% **(Figure 19)**.

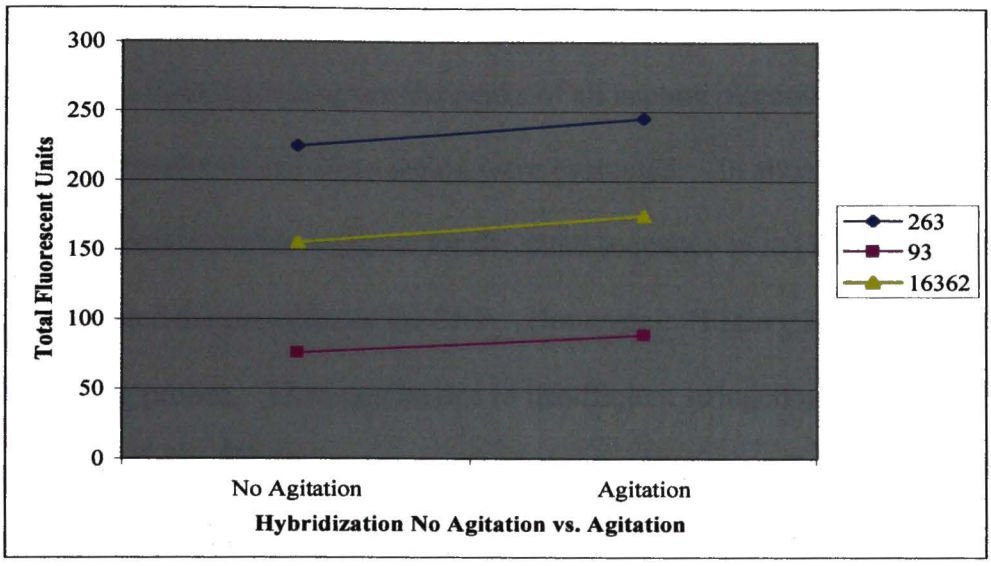

Figure 19. Plot illustrating the average values obtained with or without Agitation during hybridization.

Haplotype determination by Competitive Hybridization-based SSO

The ability to accurately and rapidly determine the mtDNA haplotype of an individual was assessed using the general SSO conditions for the bead assay. The multi-analyte DNA hybridization system was designed with 45 unique oligonucleotide sequence probes to interrogate 31 SNPs. The SNPs of interest reside in ten defined regions of the mtDNA control region. For HV1 these are IA, IB, ICa, ICb, ID; and For HVII these are IIAa, IIAb, IIB, IIC, IID. While the intent is to design a capture probe that recognizes only a specific site, a hybridization assay actually interrogates a sequence motif. The sequence motif includes all bases within the target region. This complicates the assay when SNPs are clustered, such as in HVI and HVII. Thus, each region or motif was interrogated with a capture oligonucleotide microsphere set designed to hybridize only if the region was either complementary to RCRS or exhibited a specific substitution at the SNP of interest. In regions where multiple SNPs reside in close proximity to each other, capture oligonucleotide microsphere sets were designed account for the most common mtDNA variations in that region. The MFI values for each bead set was

plotted as a histograph between beads set within a defined region. In order to determine a haplotype, from the histogram the peaks of all capture oligonucleotide microsphere competing for hybridization in the same region were evaluated. In situations where only two capture oligonucleotide probes compete for the same sequence as in (**Figures 24-26 and 29**) the highest bar indicated the presence of the SNP. However in (**Figure 24**) there are signals from both competing probes. This may be due to insufficient stringency or heteroplasmy. In regions were multiple clustered SNPs reside, multiple bars may appear, and the mtDNA haplotype is determined by evaluating all the peaks (**Figures 20, 21, 22, 23, 27 and 28**). A major part of the interpretation was defining a minimum MFI value. This was needed to avoid identifying SNPs resulting from non-specific hybridization. Therefore, a minimum specific hybridization threshold was set as a percentage of the signal level obtained from the bead with the highest MFI value. Several of the target regions overlap between the four amplicons. The minimum MFI values varied not only between the 10 regions but also between the four amplicons. Thus, values for the minimum threshold were obtained for each region and used in subsequent experiments. An example of the interpretation of sample data in Region IA through IID are illustrated (**Figures 16 through 25**).

Region IA

An example of the data obtained from an interrogation of amplicon 1 region IA, SNP positions 16126, 16129, 16124 are illustrated in **Figure 20**. The capture oligonucleotide probe with the highest MFI value is bead #6, which identifies the sequence motif as 16126T, 16129G, and 16124T (i.e. the same as the rCRS). The non-specific hybridization signal from the competing beads was <19% of the signal generated from bead #6. The more than fourfold difference in hybridization signal between the bead with the highest MFI and the other

competing beads suggests a minimum threshold for setting background noise as a percentage of the highest MFI. For region IA data obtained from multiple samples established the setting for background noise as less than 20% of the highest MFI. Any data below the established threshold was considered noise; any data above that threshold was considered a successful hybridization that would be subject to interpretation.

IA	rCRS	16126C	16129A	16124C
1C AMP.	95	18	11	3
2C AMP.	0	1	0	1
3C AMP.	0	0	0	0
4C AMP.	0	1	1	1
Bead	6	7	8	9

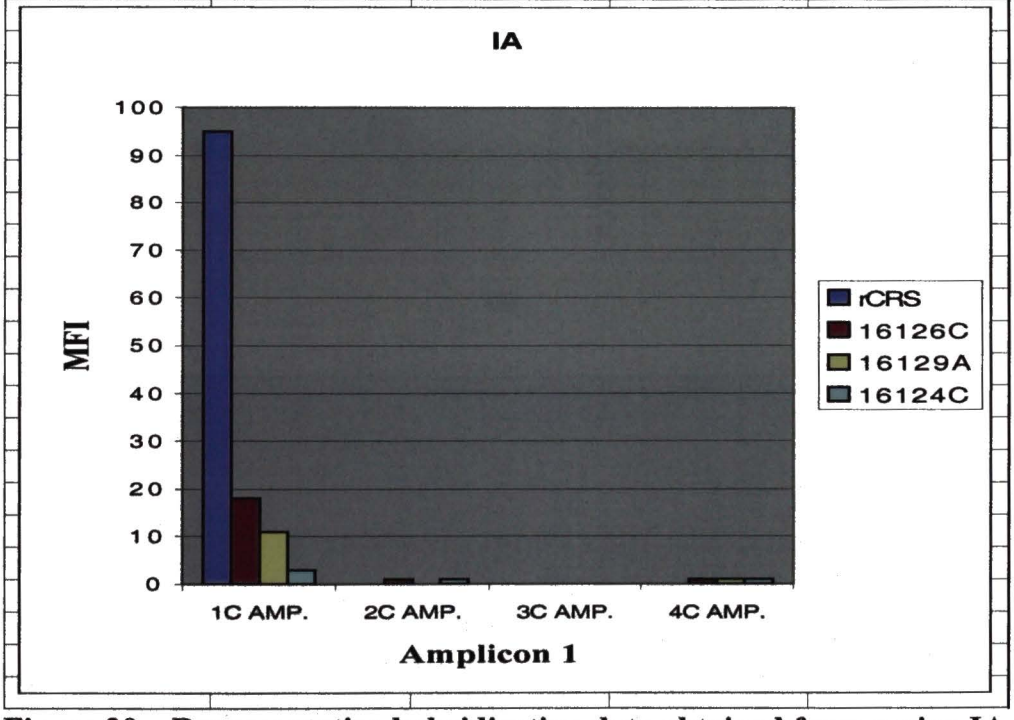

Figure 20. Representative hybridization data obtained from region IA.

Region IB

Region IB is included in amplicons 1 and 2 due to sequence overlap as illustrated in **Figure 21.** The SNP positions interrogated with this bead set are 16223, 16217 and 16224.

The capture oligonucleotide microsphere that hybridized successfully based on MFI height is bead #19, which characterizes the SNP site 16223 as 16223T (i.e. not the same as the rCRS). The MFI values for amplicons 1 and 2 are 278 and 208 MFI respectively. The other two sites are typed as 16217T and 16224T (i.e. the same as the rCRS). The background noise level of the competing beads is <26% and <28% for bead #19 for both amplicons. For region IB data obtained from multiple samples established a background threshold of 30% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

IB	rCRS	16223T	16217C	16224C
1C AMP.	70	272	45	52
2C AMP.	60	208	51	56
3C AMP.	1	2	3	3
4C AMP.	2	0	1	1
Bead	18	19	20	21

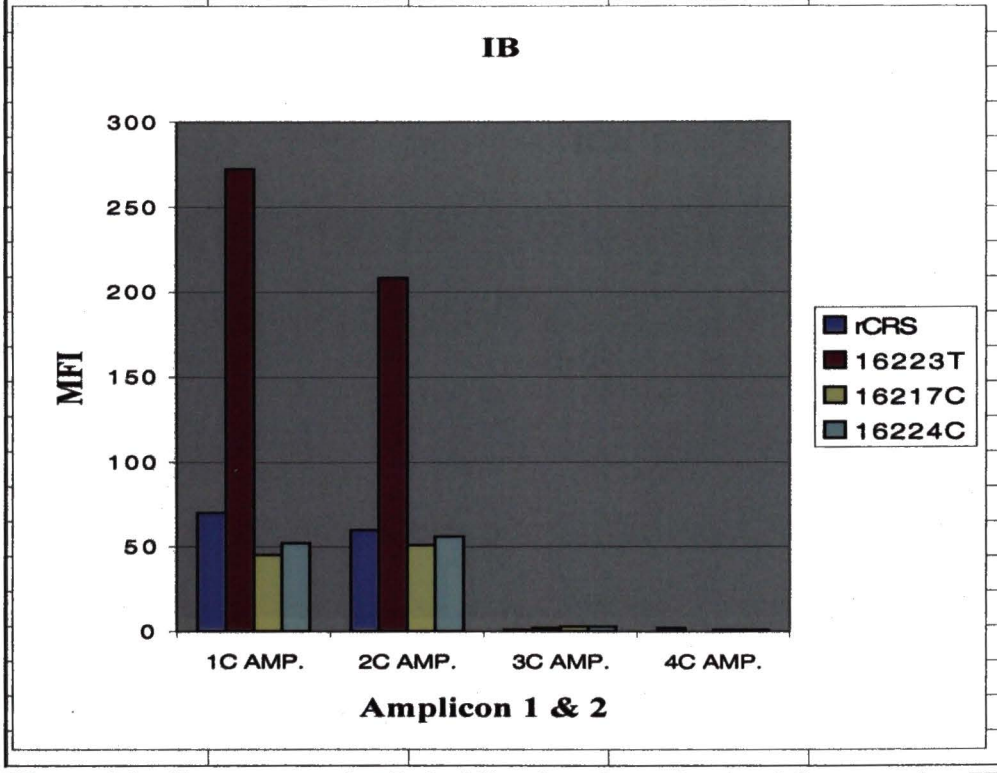

Figure 21. Representative hybridization data obtained from region IB.

Region ICa

Region ICa is included in amplicon 2 **(Figure 22)** and contains SNP positions 16292, 16294, 16296,16298 and 16304. With a MFI value of 123, bead #38 gave the highest hybridization signal. This data characterizes the SNP site 16298 as 16298C (i.e. not the same as the rCRS). The other four sites are typed as 16292C, 16294C, 16296C, and 16304T (i.e. the same as the rCRS). The background signal from the competing beads is <27% of signal obtained from bead #38. For region ICa the data obtained from multiple samples established a background threshold of 30% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

ICa	rCR	16292T 16295T	16294T	16294T 16296T	16294T 16296T 16304C	16298C	16304C
1C AMP.	5	2	1	1	5	29	22
2C AMP.	33	4	18	5	3	123	23
3C AMP.	2	1	2	0	0	3	3
4C AMP	1	1	0	1	0	1	8
Bead	33	34	35	36	37	38	39

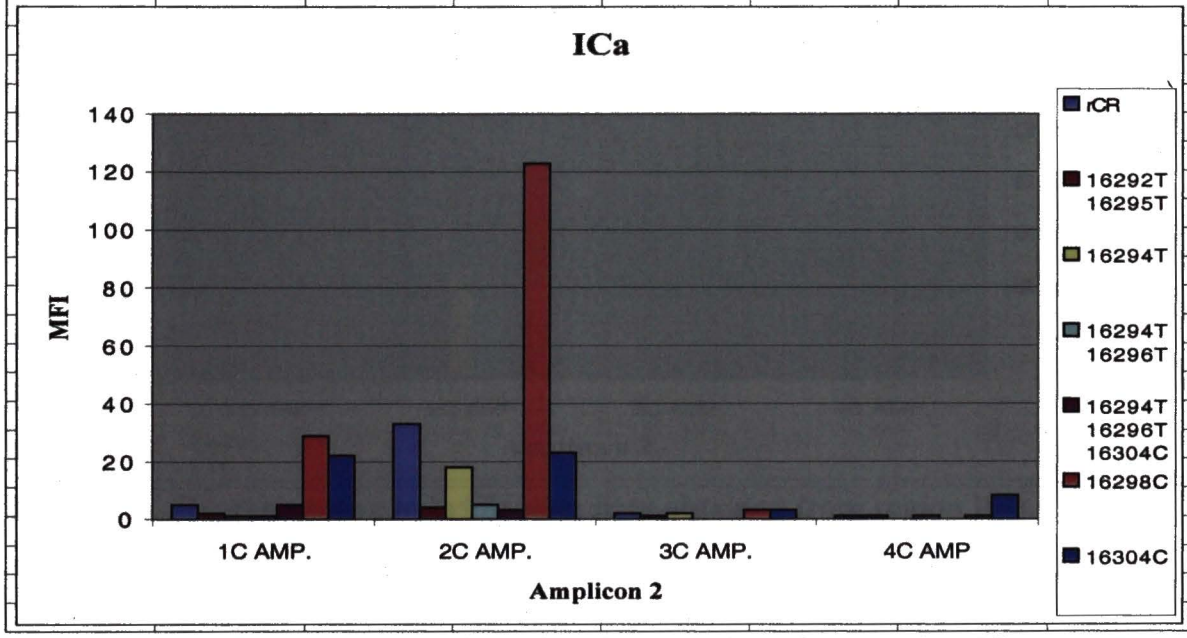

Figure 22. Representative hybridization data obtained from region ICa.

Region ICb

Region ICb is included in amplicon 2 **(Figure 23)** and contains SNP positions 16309, 16311, 16319 and 16320. With a MFI value of 172, bead #50 gave the highest hybridization signal. This data characterizes the four SNP sites typed as 16309A, 16311T, 16319G and 16320C (i.e. the same as the rCRS). The background signal from the competing beads is <30% of signal obtained from bead #50. For region ICb the data obtained from multiple samples established a background threshold of 30% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

ICb	ANDERSON	16309G	16311C	16311C 16320T	16319A	16320T
1C AMP.	4	1	4	3	9	1
2C AMP.	172	38	52	37	50	46
3C AMP.	0	0	1	1	1	0
4C AMP.	0	0	0	2	0	1
Bead	50	52	53	54	55	56

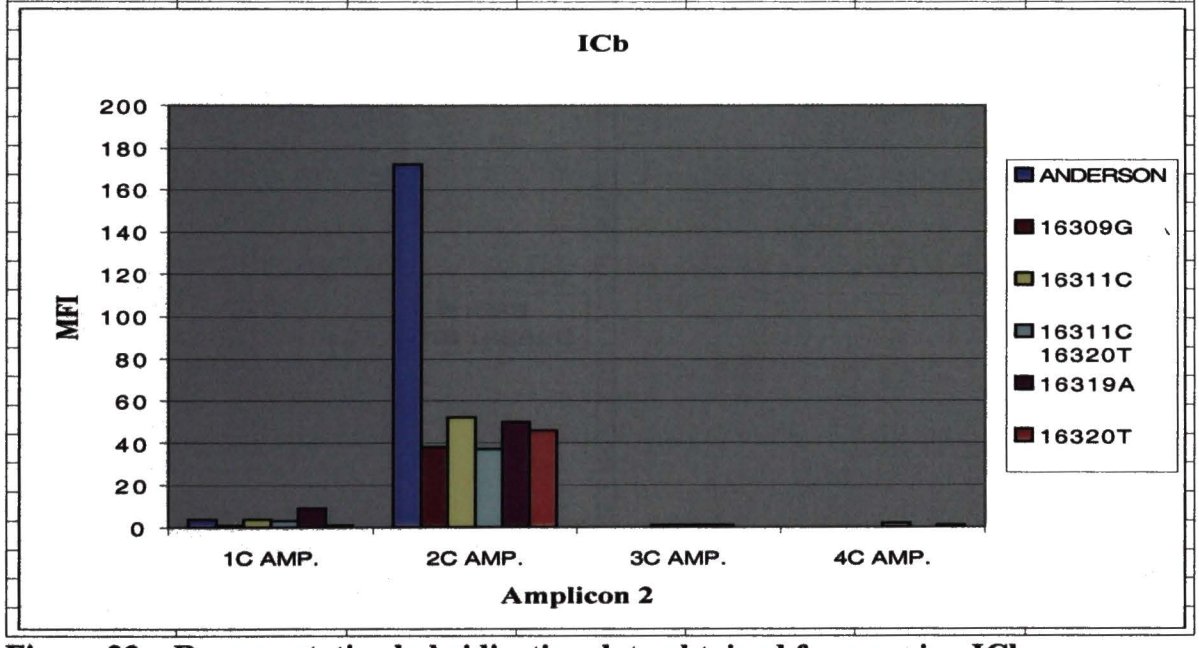

Figure 23. Representative hybridization data obtained from region ICb.

Region ID

Region ID is included in amplicon 2 **(Figure 24)** and contains SNP positions 16362. With a MFI value of 330, bead #72 gave the highest hybridization signal. This data characterizes the SNP site as 16362T (i.e. the same as the rCRS). The background signal from the competing beads is <51% of signal obtained from bead #72. For region ID the data obtained from multiple samples established a background threshold of 50% of the MFI value relative to the strongest hybridizing bead. Although the background noise level of the competing bead was higher than the threshold, (51%) the competing bead was not evaluated.

ID	rCRS	16362C
1C AMP.	0	1
2C AMP.	330	168
3C AMP.	0	0
4C AMP.	2	1
Bead	72	73

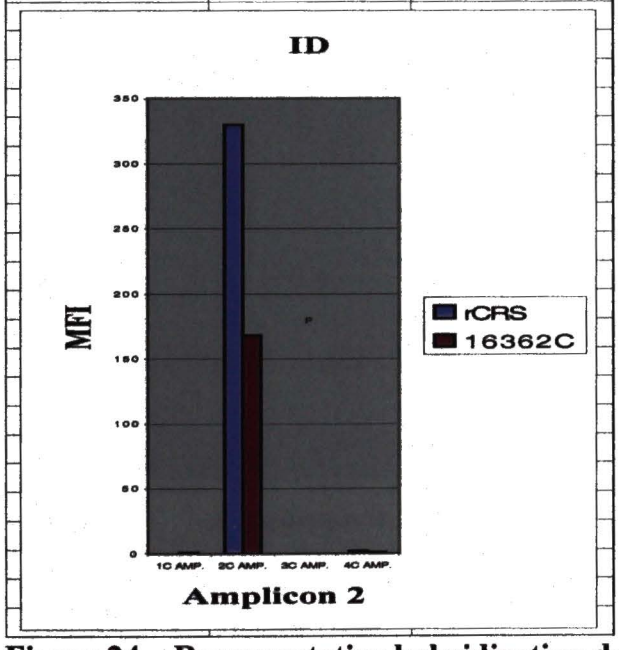

Figure 24. Representative hybridization data obtained from region ID.

Region IIAa

Region IIAa is included in amplicon 3 **(Figure 25)** and contains SNP positions 73. With a MFI value of 100, bead #75 gave the highest hybridization signal. This data characterizes the SNP site as 73G (i.e. not the same as the rCRS). The background signal from the competing beads is <10% of signal obtained from bead #72. For region IIAa the data obtained from multiple samples established a background threshold of 20% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

IIAa	rCRS	73G
1C AMP.	2	1
2C AMP.	0	2
3C AMP.	10	100
4C AMP.	0	2
Bead	75	77

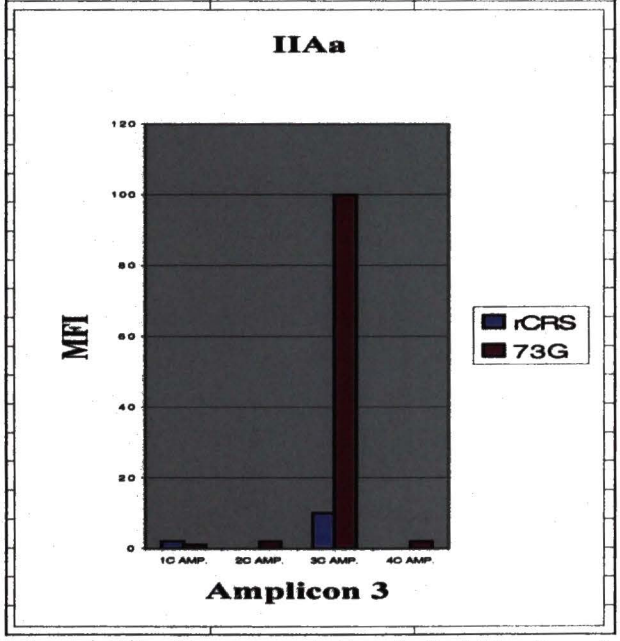

Figure 25. Representative hybridization data obtained from region IIAa

Region IIAb

Region IIAb is included in amplicon 3 **(Figure 26)** and contains SNP positions 93. With a MFI value of 49, bead #79 gave the highest hybridization signal. This data characterizes the SNP site as 93A (i.e. the same as the rCRS). The background signal from the competing beads is <10% of signal obtained from bead #79. For region IIAb the data obtained from multiple samples established a background threshold of 20% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

IIAb	rCRS	93G
1C AMP.	1	0
2C AMP.	0	2
3C AMP.	49	3
4C AMP.	1	1
Bead	79	81

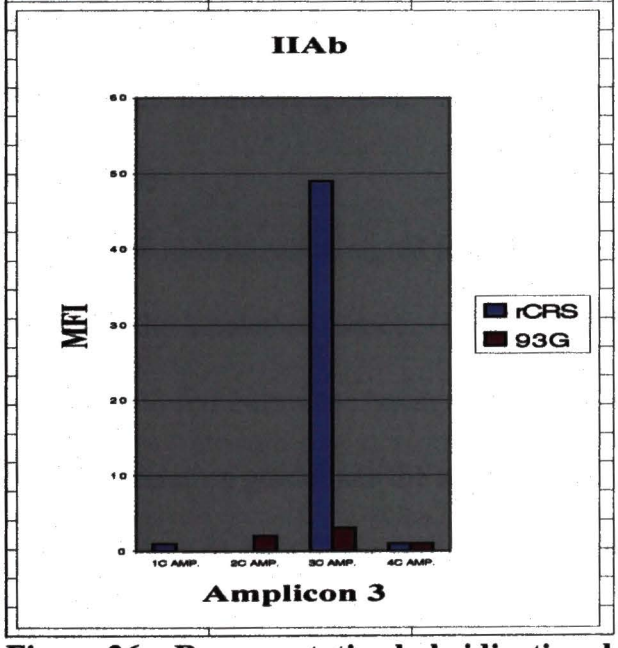

Figure 26. Representative hybridization data obtained from region IIAb.

Region IIB

Region IIB is included in amplicon 3 **(Figure 27)** and contains SNP positions 143, 146, 150, 151, 152, and 153. The bead sets that gave significant hybridization signal are #26, #28, #29, #43 and #48, with MFI values of 75, 78, 172, 42 and 41 respectively. The hybridization data would indicate that the sequence within this motif is 146C and 152C (i.e. not the same as rCRS). The other four sites within this motif typed as 143G, 150C, 151C, and 153A (i.e. the same as rCRS). In regions where multiple SNPs are located in close proximity to one another, capture oligonucleotide microspheres were designed to interrogate these different motifs. This is the case with bead #29 which, contains oligonucleotide sequences complementary to SNPs 146C and 152C. In this sample the background hybridization signal from the competing beads is <25% of signal obtained from bead #29. For region IIB data obtained from multiple experiments established the background threshold as less than 30% of the MFI value relative to the strongest hybridizing bead. Bead #43 and #48 gave a significant signal of 42 and 41 MFI respectively.

The hybridization signal from these beads results from partial binding of the capture probes to the labeled sequence. In addition, the MFI values obtained for this sample with bead #43 and #48 fell below the background threshold established at 73% for region IIB. Based upon our interpretation guidelines the sequence motif from these two beads would be discounted. Bead #43 contains the bases 146, 150T and 152C within its capture probe, while bead #48 contains the bases 150T and 152C within its capture probes. The presence of 150T in bead #43 and #48 did not sufficiently destabilize the labeled single stranded mtDNA target from these beads.

IIB	rCRS	146C	152C	146C,152C	143A	150T	143A,146C 152C	146C,150T 152C	146C,153G	150T,152C	151T,152C
1C AMP.	1	0	2	1	3	0	6	2	0	1	1
2C AMP.	0	0	1	3	5	0	7	1	1	0	0
3C AMP.	4	75	58	172	1	4	2	42	5	41	17
4C AMP.	0	0	0	0	0	0	0	0	1	0	0
Bead	24	26	28	29	30	32	42	43	44	48	59

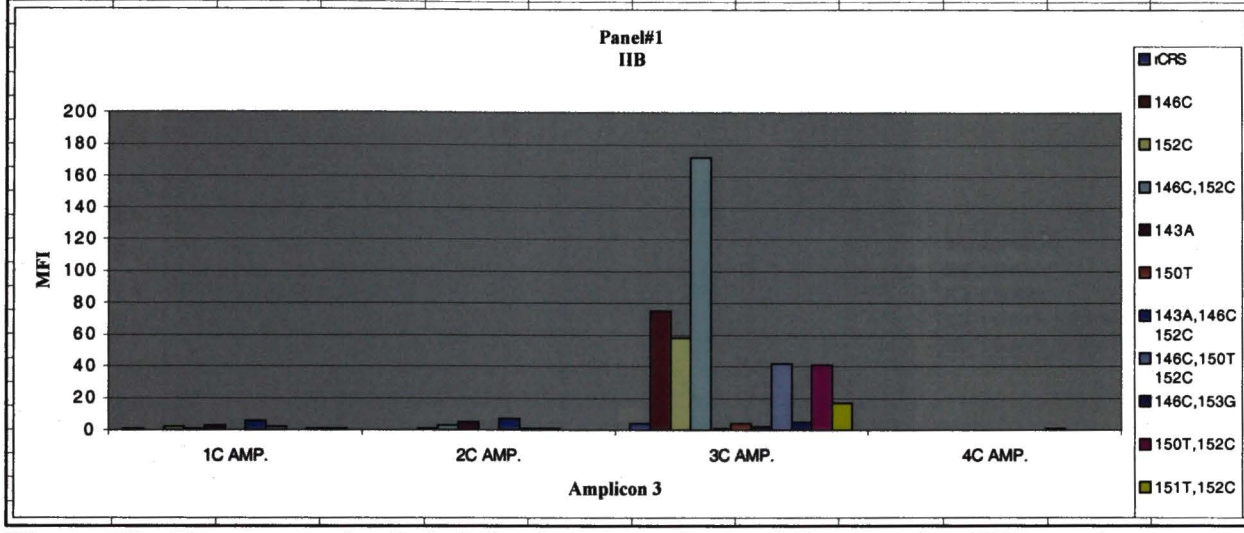

Figure 27. Representative hybridization data obtained from region IIB.

Region IIC

Region IIC is included in amplicon 3 and 4 **(Figure 28)** and contains SNP positions 189, 195, 198, 199, and 200. With a MFI value of 44 and 38 bead #92 gave the highest hybridization signal. This data characterizes the SNP site 16298 as 16298C (i.e. not the same as the rCRS). The other four sites are typed as 189A, 195T, 198C, 199T and 200A (i.e. the same as the rCRS). The background signal from the competing beads is <23% and <8% of signal obtained from bead #92. For region ICa the data obtained from multiple samples established a background threshold of 25% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

IIC	rCRS	195C	199C	189G,200G	195C,198T
1C AMP.	1	1	1	0	0
2C AMP.	0	0	0	0	2
3C AMP.	44	6	10	4	0
4C AMP.	38	3	1	0	0
Bead	92	93	94	96	100

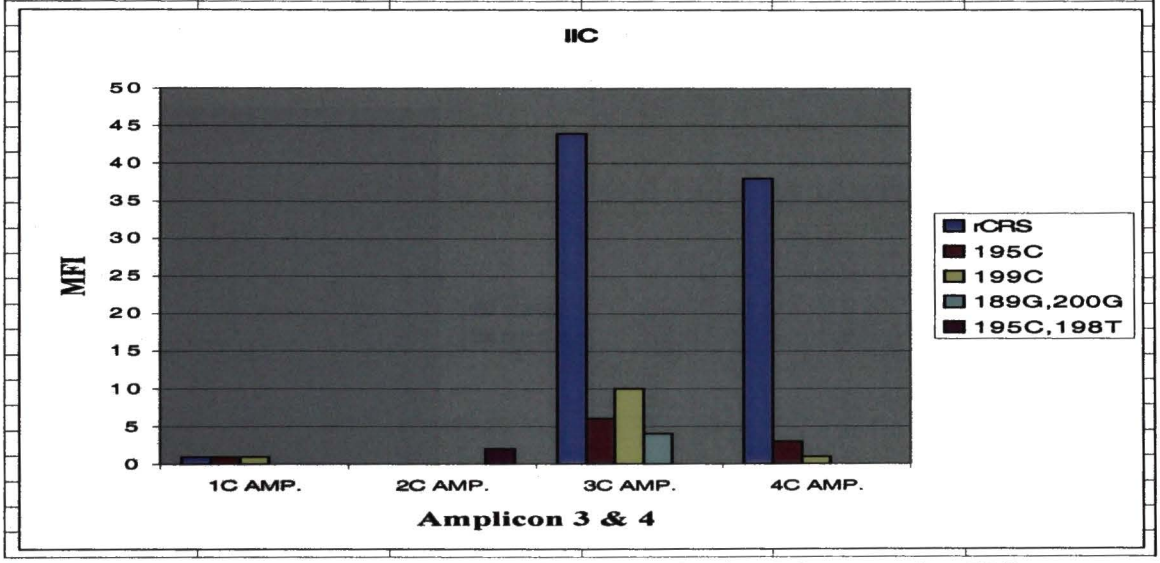

Figure 28. Representative hybridization data obtained from region IIC.

Region IID

Region IID is included in amplicon 3 and 4 **(Figure 29)** and contains SNP positions 263. With a MFI value of 153 and 60 bead #64 gave the highest hybridization signal. This data characterizes the SNP site 263 as 263G (i.e. not the same as the rCRS). The background signal from the competing beads is <46% and <43% of signal obtained from bead #64. For region IID the data obtained from multiple samples established a background threshold of 50% of the MFI value relative to the strongest hybridizing bead. Bead MFI values below this threshold were not evaluated.

IID	rCRS	263G
1C AMP.	0	2
2C AMP.	0	0
3C AMP.	71	153
4C AMP.	26	60
Bead	63	64

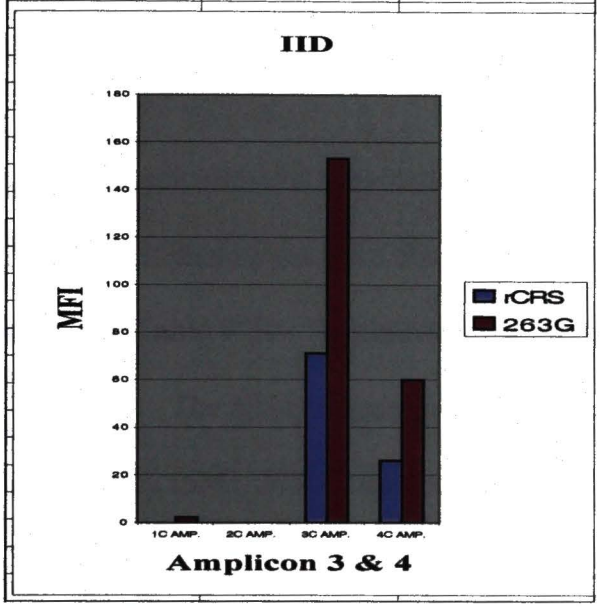

Figure 29. Representative hybridization data obtained from region IID.

MtDNA Haplotypes by SSO Hybridization Assay

Figures 20-29 contain SSO hybridization data from a single individual (C AMP). The result for this sample were representative of the data obtained from several hundred samples analyzed by this method. A total of 31 SNPs were interrogated per sample. Those SNPs whose nucleotide sequence differed from the rCRS are displayed when describing the overall haplotype. Therefore the haplotype for individual (C AMP) is shown below.

Haplotype for Sample (C AMP)

16223T, 16298C, 73G, 146C, 150T, 152C

The haplotype was consistent with mtDNA sequencing results, at those positions interrogated by the SSO capture probes.

Frequency Estimates for Observed Haplotypes in a Hispanic Population Sample Using SSO Hybridization

Fifty Hispanic samples provided by the FBI were typed by SSO Hybridization assay. A total of 30 different haplotypes were observed of these 22 haplotypes were observed only once in the data set. **Table 6** lists the observed haplotypes, the number of times each was observed and their frequency. The most frequent haplotype occurred 9 times or 18% [16223T, 16319A, 16362C, 73G, 146C, 153G, 263G], the second most frequent haplotype occurred 8% [16217C, 73G, 263G].

Table 6. Haplotypes Generated by the SSO Hybridization Assay

No.	Haplotype	Count	Freq. (p)	95%Confidence Interval (+/-)
1	16223T, 16319A, 16362C, 73G, 146C, 153G, 263G	9	0.18	0.1065
2	16217C, 73G, 263G	4	0.08	0.0752
3	16223T, 16319A, 16362C, 73G, 146C, 152C, 153G, 263G	3	0.06	0.0658
4	16223T, 16362C, 73G, 263G	3	0.06	0.0658
5	16223T, 16298C, 73G, 263G	3	0.06	0.0658
6	16217C, 73G, 146C, 263G	2	0.04	0.0543
7	16298C, 263G	2	0.04	0.0543
8	73G, 263G	2	0.04	0.0543
9	16126C, 16294T, 16319A, 16362C/T, 73G, 152C, 195C, 263G	1	0.02	0.0388
10	16126C, 73G, 263G	1	0.02	0.0388
11	16129A, 16223T, 16319A, 16362C, 73G, 146C, 153G, 263G	1	0.02	0.0388
12	16129A, 16223T, 73G, 199C, 263G	1	0.02	0.0388
13	16129A, 263G	1	0.02	0.0388
14	16217C, 16298C, 73G, 152C, 263G	1	0.02	0.0388
15	16217C, 16309G, 73G, 263G	1	0.02	0.0388
16	16217C, 16319A, 73G, 146C, 263G	1	0.02	0.00388
17	16217C, 73G, 152C, 263G	1	0.02	0.00388
18	16217C, 73G, 153G, 263G	1	0.02	0.00388
19	16223T, 16298C, 16311C, 16319A, 73G, 195C, 263G	1	0.02	0.00388
20	16223T, 16298C, 16311C, 73G, 263G	1	0.02	0.00388
21	16223T, 16298C, 73G, 153G, 263G	1	0.02	0.00388
22	16223T, 16298C, 73G, 198T, 263G	1	0.02	0.00388
23	16223T, 16311C, 16319A, 16362C, 73G, 146C, 152C, 153G, 263G	1	0.02	0.00388
24	16223T, 16319A, 16362C, 146C, 153G, 263G	1	0.02	0.00388
25	16223T, 16325C, 16362C, 73G, 263G	1	0.02	0.00388
26	16223T, 73G, 195C, 263G	1	0.02	0.00388
27	16223T, 73G, 263G	1	0.02	0.00388
28	16224C, 16311C, 73G, 146C, 152C, 263G	1	0.02	0.00388
29	16311C, 73G, 263G	1	0.02	0.00388
30	263G	1	0.02	0.00388

Population Genetics from SSO Hybridization Assay

For the sample set an unbiased estimate of the genetic diversity (equivalent to heterozygosity) was estimated using equation (1) $h = n(1-\Sigma x^2)/ (n-1)$ (Tajima, 1989) where n is the number of individuals in the sample population and x is the frequency of the haplotypes. The estimate of the fifty samples typed by SSO Hybridization Assay was:

$h = 0.9559$ or 95.59%

The probability of two randomly selected individuals from the same population group having the same mtDNA haplotypes was calculated using equation (2) $p = \Sigma x^2$ (Melton, Wilson et al., 1997)

For this data set random match probability was: $p = 0.0632$ or 6.32%

SNP Analysis by Single Base Extension Assay

Amplification of Template for Primer Extension

The amplified products for this study were generated using unlabeled primer sets A1-B1 and primer set C1-D1 for HVI and HVII, respectively. Two mtDNA amplicons were generated for each individual assayed. **Figure 30** is an example of a post amplification analytical gel in which the HVI and HVII amplicons from six individuals are displayed. The size of the HVI amplicon is~ 430 bp and HVII amplicon ~ 400 bp. The size of the amplicon may vary slightly as a result of the additions or deletion of bases within the control region.

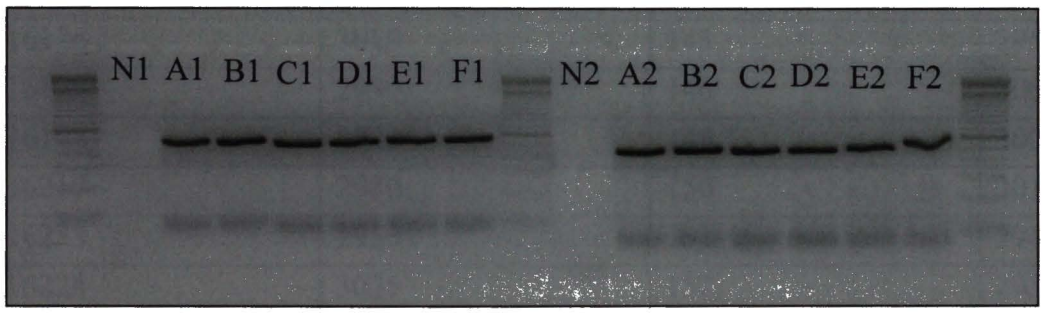

Figure 30. Post amplification analytical gel of six individuals. The amplicons for HVI are shown in lanes A1-F1. The amplicons for HVII are shown in lanes A2-F2. Lanes N1 and N2 are the negative amplification controls for HVI and HVII. Also 100 bp ladder was run on the gel to verify the size of the mtDNA amplicons.

Primer Extension

Initial primer extension assays were performed on samples with known mtDNA sequence. These initial assays were performed using only a single primer in order to evaluate the performance of each extension primer (prior to multiplexing). The relative migration of each extension primer was determined. Migration times are determined in terms of read points. The migrations read points were calibrated by performing multiple extensions of a single primer with different samples. Data from these initial assays revealed that the relative migration times

varied based upon the ddNTP incorporated (data not shown). The average migration read point for each SNP is shown in **Table 7**. The primers were designated based upon the SNP position they interrogate. Primers designated with "x.1" identify the same position as "x" but indicate that the primer sequence is different to account for variation at the primer binding site. Primer 16362 was redesigned, to extend the migration time. Primer 16362.1 contained an additional 14 bases at the 5' end of the original primer.

Table 7. Approximations of the Relative Migration Time for the Individual Primers.

HVI Primer	Migration time HV1	HVII Primer	Migration time HVII
16069.1	2970	73	3070
16093	3070	93	3110
16124	2990	93.1	2980
16126	3010	143	2950
16129	2990	146	3000
16172	2970	146.1	3000
16217	2980	150	3010
16223	2950	151	2940
16224	3025	152	2950
16292	2920	153	2960
16294	2940	189	2980
16295	2930	195	3000
16304	3000	198	3020
16309	3020	199	3000
16311	3005	200	2950
16320	2990	247	2980
16320.1	3030	263	2980
16325	3060		
16327	3020		
16362	2960		
16362.1	3080		

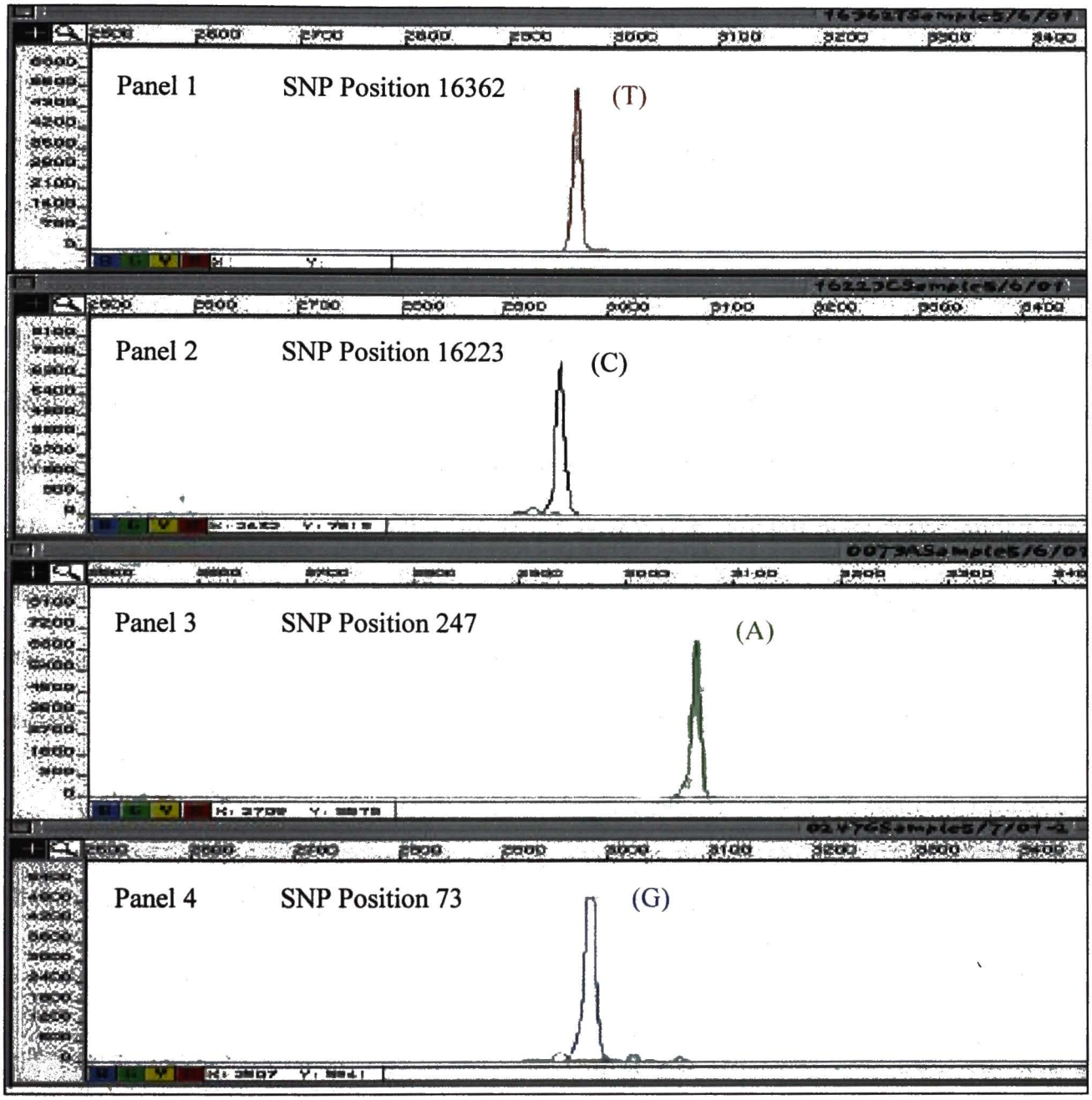

Figure 31. Single-plex primer extension assay for position 16362, 16223, 247 and 73.

The single-plex primer extension assays provided consistent and reliable base calls. **Figure 31 panels 1-4** illustrates four examples of typical results obtained when using a single-plex extension assay. The four SNPs positions interrogated were 16362, 16223, 247 and 73. Since each dideoxynucleotide is labeled with a different fluorophore, the dideoxynucleotide incorporated is represented as a colored peak (ddTTP;-ROX-red, ddCTP;-TAMRA-black,

ddGTP;-R110-blue and ddATP;-R6G-green). Therefore Thymine is visualized as a red peak, Cytosine a black peak, Guanine a blue peak and Adenine a green peak. Panel 1 identifies Thymine at position 16362, panel 2 identifies Cytosine at position 16223, panel 3 identifies Guanine at position 247 and the panel 4 identifies Adenine at position 73. Cytosine is represented by the color black in electropherogram due to visual acuity, when in reality the fluorphore TAMRA is yellow. The electrophoretic migration of the extension product is dependent upon the size of the primer, base composition, secondary structure and the specific dideoxynucleotide incorporated. The signal generated by the single base extension assay were robust (peak height > 5000 relative fluorescent units (rfu)) with very low noise. The low background is attributable to the specificity of the primers. Proper primer design insured binding to the correct sequence motif and preventing of primer slippage. Of the original 39 extension primers designed 37 gave reliable results in a single-plex reaction.

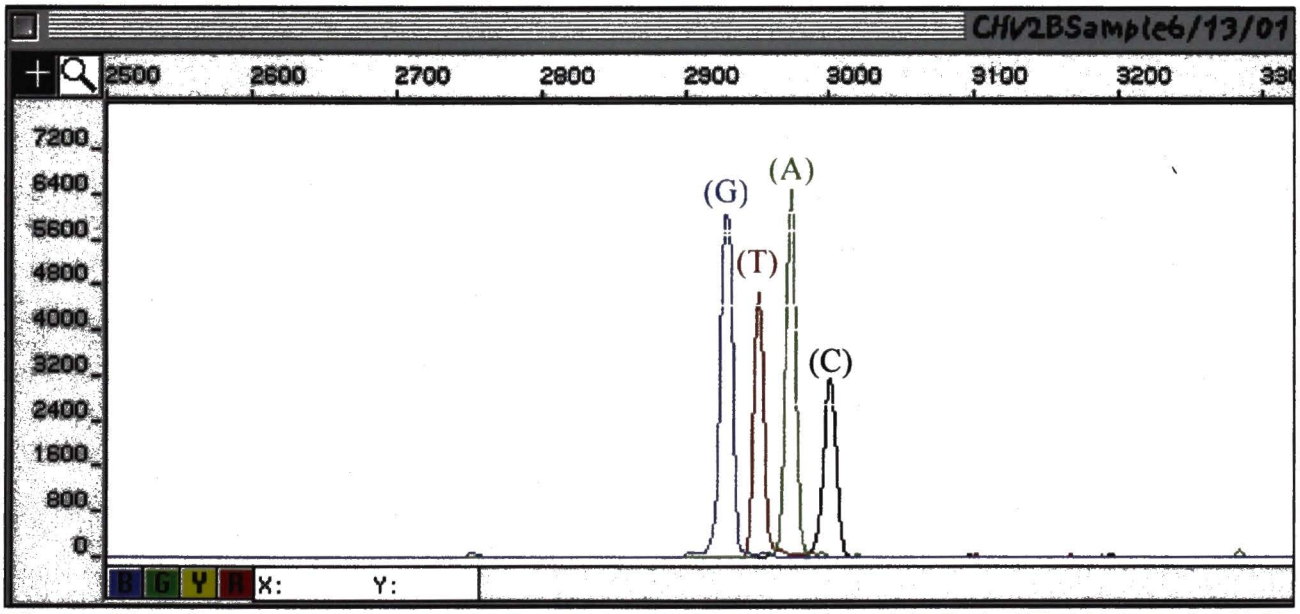

Figure 32. Illustration of multiplex assay identifying position 143G, 153T, 189A and 150C.

Migration data obtained from the single-plex reactions was used to determine potential multiplex compositions. In a multiplex primer extension assay the migration read point defines

the specific SNP position. One of the original multiplex extension assays is shown in **Figure 32.** SNP positions 143, 153, 189 and 150 are assayed in this multiplex assay. Each of the 4 SNP positions were clearly separated. The thirty seven SNP primers could be separated into eleven multiplexed primer extension reactions. In each multiplex reaction, the extension products were easily resolved by capillary electrophoresis. The multiplex extension assays are detailed in **Table 8** along with the template strand that they anneal to, the SNP state of the base defined for the rCRS (Andrews, Kubacka et al., 1999).

Table 8. Multiplex Extension Assays.

	SNP Position	Primer Binding Strand	rCRS
Multiplex 1	HVIAa		
	16223	Hyb. Rev.	C
	16129	Hyb. For.	G
	16126	Hyb. Rev	T
Multiplex 2	HVIAb		
	16069.1	Hyb. Rev.	C
	16217	Hyb. Rev.	T
	16224	Hyb. For.	T
Multiplex 3	HVIAc		
	16172	Hyb. Rev.	T
	16124	Hyb. Rev.	T
	16093	Hyb. Rev.	T
Multiplex 4	HVIBa		
	16292	Hyb. Rev.	C
	16320	Hyb. For.	C
	16309	Hyb. For.	A
	16362.1	Hyb. Rev.	T
Multiplex 5	HVIBb		
	16295	Hyb. Rev.	C
	16304	Hyb. For.	T
	16327	Hyb. For.	C
	16325	Hyb. For.	T
Multiplex 6	HVIBc		
	16294	Hyb. Rev.	C
	16311	Hyb. For.	T
	16320.1	Hyb. For.	C
Multiplex 7	HVIIAa		
	0151	Hyb. For.	C
	0152	Hyb. For.	T
	0150	Hyb. Rev.	C
	0093.1	Hyb. For.	A

Multiplex 8	HVIIAb		
	0143	Hyb. Rev.	G
	0153	Hyb. For.	A
	0146.1	Hyb. For.	T

Multiplex 9	HVIIAc		
	0093	Hyb. For.	A
	0146	Hyb. For.	T
	0073	Hyb. Rev.	A

Multiplex 10	HVIIBa		
	0263	Hyb. Rev.	A
	0199	Hyb. For.	T
	0247	Hyb. Rev.	G
	0198	Hyb. Rev.	C

Multiplex 11	HVIIBb		
	0200	Hyb. For.	A
	0189	Hyb. Rev.	A
	0195	Hyb. Rev.	T

Electropherogram Data

The fifty Hispanic blood samples provided by the FBI were analyzed with the multiplex SNaPshot™Primer Extension Assay. An example of an individual analysis is illustrated in **Figures 33 - 36.**

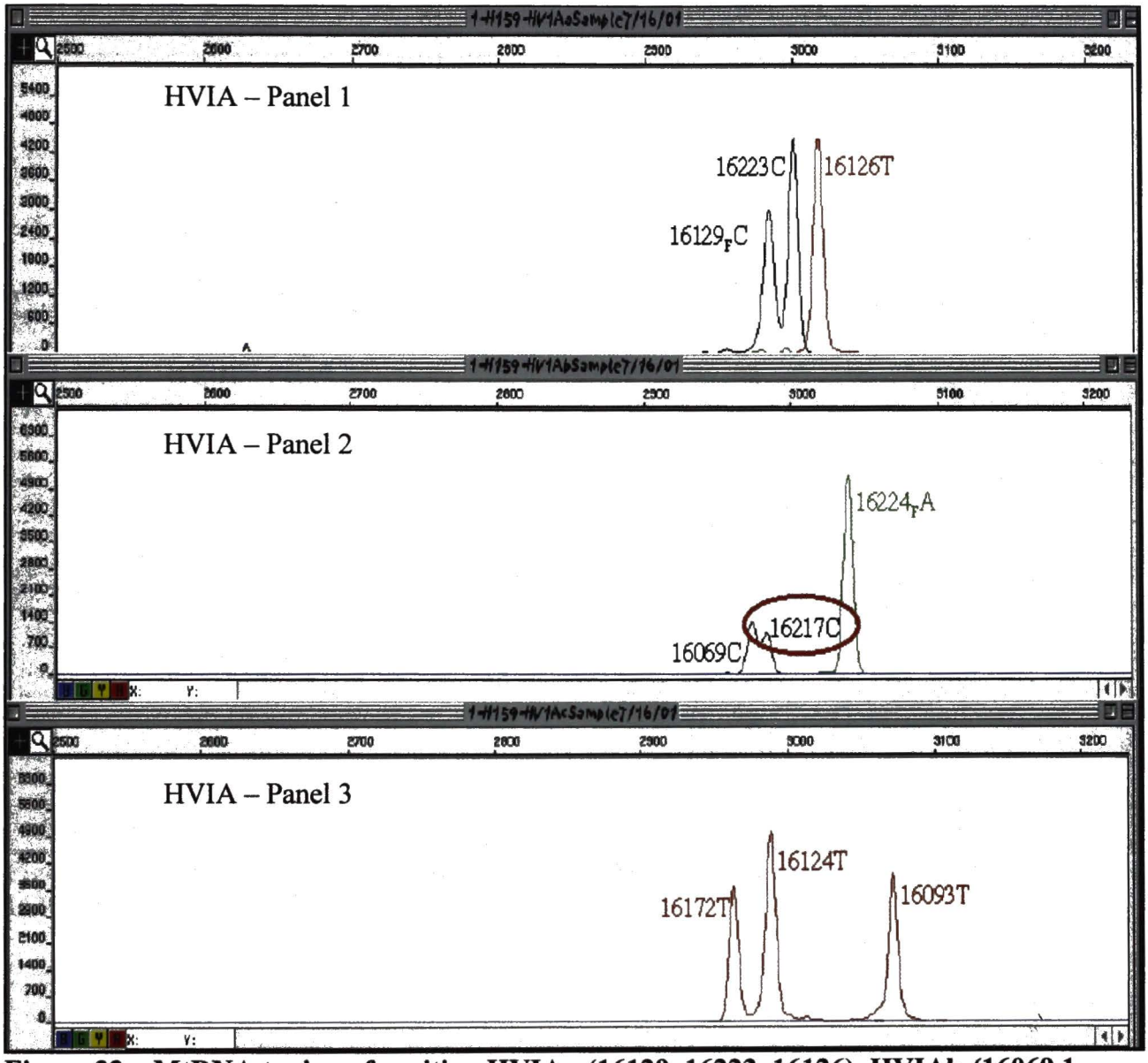

Figure 33. MtDNA typing of position HVIAa-(16129, 16223, 16126), HVIAb-(16069.1, 16217, 16224), HVIAc-(16172, 16124, 16093).

Figure 33 displays three multiplex reactions containing a total of 9 SNP positions for region HVIA Of these 9 SNP positions, 8 are the same state as the rCRS. Only one position 16217 (circled) is different from the reference sequence and contains Cytosine. The haplotype for these 9 SNPs is recorded as 16217C.

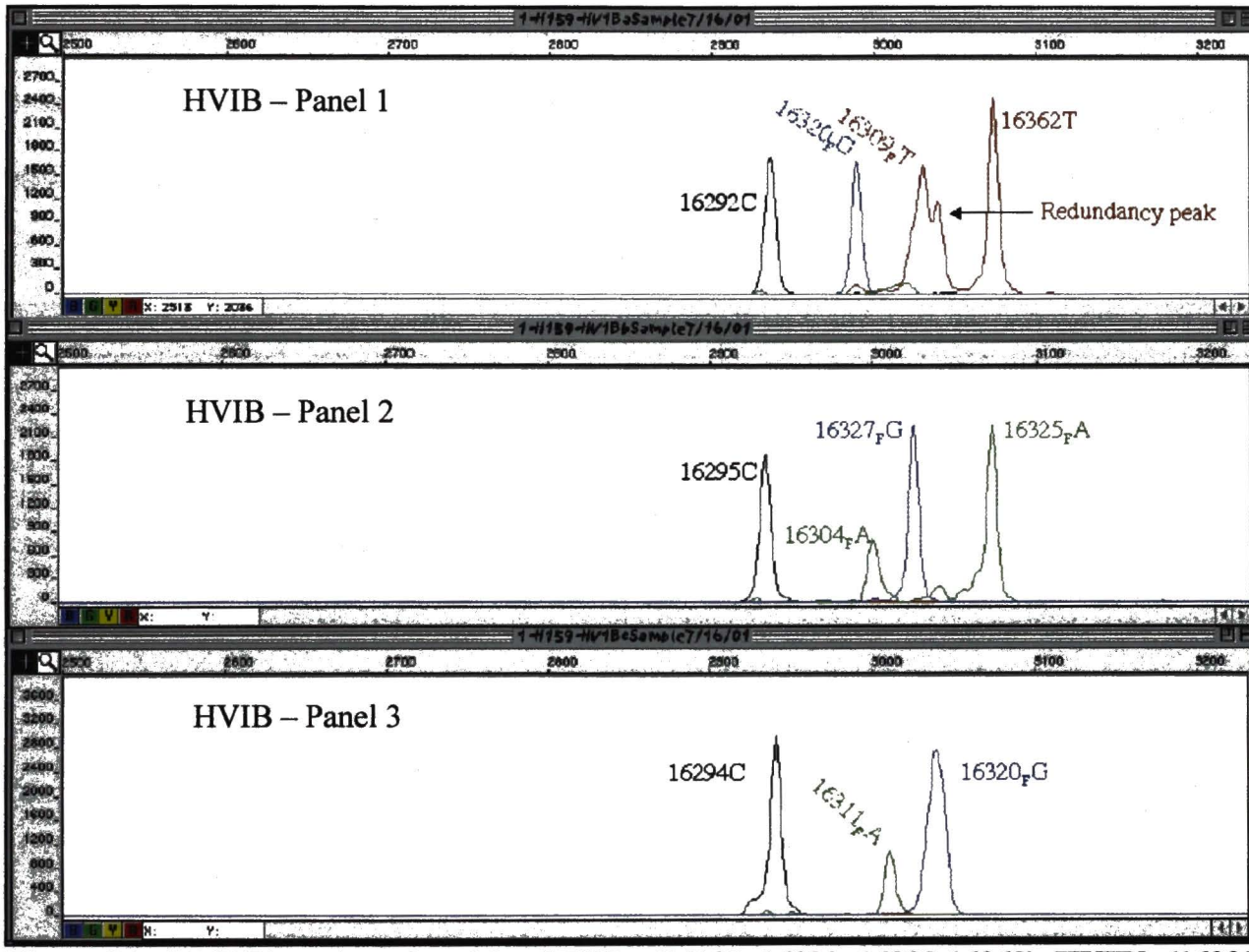

Figure 34. MtDNA typing of position HVIBa-(16192, 16320, 16309,16362), HVIBb-(16295, 16304, 16327, 16325), HVIBc-(16294, 16311, 16320.1).

Figure 34 displays three multiplex reactions containing a total of 11 SNP positions for region HVIB. All 11 SNP positions, 11 have the same nucleotide sequence as rCRS. Thus, these 11 SNPs do not contribute any polymorphisms to the haplotype designation.

SNP position 16309 (**Figure 34, Panel 1**) provides an example of an electropherogram which redundant extension primers are used to interrogate the same SNP position. At SNP position 16309 a doublet peak was observed. The sequence for the extension primer for SNP position 16309 is ('5-TATGTACGGTAAATGRCTTTATGTRC-3'). The primer has two position where the letter "R" is inserted. The letter "R" designates that both an Adenine and Guanine are present at this position. When the extension primer for SNP position 16309 was

synthesized, 4 different oligonucleotides were produced. The 4 oligos will account for all possible base combinations at the two position designated with the letter "R". The double peak results from the extension off 2 different primers that annealed prior to SNP position 16309.

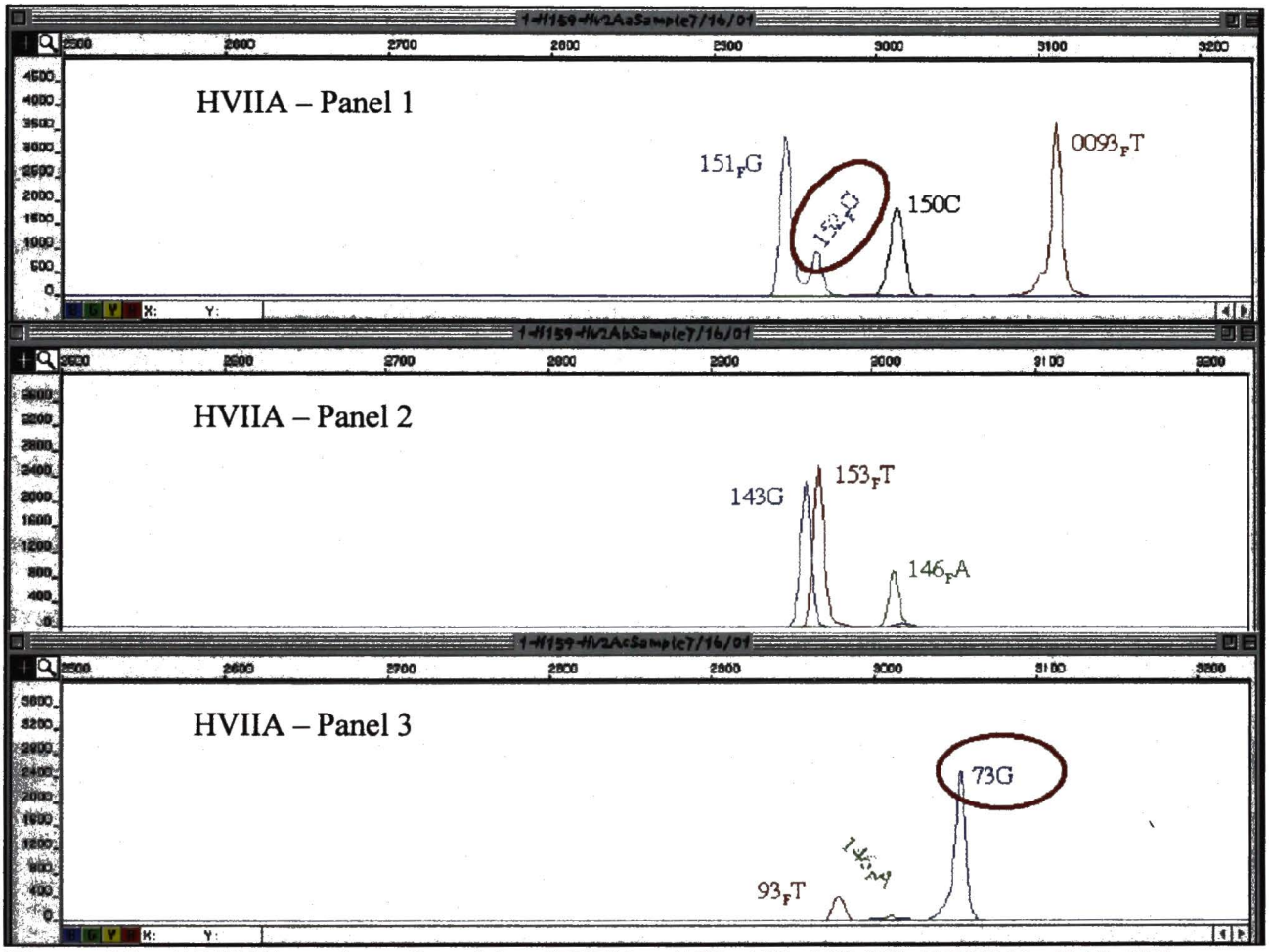

Figure 35. MtDNA typing of position HVIIAa-(151, 152, 150,93), HVIIAb-(143, 153, 146,), HVIIAc-(93, 146, 73)

Figure 36 displays three multiplex reactions assaying a total of 10 SNP positions in region HVIIA. Of these 10 positions, 8 have the same nucleotide sequence as rCRS. Two positions, 152 and 73 are different than the reference sequence and both contain "G" (Guanine). For these 10 SNPs the polymorphisms 152C and 73G were contributed to the haplotype designation. Note that position 152 is record as a "C" (Cytosine) as opposed to a "G" as shown

in the figure. This is because the primer for position 152 anneals to the heavy strand (forward strand) of the mitochondrial sequence and as a result the base incorporated is the nucleotide that is complementary to the position being interrogated. In this particular case the nucleotide complementary to the position being investigated is "G". In order to maintain the convention used in reporting mtDNA polymorphism, the haplotype for this position (152) is "C".

As discussed earlier, in order to accommodate variation in motifs due to SNPs within the primer annealing site, experimental assays dictated that multiple primers need to be incorporated within the multiplex, which are similar in sequence except each are complementary to a different polymorphism in the primer binding site. Multiple primers annealing to the same region provide redundancy in the extension assay. Two different primers were used to interrogate SNP position 146 **(Figure 35 Panels 1 and 3).** The same is true for SNP position 93 **(Figure 35 Panels 1 and 2).** Note in **Figure 35 Panel 3** the signal for 93T and 146A are low, but still detectable and typed correctly.

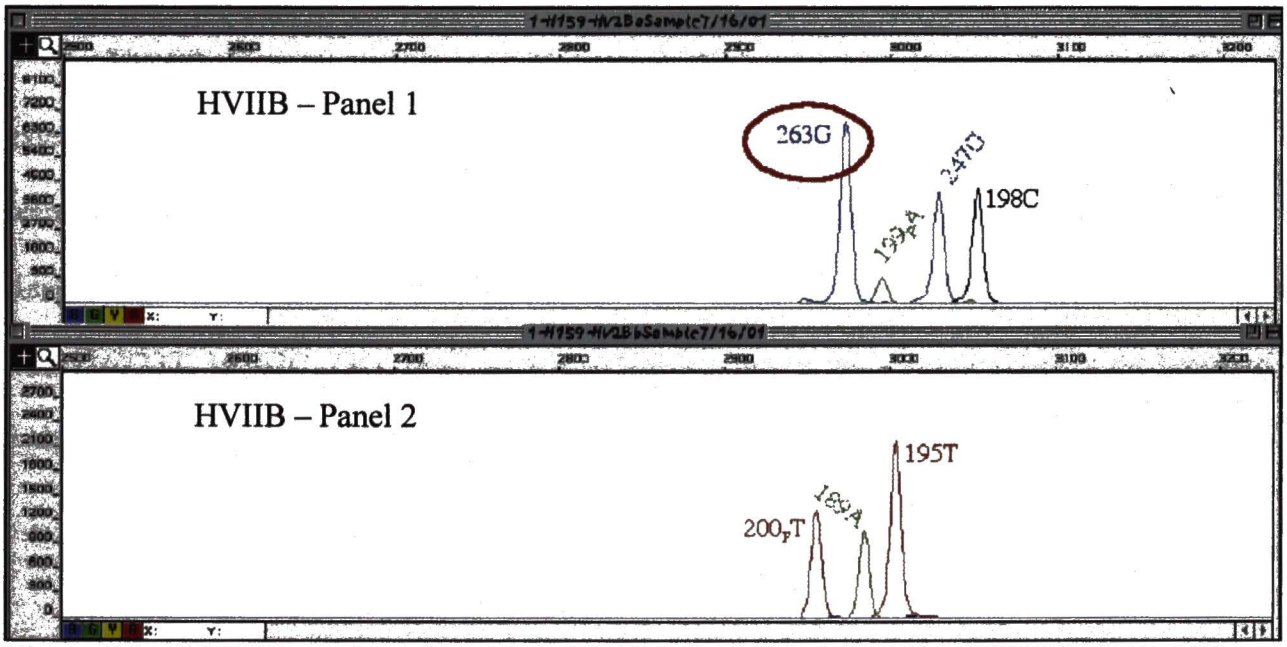

Figure 36. MtDNA typing of position HVIIBa-(263,199, 247, 198,), HVIIBb-(200, 189, 195).

Figure 36 displays two multiplex reactions containing a total of 7 SNP positions in region HVIIB. Of the 7 SNP position, 6 have the same sequence as the rCRS. Only SNP position 263 (circled) is different than the reference sequence and contains "G". The haplotype for these 7 SNPs is recorded as 263G.

Haplotype for Sample 1-H159

A total of 34 SNPs were interrogated per sample. **Figures 33-36** displays the results from the single base extension assay for a single individual. Only those SNPs whose state differed from rCRS are expressed in the haplotype. Therefore, the haplotype for this individual is listed as follows: 16217C, 152C, 73, 263.

This haplotype was consistent with mtDNA sequencing results for the SNPs interrogated.

Frequency Estimates for Observed Haplotypes in a Hispanic Population Sample Using a Single Base Extension Assay

The fifty hispanic samples provided by the FBI were analyzed using the single base extension assay. A total of 30 different haplotypes were observed. Of these, 22 haplotypes were observed only once in the data set. **Table 9** lists the observed haplotypes, the number of times each were observed and their frequency. The most frequent haplotype occurred 7 times or 14% [16223T, 16362C, 73G, 146C, 153G, 263G], the second most frequent haplotype occurred 5 times or 10% [16223T, 16325C, 16327T, 73G, 263G].

Table 9. Haplotypes Generated by Primer Extension Assay.

No.	Haplotype	Count	Frequency (p)	95% Confidence Interval (+/-)
1	16069T,16126C,73G,263G	1	0.02	0.0388
2	16093C,16223T,16325C,16362C,73G,263G	1	0.02	0.0388
3	16126C,16294T,73G,152C,195C,263G	1	0.02	0.0388
4	16129A,16223T,16362C,73G,146C,153G,263G	1	0.02	0.0388
5	16129A,16223T,73G,263G	1	0.02	0.0388
6	16129A,263G	1	0.02	0.0388
7	16172C,16311C,73G,263G	1	0.02	0.0388
8	16217C,16309G,73G,263G	1	0.02	0.0388
9	16217C,73G,143A,263G	1	0.02	0.0388
10	16217C,73G,146C,152C,263G	1	0.02	0.0388
11	16217C,73G,146C,263G	3	0.06	0.0658
12	16217C,73G,152C,263G	1	0.02	0.0388
13	16217C,73G,153G,263G	1	0.02	0.0388
14	16217C,73G,263G	4	0.08	0.0752
15	16223T,16311C,16325C,16327T,73G,195C,263G	1	0.02	0.0388
16	16223T,16311C,16325C,16327T,73G,263G	1	0.02	0.0388
17	16223T,16311C,16362C,73G,146C,152C,153G,263G	1	0.02	0.0388
18	16223T,16319A,16362C,73G,146C,153G,263G	1	0.02	0.0388
19	16223T,16325C,16327T,73G,198T,263G	1	0.02	0.0388
20	16223T,16325C,16327T,73G,263G	2	0.04	0.0543
21	16223T,16325C,16327T,73G,146C,153G,263G	1	0.02	0.0388
22	16223T,16325C,16327T,73G,263G	5	0.1	0.0832
23	16223T,16362C,73G,146C,153G,263G	7	0.14	0.0962
24	16223T,16362C,73G,146C,152C,153G,263G	2	0.04	0.0543
25	16223T,16362C,73G,146C,153G,263G	1	0.02	0.0388
26	16223T,16362C,93T,146C,153G,263G	1	0.02	0.0388
27	16223T,73G,195C,263G	1	0.02	0.0388
28	16224C,16311C,73G,146C,152C,263G	1	0.02	0.0388
29	263G	3	0.06	0.0658
30	73G,263G	2	0.04	0.0543

Population Genetics from Primer Extension Assay

For the sample set an unbiased estimate of the genetic diversity (equivalent to heterozygosity) was estimated using equation (1) $h = n(1-\Sigma x^2)/(n-1)$ (Tajima, 1989) where n is the number of individuals in the sample population and x is the frequency of the haplotypes. The estimate of the fifty samples typed by Single Base Extension Assay was:

$h = 0.9624$ or 96.24%

The probability of two randomly selected individuals from the same population group having the same mtDNA haplotypes was calculated using equation (2) $p = \Sigma x^2$ (Melton, Wilson et al., 1997) For this data set the random match probability was: $p = 0.0568$ or 5.68%

Identification of Heteroplasmy

Heteroplasmy can be defined as two nucleotides detected at a SNP position. The electropherogram for sample H213 position 16362 in HVIB suggested the presence of heteroplasmy (**Figure 37**). To confirm that the result for position 16362 was heteroplasmy and not an artifact of the multiplex extension assay, the primer for this position was used by itself with sample H213. The single-plex reaction confirms that two nucleotides are detected at this position (**Figure 38**). The results from the SSO Hybridization Assay for this sample were reviewed (**Figure 39**). The histogram also suggest that this sample was heteroplasmic at position 16362. The MFI values for the two beads interrogating this position were nearly identical. Cycle sequencing was done on sample H213. The results from the Sanger sequence analysis of both the forward and reverse strand confirmed the presence of the heteroplasmy (**Figure 40**).

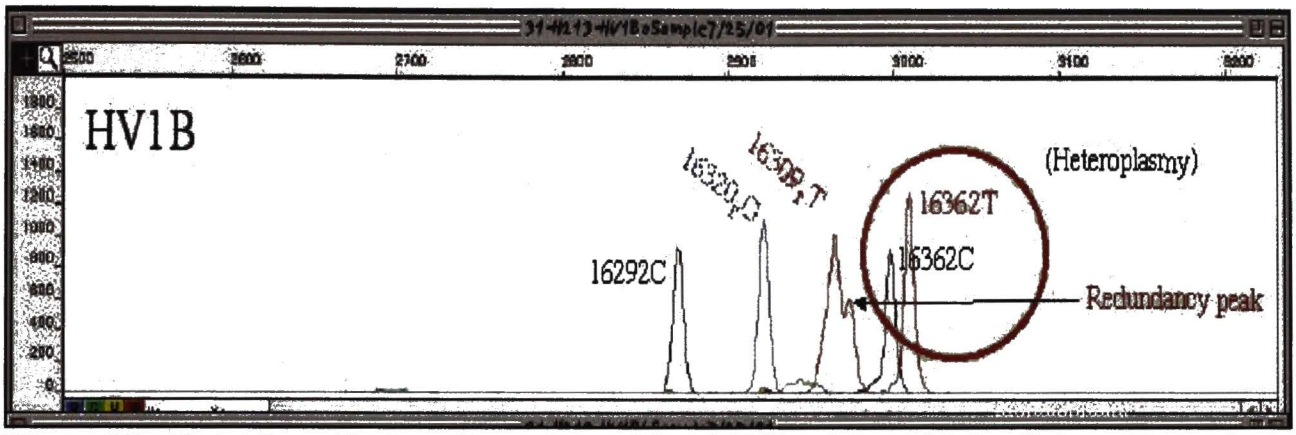

Figure 37. Detection of heteroplasmy at position 16362 using multiplex Primer Extension and the ABI Prism® 310 Genetic Analyzer.

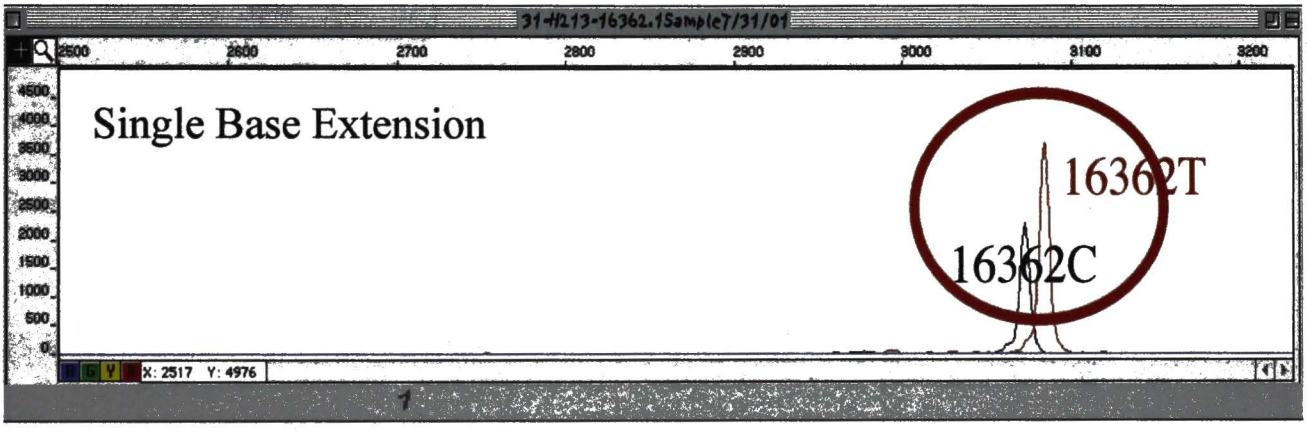

Figure 38. Detection of heteroplasmy at position 16362 using a single primer for Primer Extension with analysis by the ABI Prism® 310 Genetic Analyzer.

ID	ANDERSON	16362C
31-H213	0	2
31-H213	103	95
31-H213	0	2
31-H213	0	1
Bead	72	73

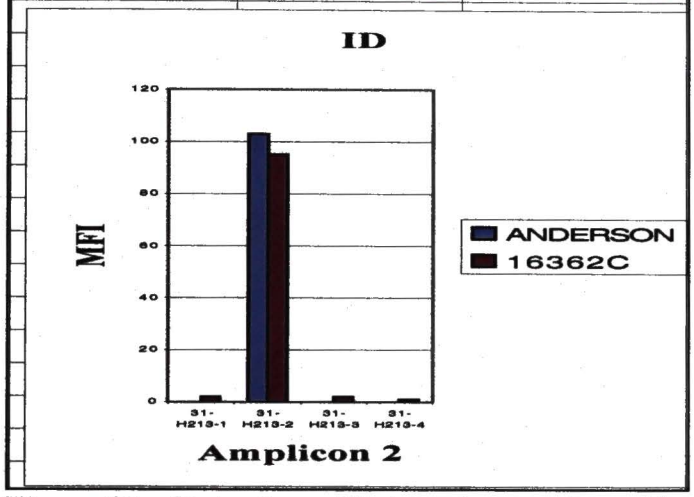

Figure 39. Detection of heteroplasmy at position 16362 by Competitive SSO hybridization and the Luminex 100™ system.

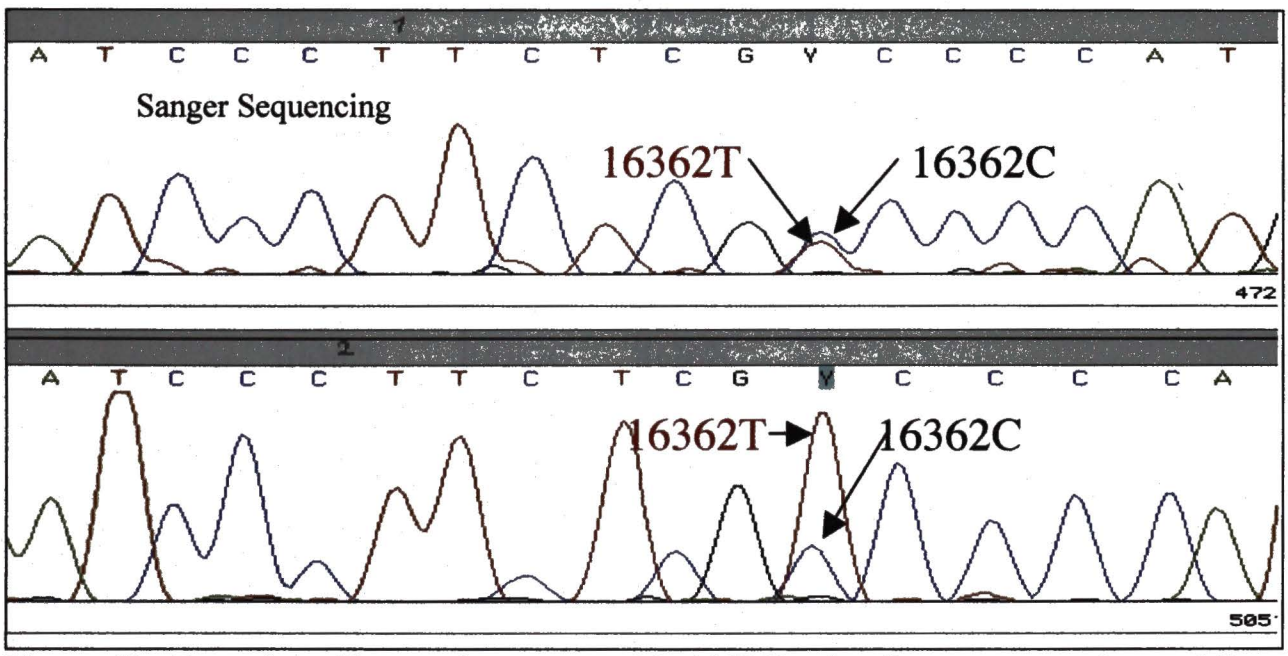

Figure 40. Detection of heteroplasmy at position 16362 using Sanger sequencing and the 310 ABI Genetic Analyzer.

Table 10. Comparison of Haplotypes obtained for Sequencing, Primer Extension and SSO Hybridization.

Cycle sequencing was used to interrogate all nucleotides positions within HVI and HVII for the 50 samples provided by the FBI. The two SNP assays developed were only able to interrogate a subset of these nucleotide positions. The SSO Hybridization assay on the Luminex 100™ instrument examined 31 SNP positions, and the Snapshot® Single Base Extension assay on the ABI PRISM® 310 Genetic Analyzer examined 34 SNP positions. The haplotypes determined for each of the three techniques are shown below in **Table 10.**

No.	INDIVDUAL	SEQUENCE	SBE	SSO
1	HIS.159	16183C		
		16189C		
		16217C	16217C	16217C
		16256T		
		73G	73G	73G
		146C		
		152C	152C	152C
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		
2	HIS.160	16223T	16223T	16223T
		16270T		
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G	263G	263G
		263G		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
3	HIS.161	16069T	16069T	
		16126C	16126C	16126C
		16145A		
		16222T		
		16261T		
		73G	73G	73G
		263G	263G	263G
		295T		
		315.1C		
4	HIS.162	16223T	16223T	16223T
		16325C	16325C	
		16362C	16362C	16362C
		73G	73G	73G
		263G	263G	263G
		315.1C		146C
5	HIS.163	16093C	16093C	
		16223T	16223T	16223T
		16274A		
		16325C	16325C	
		16362C	16362C	16362C
		73G	73G	73G
		204C		
		263G	263G	263G
		315.1C		
6	HIS.169	16223T	16223T	16223T
		16298C		16298C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		198T	198T	198T
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
7	HIS.170	16182C		
		16183C		
		16189C		
		16217C	16217C	16217C
		73G	73G	73G
		153G	153G	153G
		263G	263G	263G
		309.1C		
		315.1C		
8	HIS.171	16111T		
		16129A	16129A	16129A
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G	263G	263G
		263G		
		309.1C		
		315.1C		
9	HIS.172	16182C		
		16183C		
		16189C		
		16217C	16217C	16217C
		16319A		16319A
		73G	73G	73G
		146C	146C	146C
		263G	263G	263G
		309.1C		
		309.2C		
		309.3C		
		315.1C		
10	HIS.174	16234T		
		16298C		16298C
		263G	263G	263G
		309.1C		
		315.1C		146C
11	HIS.175	16129A	16129A	16129A
		263G	263G	263G
		309.1C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
12	HIS.176	263G 309.1C 315.1C	263G	263G 146C
13	HIS.177	16182C 16183C 16189C 16217C 73G 263G 309.1C 309.2C 315.1C	 16217C 73G 263G	 16217C 73G 263G
14	HIS.178	16183C 16189C 16217C 16270G 16278T 73G 146C 263G 309.1C 315.1C 316N	 16217C 73G 146C 263G	 16217C 73G 146C 263G
15	HIS.179	16223T 16325C 16362C 73G 263G 292C 309.1C 315.1C	16223T 16325C 16362C 73G 263G	16223T 16362C 73G 263G

No.	INDIVDUAL	SEQUENCE	SBE	SSO
16	HIS.180	16111T		
		16223T	16223T	16223T
		16239T		
		16290T		
		16311C	16311C	16311C
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		152C	152C	152C
		153G	153G	153G
		198N	263G	263G
		235G		
		253N		
		263G		
17	HIS.182	16111T		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		315.1C		
18	HIS.183	16223T	16223T	16223T
		16298C		16298C
		16311C	16311C	16311C
		16319A		16319A
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		195C	195C	195C
		214G		
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
19	HIS.184	16223T	16223T	16223T
		16298C		16298C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		249D		
		263G	263G	263G
		290D		
		291D		
		315.1C		
20	HIS.190	16092C		16129A
		16129A	16129A	
		16148T		
		16223T	16223T	16223T
		16354T		
		73G	73G	73G
		199C		199C
		204C		
		250C		
		263G	263G	263G
		309.1C		
		315.1C		
21	HIS.191	16051G		
		16223T	16223T	16223T
		16298C		16298C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		
22	HIS.192	16223T	16223T	16223T
		16325C	16325C	
		16362C	16362C	16362C
		73G	73G	73G
		204C		
		263G	263G	263G
		315.1C		146C

No.	INDIVDUAL	SEQUENCE	SBE	SSO
23	HIS.193	16240G		
		16298C		16298C
		263G	263G	263G
		309.1C		
		315.1C		
24	HIS.194	16086C		
		16183C		
		16189C		
		16217C	16217C	16217C
		73G	73G	73G
		139C		
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		
25	HIS.196	16182C		
		16183C		
		16189C		
		16234T		
		73G	73G	73G
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		
26	HIS.197	16111T		
		16209C		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		94A		
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
27	HIS.199	16111T		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		315.1C		
28	HIS.200		16223T	16223T
			16362C	16319A
			93T	16362C
			146C	146C
			153G	153G
			263G	263G
				93T
29	HIS.201	16153A		
		16223T	16223T	16223T
		16298C		16298C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		153G	153G	153G
		249D		
		263G	263G	263G
		290D		146C
		291D		
		309.1C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
30	HIS.212	16111T		
		16223T	16223T	16223T
		16239T		
		16259T		
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		152C	152C	152C
		153G	153G	153G
		214G		
		235G		
		263G	263G	263G
		309.1C		
31	HIS.213	16126C	16126C	16126C
		16163G	16362C/T*	16362C/T*
		16186T		
		16189C		
		16257T		
		16294T		16294T
		73G	73G	73G
		152C	152C	152C
		183G		
		195C	195C	195C
		263G	263G	263G
		309.1C		
		315.1C		
32	HIS.214	16183C		
		16189C		
		16194C		
		16195C		
		16217C	16217C	16217C
		16298C		16298C
		73G	73G	73G
		114G		
		152C	152C	152C
		207A		
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
33	HIS.215	16223T	16223T	16223T
		16298C		16298C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		
34	HIS.217	16223T	16223T	16223T
		16298C		16298C
		16311C	16311C	16311C
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		
		333C		
35	HIS.219	16111T		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
36	HIS.220	16163G		
		16164T		
		16172C	16172C	
		16219G		
		16311C	16311C	16311C
		73G	73G	73G
		263G	263G	263G
		309.1C		
		315.1C		
37	HIS.221	16075C		
		16111T		
		16223T	16223T	16223T
		16239T		
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		315.1C		
38	HIS.222	16223T	16223T	16223T
		16325C	16325C	
		16327T	16327T	
		73G	73G	73G
		249D		
		263G	263G	263G
		290D		
		291D		
		309.1C		
		315.1C		
39	HIS.223	16111T		
		16223T	16223T	16223T
		16274A		
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
40	HIS.224	16183C		
		16189C		
		16223T	16223T	16223T
		16248T		
		16278T		
		73G	73G	73G
		195C	195C	195C
		263G	263G	263G
		309.1C		
		315.1C		
41	HIS.225	16111T		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		152C	152C	152C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		315.1C		
42	HIS.226	16111T		
		16223T	16223T	16223T
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		315.1C		
43	HIS.227	16183C		
		16189C		
		16217C	16217C	16217C
		16298C		16298C
		73G		
		114G	73G	73G
		146C	146C	146C
		263G	263G	263G
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
44	HIS.228	16182C		
		16183C		
		16189C		
		16224C	16224C	16224C
		16311C	16311C	16311C
		73G	73G	73G
		146C	146C	146C
		152C	152C	152C
		207A		
		263G	263G	263G
		315.1C		
45	HIS.229	16111T		
		16223T	16223T	16223T
		16239T		
		16256T		
		16290T		
		16319A		16319A
		16362C	16362C	16362C
		73G	73G	73G
		146C	146C	146C
		152C	152C	152C
		153G	153G	153G
		235G		
		263G	263G	263G
		309.1C		
		315.1C		
46	HIS.230	16183C		
		16189C		
		16217C	16217C	16217C
		16278T		
		73G	73G	73G
		263G	263G	263G
		309.1C		
		309.2C		
		315.1C		

No.	INDIVDUAL	SEQUENCE	SBE	SSO
47	HIS.231	16092C 16111T 16223T 16249C 16290T 16319A 16362C 73G 146C 153G 235G 263G 309.1C 309.2C 315.1C	16223T 16362C 73G 146C 153G 263G	16223T 16319A 16362C 73G 146C 153G 263G
48	HIS.232	16182C 16183C 16189C 16217C 73G 143A 263G 309.1C 315.1C	 16217C 73G 143A 263G	 16217C 73G 263G
49	HIS.233	16182C 16183C 16189C 73G 263G 309.1C 309.2C 315.1C	 73G 263G	 73G 263G
50	HIS.234		16217C 16309G 73G 263G	16217C 16309G 73G 263G

The cycle sequence analysis was clearly able to identify many more polymorphisms since it interrogates every position within the HVI and HVII amplicons. Sequencing also identified the presence of insertions and deletions relative to the revised Cambridge Reference Sequence. Neither the Sequence Specific Oligonucleotide Hybridization assay nor the Single Base Extension assay was capable of identifying the presence of insertions or deletions. Comparisons with sequencing could only be made at the 31 positions interrogated by the SSO assay and the 34 positions interrogated by the SBE assay. Although the FBI sent 50 samples, only sequence data for 48 of these samples were provided. The sequence data for samples HIS.200 (our sample #28) and HIS.234 (our sample #50) was not available to compare with the two SNP assays. In every case were sequencing identified a SNP, and the corresponding position was interrogated by the Single Base Extension assay, the correct base call was made. However, in 7 of the 48 samples, the Sequence Specific Oligonucleotide Hybridization assay incorrectly identified the appropriate haplotype. In 5 samples (HIS.162, 174, 176, 192, and 201), the SSO assay incorrectly identified position 146C. In each case, position 146 was the reference sequence or "T". In sample HIS.169, position 198T was not identified, and in sample HIS.213, position 16294T was also not identified. In these two samples the SSO method indicated the positions were identical to the revised Anderson reference sequence. Ambiguity was also detected in sample HIS.200. Although sequence data was not available for this sample, comparisons between the two SNP assays could be made at the corresponding base positions. Relative to the Single Base Extension assay, the Sequence Specific Oligonucleotide Hybridization assay failed to detect the SNP at positions 16362C and 93T. Position 93T is of considerable interest since this base represents a transversion from the reference sequence which has an "A" at this position.

A transversion represents the substitution of a Purine nucleotide (Adenine and Guanine) for a pyrimidine (Cytosine and Thymine) or vice versa. The Sequence Specific Oligonucleotide Hybridization assay was not capable of detecting transversions, only transition type substitutions. Since all 4 dideoxy chain terminators are used in the ABI PRISM® SNaPshot™ Multiplex Kit, the Single Base Extension assay is capable of detecting both transition and transversion base substitutions.

DISCUSSION

The goal of this thesis project was to develop alternative methods to traditional forensic mtDNA sequence analysis. Conventional forensic mtDNA analysis requires the direct sequencing of HVI and HVII in both the forward and reverse directions. This method is time consuming, labor intensive and expensive. Less than 20 laboratories in the United States have the capability of performing forensic mtDNA analysis. The current rate of throughput for mtDNA typing is approximately 1-2 cases per staff member per month (Melton and Nelson, 2001). Using traditional sequencing technology, it does not appear feasible to increase sample throughput without compromising the integrity of the assay. Furthermore 98-99% of the information obtained during sequencing is uninformative. In fact, only 1-2% of the sequence data will reveal useful genetic polymorphisms.

Two methods for determining mtDNA haplotypes through the direct interrogation of SNPs within HVI and HVII have been developed. A Sequence Specific Oligonucleotide Hybridization assay was developed on the Luminex 100 ™ flow cytometer, as well as a Single Base Extension assay for the ABI Prism® 310 Genetic Analyzer. The SNP typing of mtDNA sequences can provide a significant benefit in many forensic and human identification cases. The reassociation of mass disaster remains, mass grave analysis, and the screening of large numbers of crime scene samples are several examples of their potential application. Their inclusion as a standard screening tool would be highly beneficial in directing a more extensive

DNA analysis toward those samples that possess the greatest evidentiary value. This would be most beneficial to those forensic laboratories that are unable to perform direct sequence analysis (due to time and/or budget constraints) on large numbers of samples.

Sequence Specific Oligonucleotide Hybridization Assay on the Luminex 100™ Platform

The Sequence Specific Oligonucleotide Hybridization assay involves annealing one strand of a PCR product spanning the SNP of interest to an oligonucleotide capture probe complementary to each of the alternative SNP states and measuring the relative hybridization efficiency. The LabMAP™ technology developed by Luminex was utilized in the design of our SSO Hybridization assay for mtDNA haplotype analysis. The LabMAP™ technology consists of two major components; color-coded microspheres with covalently attached sequence specific oligonucleotide capture probes and a bench top flow cytometer (Luminex 100™). A set of 100 different color-coded microspheres was available, therefore, up to 100 different capture probes could be utilized. The LabMAP™ technology is similar to the reverse dot blot assay used in the AmpliType® PM+DQA1 PCR Amplification and Typing Kit (Applied Biosystems), where several different probes attached to a membrane strip are hybridized to a PCR product, and the captured PCR product is detected. However, while the reverse dot blot procedure involves denaturation, hybridization, washing, binding of the detector (Streptavidin-Alkaline Phosphatase), another washing, followed by the addition of a colorimetric substrate resulting in colored dots, the LabMAP™ technology format is homogeneous. All of the oligonucleotide capture probes, once bound to microspheres, are blended together into a single mixture. For this assay, Bodipy-TMR fluorescently labeled primers are incorporated into the HVI and HVII amplification products. An asymmetric PCR amplification process was used so that the labeled

PCR products were predominantly single stranded. The labeled PCR products were hybridized with the microsphere mixture for approximately 45 minutes to permit sequence specific hybridization to occur. The hybridization reaction is then read directly on the Luminex 100™ instrument. The Luminex 100™ is equipped with a XY-platform which accommodates a 96 well microtiter plate. This allows for the automated analysis of up to 96 samples in less than one hour. The Luminex 100™ instrument reports the median fluorescent intensity (MFI) of the labeled PCR product captured by each color-coded microsphere. The LabMAP™ technology offers numerous advantages over other SNP detection methods: the Luminex 100™ instrument is relatively inexpensive (approximately $35,000) as compared with other flow cytometers and other detection equipment routinely found in DNA forensic labs; the XY platform allows an increase in sample throughput by using 96-well microtiter plates instead of single tubes; multiplexing capability reduces both labor and reagent cost, and reduces sample to sample inconsistencies; and new SNP positions can be easily added to the multiplex reaction.

SNP discrimination depends upon the existence of a difference in stability between fully matched probe-target duplexes and mismatched probe-target duplexes. A hybridization temperature of 55°C was chosen so that mismatched probe-target duplexes would be unstable. This is essential to minimize non-specific binding of the labeled HVI and HVII amplification products to the wrong bead-capture probe set. Therefore, the Sequence Specific Oligonucleotide Hybridization assay is a very temperature-sensitive process. The temperature of the heat block and the XY platform is critical and must be checked frequently to ensure optimal hybridization. If the temperature was to drop below the 55°C optimal temperature, non-specific hybridization could occur and result in an erroneous haplotype. Alternatively, if the

hybridization temperature was above 55°C, a reduction in the MFI values could occur, and a SNP position may fail to be identified.

A panel of 45 oligonucleotide capture probes was synthesized and coupled with 45 different color-coded microspheres. The capture probes were designed to interrogate a total of 31 SNPs. The 31 SNPs were distributed into 10 regions within HVI and HVII. HVI was divided into 5 hybridizing regions containing 17 SNPs: 1A – 3 SNPs; 1B – 3 SNPs; 1Ca – 6 SNPs; 1Cb – 4 SNPs; and 1D – 1 SNP. HVII was also divided into 5 hybridizing regions containing 14 SNPs: 2Aa – 1 SNP; 2Ab – 1 SNP; 2B – 6 SNPs; 2C – 5 SNPs; and 2D – 1 SNP. Regions HVI-1D, HVII-2Aa, HVII-2Ab, and HVII-2D each contain a single SNP and are therefore interrogated by 2 microspheres (per SNP), representing the two possible alternate bases. The other six regions are interrogated by between 4 and 11 microspheres. Region HVII-2B spans 11 bases, which is interrogated by 11 microspheres and includes 6 SNPs. The SSO Hybridization assay was complicated by the presence of other polymorphisms surrounding the SNP of interest. The oligonucleotide capture probes were designed to be complementary to the most common substitutions seen in the FBIs mtDNA database. The clustering of SNPs within a small region significantly complicated the design of these capture probes. One of the major limitations in the SSO Hybridization assay is that it was not possible to include a capture probe for every permutation of SNPs within a given region such as HVII-2B. As a consequence, the interpretation of the hybridization results was made more difficult. In several instances, mismatched capture probes and labeled target sequences were not sufficiently destabilized and resulted in non-specific binding, and incorrect SNP calls.

Limitations of the Sequence Specific Oligonucleotide Hybridization Assay

There were a number of factors that made the SSO Hybridization assay more laborious and challenging relative to the SBE assay. The initial challenge was the development of the microsphere capture reagent. Each of the 45 oligonucleotide capture probes had to be covalently attached to a particular color-coded microsphere. The attachment chemistry was fairly routine, however, there was not a simple assay available to monitor the coupling efficiency. Each microsphere could have between 100,000 to 500,000 capture oligonucleotides attached on the surface. Variability in the coupling efficiencies could result in significant fluctuations in the MFI values and potentially erroneous SNP calls. Once the individual coupled microspheres were synthesized, they were counted and mixed together at a concentration of 400 of each bead set per μL. During the development of the SSO Hybridization assay it appeared that the color-coded microspheres were not balanced in the bead mix. Several of the bead sets were under represented in the total bead counts relative to others used in the interrogation of SNPs within the same region. If a bead set was present in a lower concentration, then the corresponding MFI values obtained would also be reduced. Since the SNP designations are dependant on the MFI values, an erroneous SNP assignment could be made for a particular sequence motif. The imbalance was observed for the following regions: HVII-2A bead 93A; HVII-2B beads 143A, 143A 146C 152C, and 146C 153G; HVII-2C bead 195C; and HVII-2C the bead complementary to the revised Anderson Reference Sequence (Data not Shown). The MFI values for region HVII-2B were typically much lower than those of other regions. This region had the largest number of incorrect SNP calls **(Table 10)**. During the development of our mtDNA SSO Hybridization assay, a research group from Glaxo Wellcome (Research Triangle Park, North Carolina) developed an alternative to the direct coupling of the oligonucleotide

capture probes to the microspheres (Iannone et al., 2000). Each capture probe had a 25-nucleotide ZipCode sequence at the 5' end and an allele-specific sequence at the 3' end. The color-coded microspheres were coupled with a 25-base cZipCode. The cZipCodes were complementary to their corresponding ZipCodes and were used to capture the probe-target complex after hybridization. The use of these ZipCodes would simplify the development of new capture probes. Rather than having to synthesize new coupled beads when changes in a capture probe or when additional capture probes were required, a new oligonucleotide would be synthesized with the corresponding ZipCode at its 5' end. The use of ZipCodes will greatly enhance the utility of the SSO Hybridization assay.

The SSO Hybridization assay required the input of labeled single stranded amplification product. This mandated the synthesis of Bodipy-TMR labeled primers and the optimization of an asymmetric PCR amplification system. In order to obtain a sufficient quantity of labeled PCR product, a high cycle number was required. The requirement of a high cycle number resulted in samples with contamination and spurious amplification products. In samples where this was observed, significantly lower MFI values were obtained.

The stringent hybridization temperature requirement and the physical manipulations needed for the handling of the microspheres contributed to spurious results in the SSO Hybridization assay. As previously discussed, failure to maintain the required 55°C hybridization temperature resulted in the binding of mismatched microspheres, and incorrect SNP calls. The hybridization reactions took place in an external heat block and were rapidly transferred into the Luminex 100™ flow cytometer. Although the XY platform on the Luminex 100™ instrument was heated, we were not able to incorporate a thermocouple device to accurately monitor the temperature of the chamber. The assay would have been greatly

simplified and perhaps more consistent if the hybridization and analysis could be done entirely within the Luminex 100™ instrument. The coupled microspheres have a propensity to clump. As a result, they must be constantly vortexed and sonicated to keep them resuspended. These additional steps made the analysis more laborious and time consuming. Further, the microspheres required protection from ambient light due to photobleaching. The bead mix and the hybridization reaction tubes were wrapped in aluminum foil to protect against photobleaching. It had been observed that a 1 hour exposure to light caused a spectral shift in certain of the microspheres which could lead to a misclassification in the Luminex 100™ instrument (Taylor et al., 2001).

Sequence Specific Oligonucleotide Hybridization Data Interpretation

The interpretation of the Sequence Specific Oligonucleotide Hybridization results proved to be the most challenging aspect of the assay. The MFI values generated from the Luminex 100™ instrument were exported into an Excel spreadsheet. After background correction the MFI values for each bead region were presented as histograms for interpretation. Several factors that influenced the amount of labeled HVI and HVII target sequences bound during the hybridization reaction have been previously described. Those regions exhibiting very low MFI values could usually be traced to poor asymmetric amplification, or bead mix imbalance. For each sample, a total of 6,000 events (beads) were analyzed. This represented approximately 8 percent of the total beads added to the hybridization reaction. Since only a small portion of the beads with a labeled target were counted, it is possible that the MFI values were low based upon a stochastic effect. This situation could have been further exasperated in those microsphere sets that were under represented in the bead mix. However, the most significant factor affecting the

accurate SNP determination was the inability to destabilize mismatched capture probes and labeled target sequences. This was most evident in regions were numerous polymorphisms clustered.

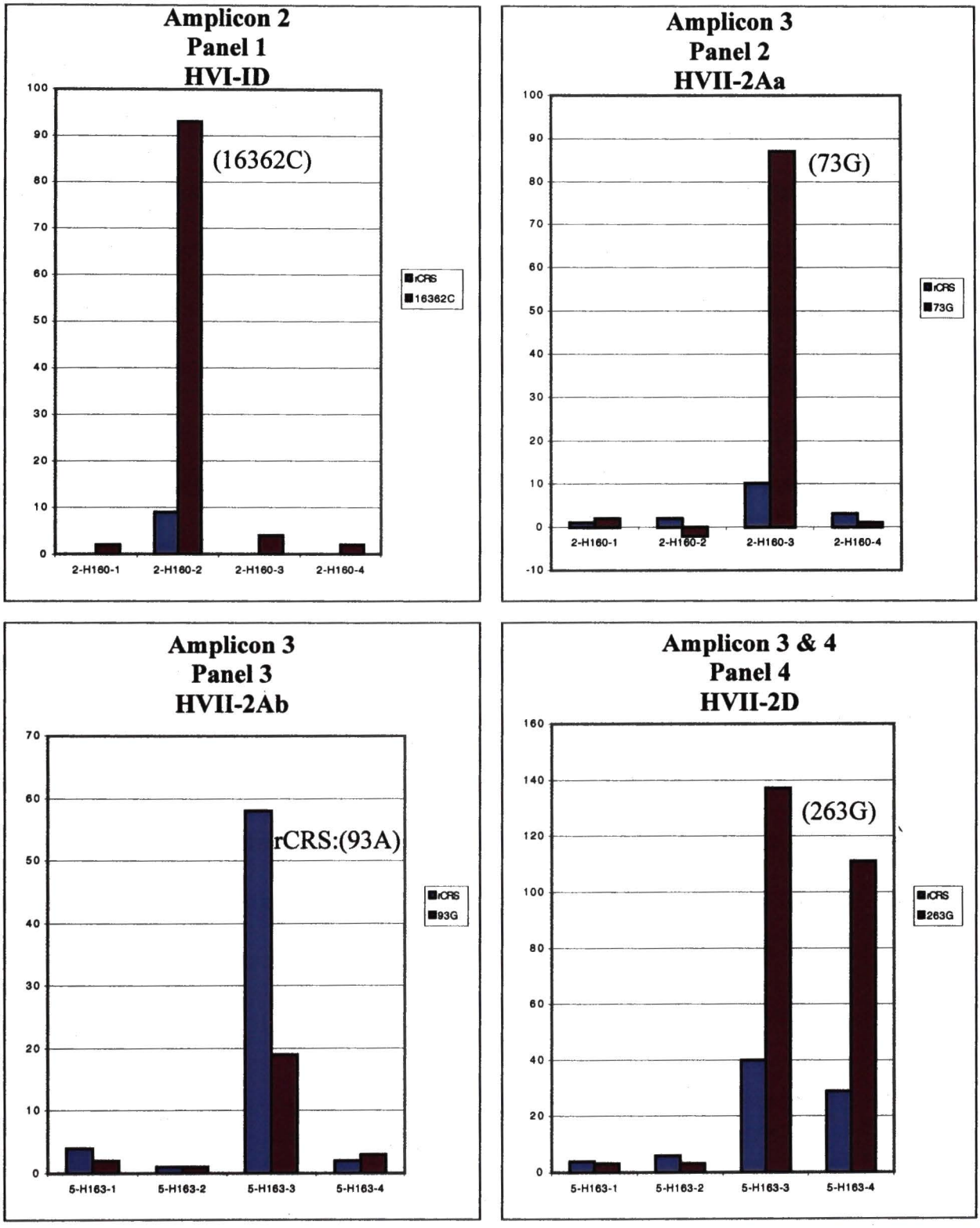

Figure 41. Typical SSO Hybridization results from the 4 regions which were interrogated by a pair of microspheres representing the two possible alternative bases.

Regions (HVI-1D, HVII-2Aa, HVII-2Ab, and HVII-2D) that contain a single SNP, were interrogated by 2 microspheres, representing the two possible alternate bases. The interpretation of the hybridization results from these regions was extremely reliable and reproducible (**Figure 41**). **Figure 41, Panels 3 and 4** demonstrate a limitation in the SSO Hybridization assay. Although the correct base is identified in these two examples, a significant amount of binding has occurred between the labeled target sequence and the alternate bead. It is not possible to determine if the signal was the result of heteroplasmy at these two positions or non-specific binding of the mismatched alternate bead. The sequence data and the Single Base Extension assay confirm this position as 263G. However, neither of these methods indicated the presence of heteroplasmy at this position in sample HIS.163. SNP position 263 (**Figure 41, Panel 4**) is present in both amplicons 3 and 4. The hybridization results between these two amplicons were consistent in this sample and all others analyzed. The results from sample HIS.200 (**Table 10**) demonstrated a limitation in the detection of sequence transversions. Position 93 is T, whereas the two capture beads identify either A or G. None of the capture probes were designed to identify transversions. It has been reported that transition mutations occur approximately 3 times as often as transversions in HVI and HVII(Holland and Parson, 1999). Even at this rate, a significant number of erroneous haplotypes could occur. This is not a problem in the Single Base Extension assay since all 4 labeled dideoxynucleotides are present.

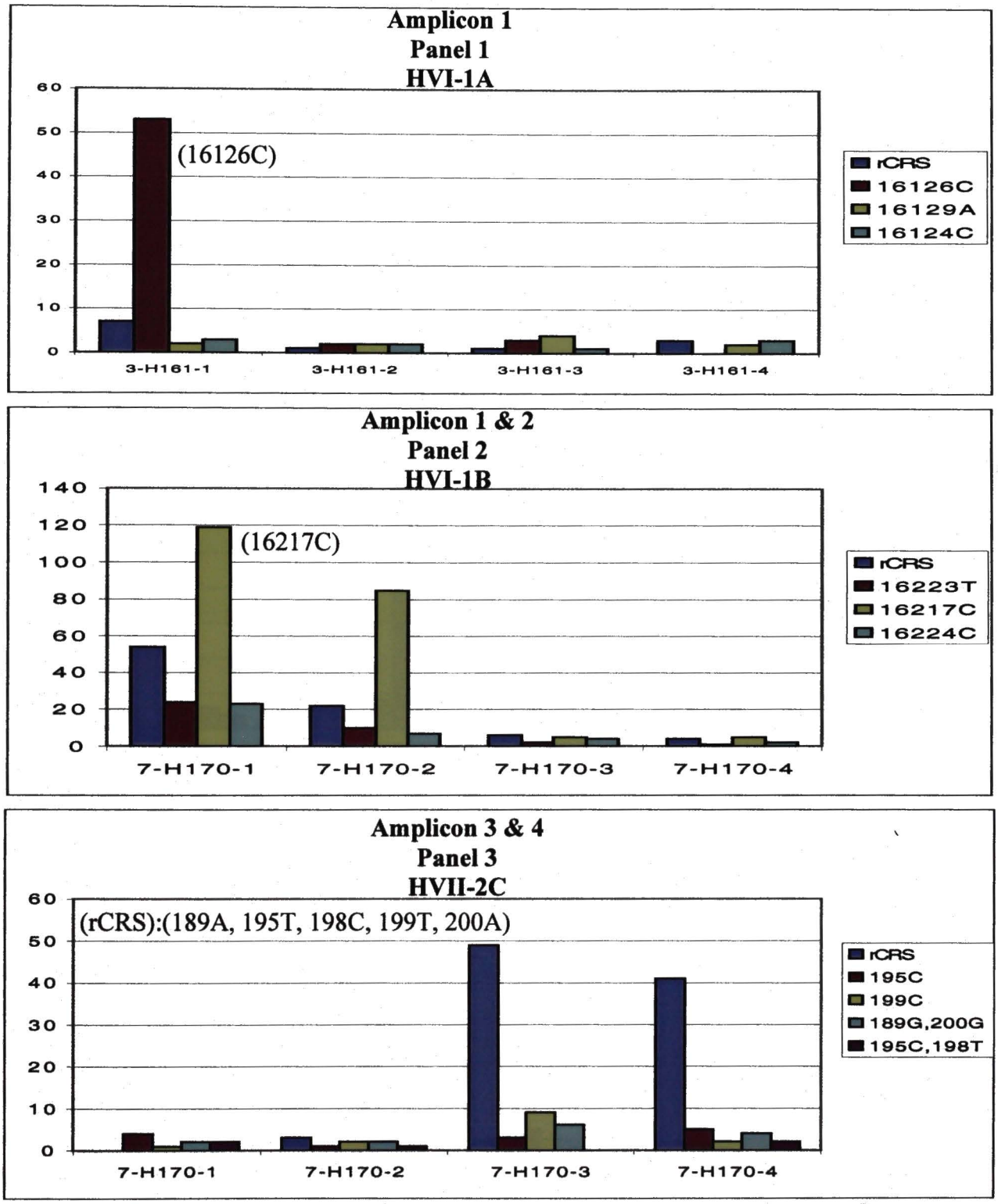

Figure 42. Typical SSO Hybridization results from 3 regions interrogated by either 4 or 5 microspheres that are used to identify 3-5 SNP positions.

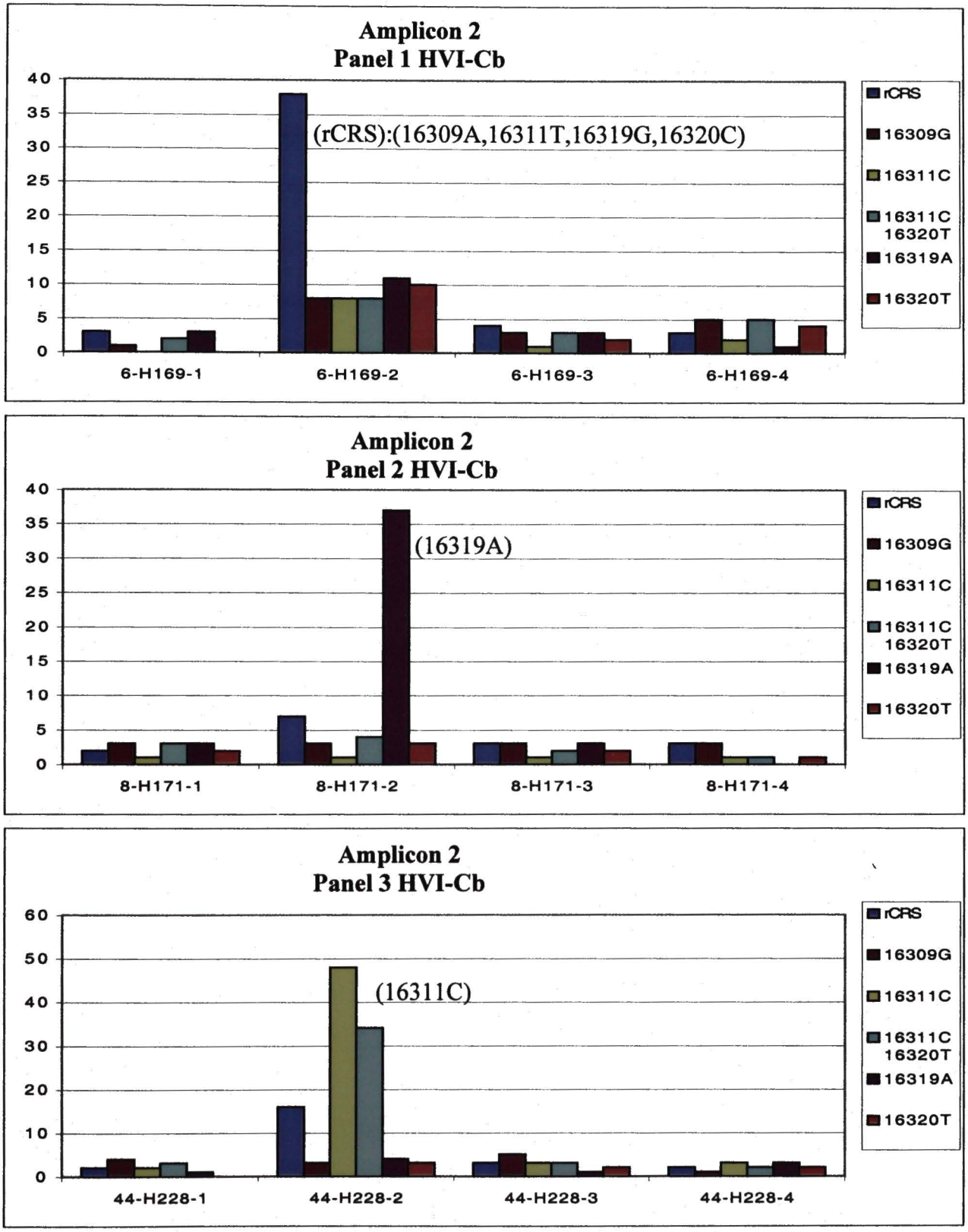

Figure 43. Typical SSO Hybridization results from region HVI-Cb which is interrogated by six beads to identify 4 SNP positions.

The interpretation of the hybridization results gets progressively more difficult in regions containing a significant number of polymorphic positions. **Figures 42 and 43** are examples of regions with a moderate density of polymorphic positions. With the exception of sample HIS.169 all of these regions were correctly typed. The correct haplotype at this region is 198T, however, the only capture bead for this position hybridizes to 195C – 198T. Position 195 in sample HIS.169 is T, which corresponds to the revised Anderson Reference Sequence. Since only one of the two bases was present in the target sequence, the SSO Hybridization assay failed to detect position 198T. **Figure 43, Panel 3** is an example where a significant amount of hybridization occurred with a bead with a single mismatched capture probe (bead 16311C – 16320T). Similar results are seen in **Figure 44, panel 3** that represents the SNP dense region HVI-1Ca. Although the correct haplotype was determined (rCRS), a significant amount of hybridization was observed with the capture probe for position 16294T. This single base mismatch was insufficient to completely destabilize the heteroduplex.

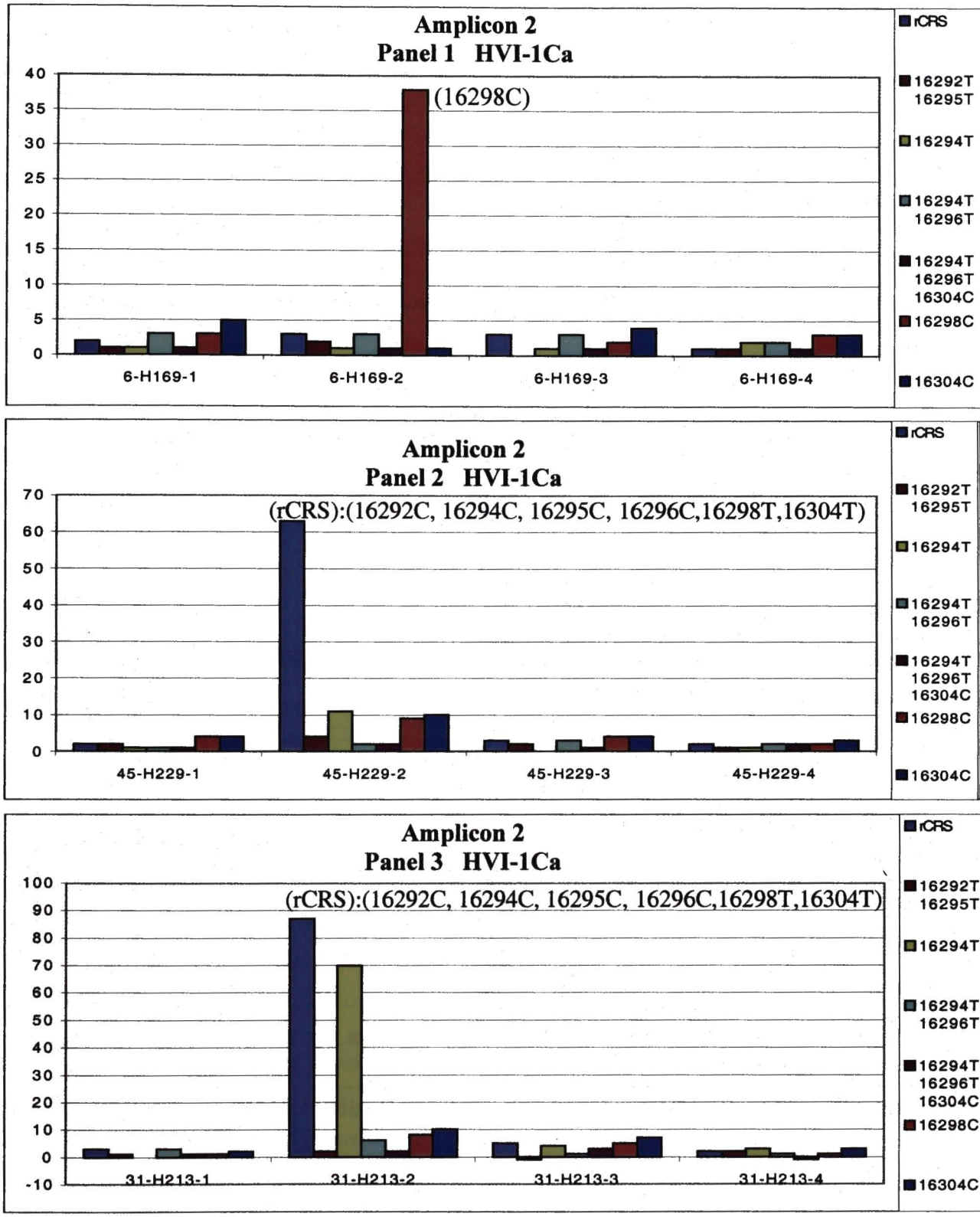

Figure 44. Typical SSO Hybridization results from region HVI-Ca which is interrogated by 7 beads to identify 6 SNP positions.

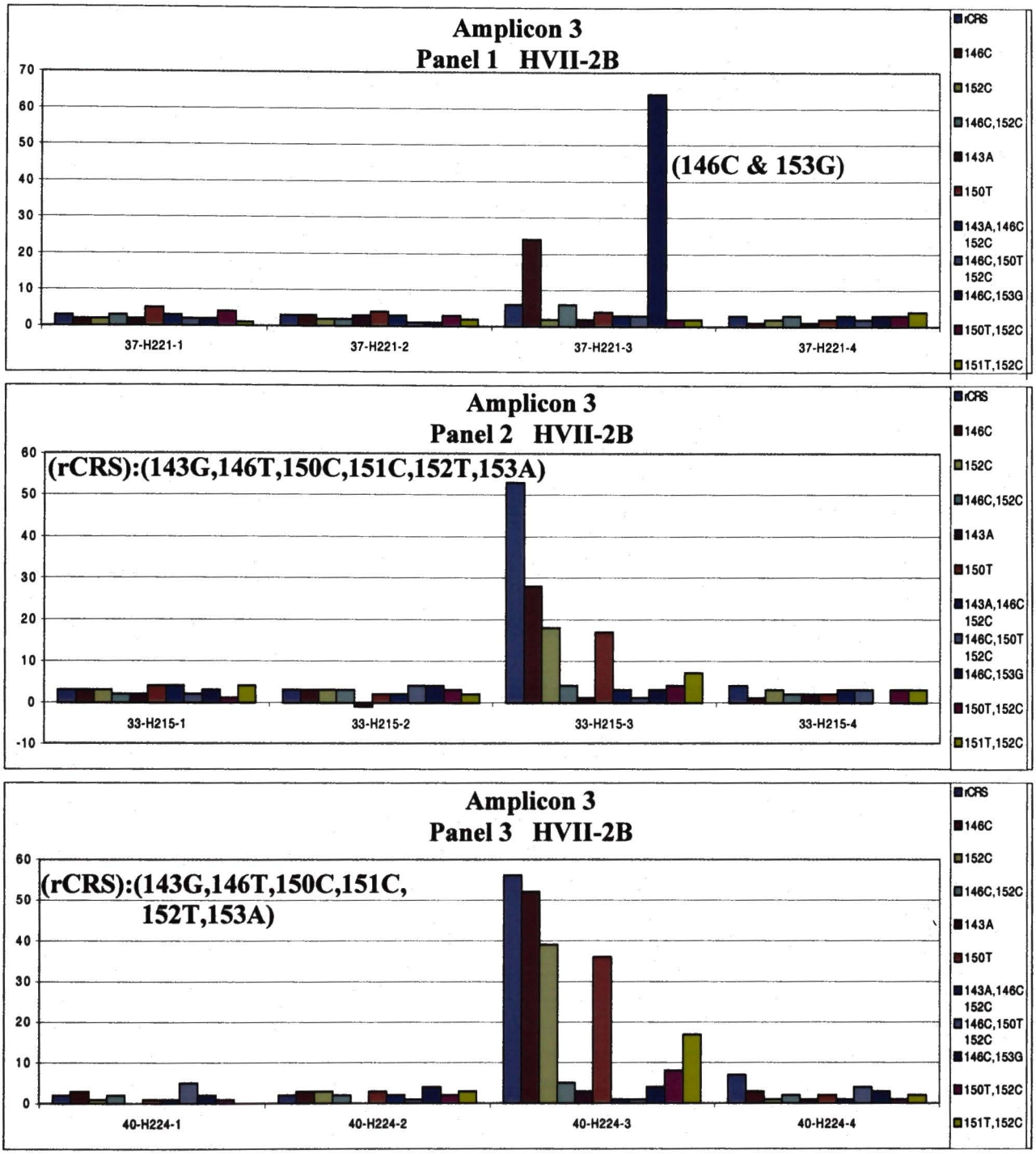

Figure 45. Typical SSO Hybridization results from region HVII-2B which is interrogated by 11 beads to identify 6 SNP positions.

HVII-2B proved to be the most difficult region to correctly haplotype. Five of the seven discrepancies that were detected between the sequence data and the SSO Hybridization results occurred in this region. **Figure 45, panels 2 and 3** demonstrate the difficulty in the

interpretation of the hybridization results from this region. Although both these samples have the revised Cambridge Reference Sequence in this region, a significant amount of hybridization occurs with several of the other bead sets. SNP position 146C was miscalled in each of the five samples erroneously typed. **Figure 46** is an example of an erroneous call within region HVII-2B. The bead with the greatest MFI value identified SNP positions 146C – 153G. The correct call for this sample was only 153G. This is another example in which a capture probe was not available solely for a unique SNP position. The nucleotide at position 146 is T, this single base mismatch was insufficient to destabilize the labeled target sequence. The Sequence Specific Oligonucleotide Hybridization assay using the 45 capture probes was insufficient to correctly type HVI and HVII. In order to utilize this hybridization assay as an effective mtDNA screening tool, several of the capture probes must be redesigned. In addition, numerous other probes must be synthesized to account for the most common transversion positions.

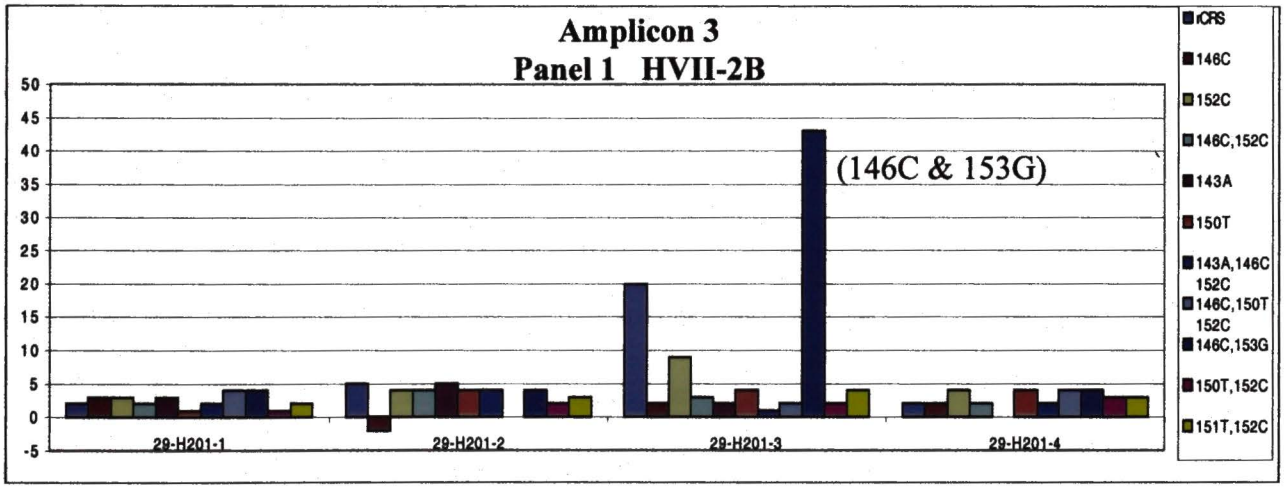

Figure 46. SSO Hybridization results at region HVII-2B for sample HIS.201. The data suggests a haplotype of 146C-153G, the correct haplotype is 153G.

The Luminex LabMAP™ technology may still prove useful in the analysis of the mtDNA polymorphisms. The primer extension assay can be adopted for the Luminex platform.

In conjunction with the ZipCode sequences, the appropriate extension primers can be synthesized. This technology may require the development of new dideoxynucleotides that are labeled with a series of fluorochromes appropriate for the lasers found in the Luminex 100™ instrument. Alternatively, the Luminex 100™ instrument can be redesigned to accommodate the use of the dideoxynucleotides found in the SNapshot™ system.

SNaPshot™ Single Base Extension Assay

The SNaPshot™ Single Base Extension assay utilizes the same chain termination chemistry employed by standard DNA dideoxy-sequencing methods (Sanger, Nicklen et al., 1977). A number of researchers have attempted to exploit the advantages of primer extension-based detection of SNPs in diagnostic testing and have developed various gel-based and solid-phase-based methods (Tully, Sullivan et al., 1996; Planz, Pogue et al., 1999). The method evaluated in this study represents the next stage in the evolution of SNP panels, through the development of flexible virtual SNP arrays that incorporate the precision of single nucleotide extension with high speed processing via capillary electrophoresis. For each locus, the amplified product serves as the DNA template onto which a SNP-specific oligonucleotide primer is hybridized. Unlike the Luminex SSO Hybridization assay, amplification primers are unlabeled and identical to the standard primer sets used for sequencing. In addition, other factors, such as bead concentration, coupling efficiency, and stringent hybridization parameters do not influence the sensitivity or success of the extension reactions. The entire process is more streamlined starting with an initial, standard amplification, followed by an enzymatic cleanup step, ExoSAP-IT, to eliminate amplification primers and would-be competitor dNTPs. This

processing does not require the transfer of amplified products to new tubes and simultaneously prepares the amplicons for sequencing should it be necessary for confirmation.

The SNaPshot™ reaction has a robust range of potential input DNA concentration from a recommended low of 0.01 pmol to 0.4 pmol of template DNA. During this study, a concentration of 0.15 pmols of input DNA was selected, which corresponds to approximately 2.25 ng of template. This was done so that a direct comparison of the performance between of the various extension primers could be assessed. Generating haplotypes using the numerous multiplex reactions developed to cover the HVI and HVII SNPs interrogated by this method utilized approximately 20 ng of the total amplified product generated for a sample. In contrast, standard sequencing reactions for the 4 HVI and HVII subunits performed in both forward and reverse directions requires approximately 500 ng of total amplified product when the recommended range of 40 – 90 ng of PCR product is added to each reaction. The SNaPshot™ primer extension assay is approximately 20 –25 times more sensitive, with regard to low copy number detection, than the standard sequencing approach. This would suggest that this system has the potential to be a valuable alternative to sequence analysis when samples are limited, as well as being more robust in detection and typing of heteroplasmic sites.

Of major concern in forensic applications is the inconsistency found in sequencing and other typing methods to effectively discern heteroplasmic positions. Although Tully et al.'s estimate of 13.8 % reflects a robust measure of the actual heteroplasmy for mtDNA sequences, recent experiences in controlled studies by NIST to evaluate the ability to detect and correctly characterize heteroplasmy in mtDNA sequences, have illuminated the fact that current sequencing chemistries and detection methods often fail at detecting heteroplasmy when the alternate bases are present at low ratios. A major advantage of a SNP-based approach to

mtDNA typing revolves around the individual interrogation of the polymorphic site by specific nucleotides. The presence of two nucleotide reporters at a particular SNP locus definitively identifies and detects the presence of heteroplasmy in a mtDNA sample. The increased sensitivity demonstrated by the SNaPshot™ method would suggest that this method is superior in situations where low levels of heteroplasmy exist.

Data published by the FBI (FBI DNA Analysis Unit II, 2000) indicates that approximately one third of the polymorphic sites employed in the SNP panel have demonstrated heteroplasmy. This data was based on a variety of detection methods. These SNP positions account for approximately 25% of the base positions listed in the FBI document as heteroplasmic sites, exclusive of length heteroplasmies. Additionally, four SNP sites assayed by our panel have heteroplasmic positions reported either between the SNPs tested or adjacent to the base being queried. This suggests that the initial panel alone may be capable of detecting a large fraction of heteroplasmic polymorphisms.

Once the reliability of the extension primer sets was demonstrated, 50 individuals of Southwestern Hispanic origin (that had been previously sequenced for the SWGDAM mtDNA database) were typed with the primer extension cocktails. The evaluation of the testing panel was done in a blind format, in that sequence data for these individuals was not made available until after the SNP typing results were completed. With this SNP panel, 44 percent of the 50 individuals could be differentiated by the array. Twenty-eight haplotypes were shared among two to seven individuals. Sequence data for theses samples including only sites exhibiting nucleotide substitutions provided 45 unique haplotypes with five individuals sharing a common haplotype. These results indicate that the SNP panel resolved approximately 48% of the potential haplotypes contained in this data set. It should be noted that it would not be difficult to

modify the panel so that a higher percentage of haplotypes could be resolved. The addition of one SNP position, 16319, increases the number of unique haplotypes detected by the SNP panel to 37 or approximately 78% of the potential haplotypes in this data set. The remaining 8 individuals were distributed among 3 haplotypes and possessed additional substitutions at sites reporting fairly low levels of polymorphism.

Nucleotide determination at these 34 SNPs was concordant with the results obtained from the FBI in all but two instances. In the first case, the sequence data from the FBI indicated that sample HIS.159 had a C residue at position 146 in HVII. This was recorded in the SWGDAM database, since it differed from the revised Cambridge Reference Sequence at this site. Typing with our extension primer assay for position 146 revealed a T residue, corresponding to the rCRS (**Table 10**). To confirm this finding we sequenced the questioned sample and confirmed that our typing was correct. Subsequently, the SWGDAM database has been corrected for this error.

The second discrepancy in the study concerned a heteroplasmic site detected in sample HIS.213 at position 16362 in HV1. The SWGDAM database for HIS.213 did not report any polymorphism at this position, however the SNaPshot™ extension assay clearly demonstrated the presence of two nucleotides as was previously shown in **Figures 37 and 38**. HIS.213 was sequenced, and the heteroplasmy at this site was clearly evident in our processing of this sample (**Figure 40**).

To determine the feasibility of multiplexing arrays of SNPs using this technology, we prepared various mixtures of extension primers. Each of the extension primers had been previously evaluated in individual reactions. At the time this research was performed, the only SNaPshot™ kit available was designed for individual SNP detection. However, using the existing kit, we were able to successfully multiplex groups of 4 to 5 SNPs without any

observable impediment to the extension reaction. To evaluate the effect of extension primer inhibition and the ability to effectively type SNP rich areas, a pair of multiplex reactions that extensively interrogated SNPs in an eleven bp region of HV2 was developed. Six individual SNPs, 143, 146, 150, 151, 152, and 153 were successfully interrogated in this region **(Figure 47)**. This verified that the methods we used to optimize the extension primers was effective, and that it was feasible to develop arrays in SNP dense regions.

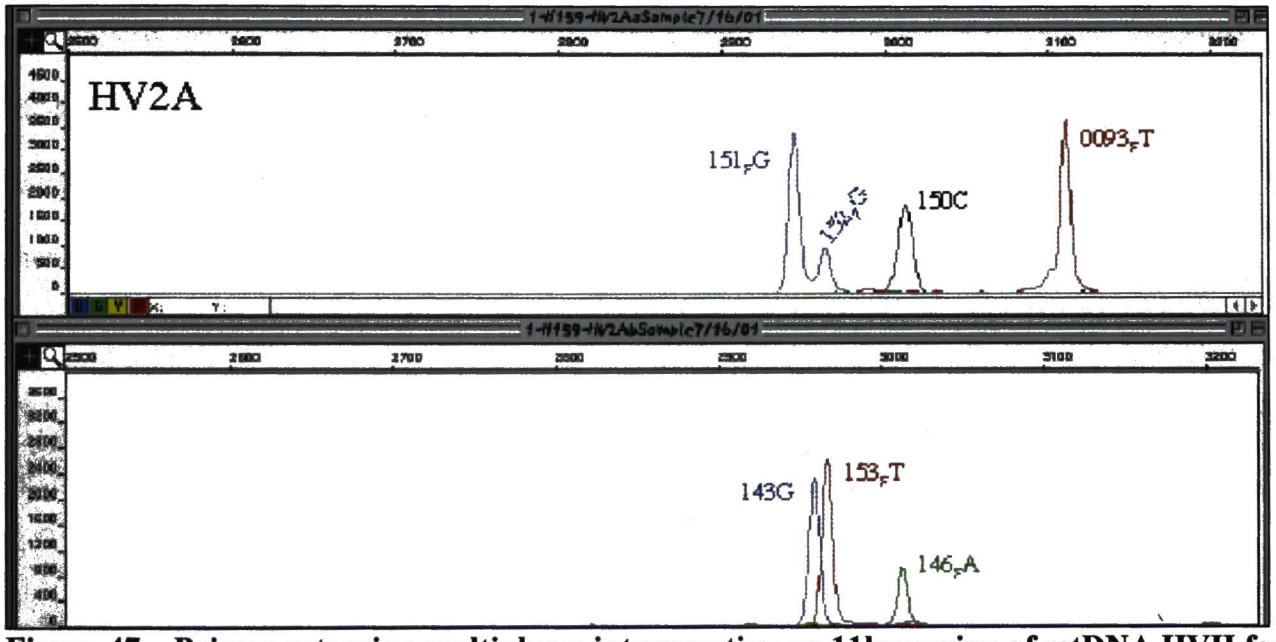

Figure 47. Primer extension multiplexes interrogating an 11bp region of mtDNA HVII for 6 individual SNPs.

In collaboration with R&D scientists at Applied Biosystems, a panel of 15 extension primers was prepared incorporating mobility modifiers in an attempt to more effectively separate the extension products during capillary electrophoresis. **Figure 48** shows the results of our first attempt at employing the new multiplex cocktail. The new multiplex cocktail was used on sample HIS.213, in which heteroplasmy had been previously detected. The electropherogram in **Figure 48** demonstrates that the incorporation of the Applied Biosystems mobility modifiers,

such as those incorporated in their STR kits, greatly enhances the separation between the extension products. These mobility modifiers can be sequentially added to discrete extension primer sequences without potential hairpins, dimers or other primer-template interactions. A collection of mobility modifiers is available from ABI. They can be incorporated into the extension primer, individually or in tandem, to regulate extension product migration to particular positions in an electrophoretic run. The preliminary experiments using the 15-plex cocktail of extension primers was not optimized. Analysis of the electropherogram (**Figure 48**) indicates that the extension reaction parameters need to be evaluated. However, it should be noted that the reaction was performed using the SNaPshot™ kit designated for only a single extension reaction. The use of this early version of the kit may have resulted in reduced nucleotide incorporation and weak signal strength at some of the SNP positions.

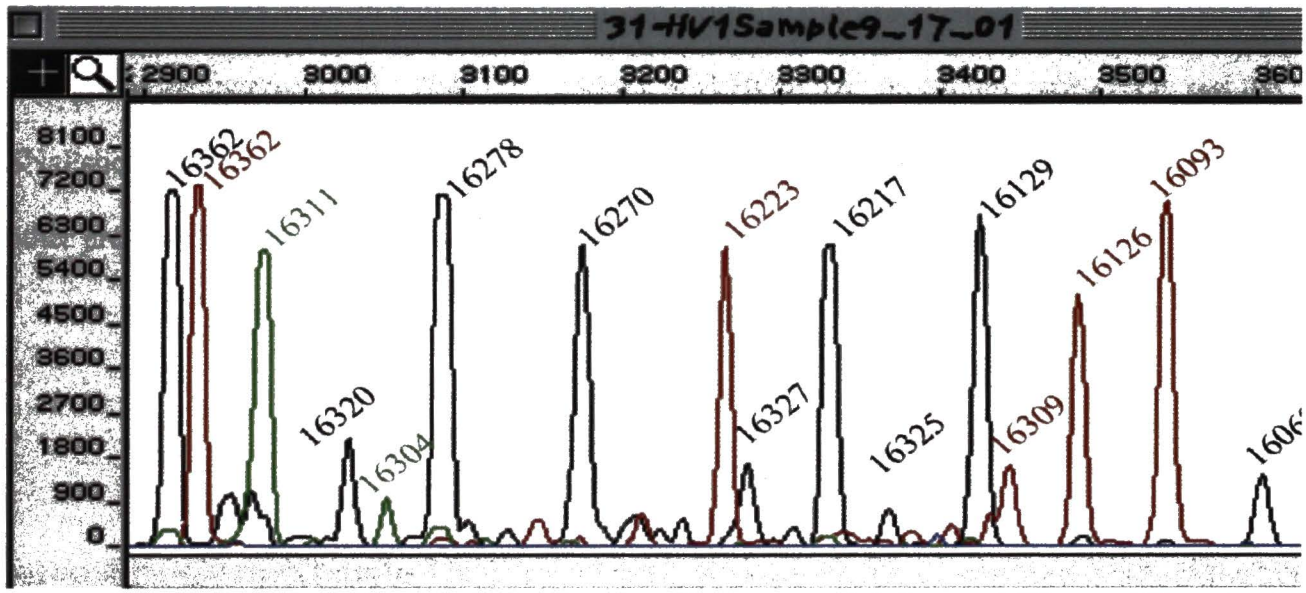

Figure 48. Multiplex of 15 extension primers with some containing mobility modifiers to alter migration during electrophoresis.

Our preliminary investigations into the utility of this method suggest that individual arrays of up to 25 SNP extension primers can be successfully muliplexed for the analysis of HV1

and HVII. Extension primers for coding region SNPs could also be included in future arrays. The simplistic protocols and chemistries employed in the SNaPshot™ Single Base Extension system would allow for the rapid incorporation of this technology into the forensic laboratories. This system has the potential for automation, and could be included in a mtDNA databasing operation or when the mass screening of samples is required.

Statistical Evaluation of the SNP Panels

A major goal in the development of any rapid assay for mtDNA SNP detection is that it approaches the level of resolution obtainable through direct sequencing. Both hybridization-based assays and primer extension assays have previously been applied to mtDNA analysis (Stoneking, Hedgecock et al., 1991; Reynolds, Clark et al., 1996; Tully, Sullivan et al., 1996; Melton, Wilson et al., 1997; Planz, Pogue et al., 1999). The approach investigated in this study address three issues: high-throughput analysis capability, platform robustness, and the statistical strength of the marker panel. Both the Luminex SSO Hybridization capture probe array and the SNaPshot™ primer extension array, interrogated a much larger collection of SNP positions than other previously reported systems, which ranged from 10 to 15 SNPs (**Table 11**).

Table 11. HV1 and HV2 SNP Loci Analyzed in Various Screening Systems.

Melton, Wilson et al., 1997	Tully, Sullivan et al., 1996	Planz, Pogue et al., 1999	31 SNP SSO Hybridization Assay		34 SNP SBE Assay	
16126	16069	16069	16124	73	16069	73
16129	16129	16129	16126	93	16093	93
16217	16189	16191	16129	143	16124	143
16223	16224	16224	16217	146	16126	146
16247	16311	73	16223	150	16129	150
16261	73	146	16224	151	16172	151
16304	146	152	16292	152	16217	152
16311	152	195	16294	153	16223	153
16362	195	199	16295	189	16224	189
73	247	247	16296	195	16292	195
146	302		16298	198	16294	198
152	523		16304	199	16295	199
195			16309	200	16304	200
199			16311	263	16309	247
247			16319		16311	263
			16320		16320	
			16362		16325	
					16327	
					16362	

The most comprehensive set of mtDNA SSO data comes from European, Asian and African populations assayed at 13 SNP positions (Melton and Stoneking, 1996; Melton, Ginther et al., 1997; Melton, Wilson et al., 1997). Measures of genetic diversity (*h*) (Tajima, 1989) provide an estimate of how well a proposed panel of markers can capture the inherent genome

variability among populations. For the 13 SNP panels, *h* ranged from a low of approximately 0.91 in Caucasians to a high of 0.996 in Asians. Unbiased estimates of diversity are tempered by the sample size of the examined populations and with *n* ranging from 381 individuals to 993 individuals, any bias due to sampling error can be easily discounted. The estimates of diversity presented by Melton can be considered reliable for a panel of 13 SNPs. The data presented in this study for a 31 SSO SNP array yield a slightly higher diversity of 0.9559 for the 50 Hispanic individuals tested. This estimate of genetic diversity is most likely an underestimate tempered by the comparatively small sample size ($n = 50$) available for this study. These results, however, are in line with the data reported by Melton (2001) for southwestern Hispanics of $h = 0.938 - 0.969$, $n = 94$ and 60, respectively. An analysis of the maximum genetic diversity obtainable from the population sample tested, using the full sequence data for HVI and HVII, provides a $h = 0.9918$ when only substitutions are evaluated. Inclusion of all sequence variations, substitutions, insertions, and deletions, for this population sample only increases the genetic diversity estimate to 0.9976.

Of forensic interest, is the frequency of unique haplotypes identified by a particular SNP panel. In the 50 Southwestern Hispanics tested with the Luminex SSO array, 30 different haplotypes were observed with 22 haplotypes observed only a single time. When evaluating only substitution data, the sequence data from this sample population includes a total of 46 haplotypes, with 45 haplotypes present only once. This number is only increased by two, to 48 haplotypes when the entire sequence is evaluated. In a larger population of 2426 individuals, the same Luminex SSO Hybridization array, identified a total of 635 haplotypes (Steven Lee, personal communication). Of these, 378 represented unique haplotypes observed only once in the population sample. From this data, a reasonable expectation of the Luminex SSO

Hybridization panel would be to obtain 60 – 70 percent of the sample size in haplotypes with approximately 23 – 30 percent unique haplotypes. This is comparable to the results obtained by Melton and Reynolds.

In forensic analyses the probability that two random samples should match by chance is often of interest. The random match probability (p), described in Equation 2 of the methods section, determined for the Luminex SSO panel in a sample of 50 southwestern Hispanics was determined to be 0.0632. This estimate is comparable to that reported by Tully for British Caucasians ($p = 0.054$). This measure should not however be taken out of context, considering the overall lower diversity of the Hispanic populations when compared to Africans and Asians(Melton and Stoneking, 1996; Melton, Ginther et al., 1997). Estimates of p from a larger ($n = 2426$) population analyzed with the Luminex SSO array (Steven Lee, personal communication) provides a value of 0.012, which is more in line with what has been seen in actual sequence databases from the major population groups (FBI DNA Analysis Unit II, 2000).

Although a direct comparison of the two SNP panels evaluated here cannot be done due to differences in array composition, statistical interpretation of the robustness of the panels is possible. Genetic diversity (h) estimated for the 34 SNP primer extension array was 0.9624 slightly higher than that determined for the Luminex SSO panel. This variation is most likely due to either the increased number of SNPs assayed or individual differences in the array composition, or a combined effect of the two parameters. As discussed earlier with regard to the SSO Hybridization panel, our results are consistent with those reported for other SNP typing systems. Both the SSO Hybridization and Single Base Extension arrays detected 30 haplotypes among the 50 Southwestern Hispanics tested with 22 unique haplotypes. The distribution of the haplotypes differed due to the varying composition of the arrays. The most common haplotype

determined using the primer extension array (16223T, 16362C, 73G, 146C, 153G, 263G) had a frequency of 14 percent. This is slightly less of than the most common SSO haplotype (16223T, 16319A, 16362C, 73G, 146C, 153G, 263G array, which had a frequency of 18 percent. These two haplotypes only differ in position 16319, which was not assayed by the primer extension array.

A forensic mtDNA SNP panel consisting of the positions evaluated in this study provides a reasonable alternative to the full sequencing of the HVI and HVII regions. Expansion of the array to include additional polymorphic sites would enable the panel to assess a level of diversity comparable to that determined through sequence analysis. Similarly, a slight expansion of the testing panels would decrease random match probabilities, especially in the most diverse populations. In populations, such as Southwestern Hispanics, where overall haplotype diversity is reduced, the analysis of HVI and HVII by either sequencing or SNP analysis, will not yield the highest degree of discrimination. In these populations, an auxiliary SNP panel, which would interrogate additional control region positions or coding region SNPs, can be incorporated to provide the necessary resolution.

BIBLIOGRAPHY

AABB (2000). "Annual Report Summary for 2000." Annual Report 2000: 1-11.

Alberts, B., D. Bray, J. Lewis, M. Raff and J. D. Watson (1989). Molecular biology of the cell.

Altman, L. (2001). "Forensic Lab is Prepared for 20,000 DNA Tests." The New York Times: collage4.nytimes.com/guests/articiles/2001/09/14/867960.xml.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden and I. G. Young (1981). "Sequence and Organization of the Human Mitochondrial genome." Nature **290**: 457-65.

Andrews, R. M., I. Kubacka, P. F. Chinnery, R. N. Lightowlers, D. M. Turnbull and N. Howell (1999). "Reanalysis and Revision of the Cambridge Reference Sequence for Human Mitochondrial DNA." Nature Genetics **23**(2): 147.

Applied Biosystems (1999). "ABI Prism® SNapshot™ ddNTP Primer Extension Kit." ABI Prism® SNapshot™ ddNTP Primer Extension Kit Manual.

Avery, O. T., C. M. Macleod and M. McCarty (1944). "Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types." J Exp Med **79**(2): 137-58.

Bendall, K. E. and B. C. Sykes (1995). "Length Heteroplasmy in the First Hypervariable Segment of the Human Mitochondrial mtDNA Control Region." Am J Hum Genet **57**: 248-56.

Bogenhagen, D. and D. A. Clayton (1980). "The number of mitochondrial deoxyribonucleic acid genomes in mouse L and HeLa cells: quantitative isolation of mitochondrial deoxyribonucleic acid." J Biol Chem **249**: 7991-5.

Botstein, D., R. White, M. Skolnick and R. W. Davis (1980). "Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms." Am J Hum Genet **32**(3): 314-31.

Brown, W. M., M. George and A. C. Wilson (1979). "Rapid Evolution of Animal Mitochondrial DNA." Proc Nat Acad Sci USA **76**: 1967-71.

Butler, J. M. (1999). "STR Analysis by Time-of-Flight Mass Spectrometry." Profiles in DNA **2**: 4-7.

Butler, J. M. and B. C. Levin (1998). "Forensic Applications of Mitochondrial DNA." Tibtect **16**(158-62).

Butterfield, F. (1996). DNA tests expected to get wider use. News Observer. Raleigh, NC: p. A9.

Cann, R. L., M. Stoneking and A. C. Wilson (1987). "Mitochondrial DNA and Human Evolution." Nature **325**: 31-6.

Conners, E., T. Lundregan, N. Miller and T. McEwen (1996). Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial.

Daoudi, Y., M. Morgan, C. Diefenbach, J. Ryan, T. Johnson, G. Conklin, K. Duncan, K. Smigielski, E. Huffine, D. Rankin, R. Mann, T. Holland, K. Mcelfresh, J. Canik, V. Ambrustmacher and M. Holland (1998). "Identification of the Vietnam Tomb of the Unknown Soldier, the Many Roles of Mitochondrial DNA." Nineth International Symposium on Human Identification: http://www.promega.com/geneticidproc/ussymp9proc/abstracts.html.

Eisenberg, A. J. (2000). "Microsatellite DNA Analysis." UNT Health Science Center DNA Identity Lab Manual: 15-30.

Erickson, D. (1991). "Do DNA fingerprints protect the innocent?" Sci Amer **265**(2): 18.

FBI DNA Analysis Unit II (2000). Scientific Oral Study Guide 11/16/00.

FBI DNA Analysis Unit II (2001). Mitochondrial DNA Sequencing Protocal 5/2001, Federal Bureau of Investigation, Washington DC: 79.

Fischer, C., A. R. Isenberg, J. E. B. Stewart, K. W. P. Miller, C. Theisen, M. R. Wilson, J. DiZinno and B. Budowle (2000). "Mitochondrial DNA: Today and Tomorrow." Eleventh International Symposoium on Human Identification: www.promega.com/geneticidproc/ussymp11proc/content/fisher.pdf.

Fulton, R. J., R. L. McDade, P. L. Smith, L. J. Kienker and J. R. Kettman (1997). "Advanced multiplexed analysis with the FlowMetrix system." Clin Chem **43**: 1749-1756.

Gabriel, M. N., M. N. Calloway, R. L. Reynolds, S. Andelinovic and D. Primorac (2001). "Population Variation of Human Mitochondrial DNA Hypervariable Regions I and II in 105 Croatian Individuals Immobilized Sequence-specific Oligonucleotide Probe Analysis." CMJ **42**(3): 328-35.

Giles, R. E., H. Blanc, H. M. Cann and D. C. Wallace (1980). "Maternal Inheritance of Human Mitochondrial DNA." Proc Natl Acad Sci U.S.A. **77**: 6715-9.

Gill, P., P. L. Ivanov, C. Kimpton, R. Piercy, N. Benson, G. Tully, I. Evett, E. Hagelberg and K. Sullivan (1994). "Identification of the Remains of the Romanov Family by DNA Analysis." Nature Genetics **6**: 130-5.

Ginther, C., L. Issel-Tarver and M. C. King (1992). "Identifiying Individuals by Sequencing Mitochondrial DNA from Teeth." Nature Genetics **2**: 135-8.

Gray, M. W. (1992). "The Endosymbiont Hypothesis Revisited." Int Rev Cytol **141**: 233-57.

Greenberg, B. D., J. E. Newbold and A. Sugino (1983). "Intraspecific Nucleotide Sequence Variability Surrounding the Origin of Replication in Human mitochondrial DNA." Gene **21**: 33-49.

Grimberg, J., S. Nawoschic, L. Bellvscio, R. McKee, A. Turck and A. J. Eisenberg (1989). "A Simple and Efficient Nonorganic Procedure for the Isolation of Genomic DNA from Blood." Nucl Acids Res **17**(20): 8390.

Holland, M., D. L. Fischer, L. G. Mitchel, W. C. Rodriquex, J. J. Canik, C. R. Merril and V. W. Weedn (1993). "Mitochondrial DNA Sequence Analysis of Human Skeletal Remains: Identification of Remains from the Vietnam war." J Forensic Sci **38**: 542-53.

Holland, M. M. and T. J. Parson (1999). "Mitchondrial DNA Sequence Analysis Validation and Use for Foresic Casework." Forensic Science Review **11-1**: 21-50.

Holland, M. M., R. K. Roby, J. J. Canik and V. W. Weedn (1995). "An Update on the Military's Program of Skeletal Remains Identification Using Mitochondrial DNA Sequence Analysis: Identification of Soliders Killed During the Vietnam War, Korean War and World War II." Sixth Internaltional Symposium of Human Identification-1995: www.promega.com/geneticidproc/usssymp6proc/holland.htm.

Howell, N. (1992). "Mitochondrial Gene Segregation in Mammals: Is the Bottleneck Always Narrow?" Hum Genet **90**: 117-20.

Iannone, M. A., J. D. Taylor, J. Chen, S. Li, P. Rivers, K. A. Slentz-Kesler and M. P. Weiner (2000). "Multiplexed Single Nucleotide Polymorphism Genotyping by Oligonucleotide Ligation and Flow Cytometry." Cytometry **39**: 131-40.

Isenberg, A. R. and J. M. Moorem (1999). "Mitochondrial DNA analysis at the FBI Laboratory." Forensic Science Communication 1999: http://www.fbi.gov/programs/lab/fsc/backissu/july1999/dnatext.htm.

Ivanov, P. L., M. J. Wadhams, R. K. Roby, M. Holland, v. W. Weedn and T. Parsons, J. (1996). "Mitochondrial DNA Sequence Heteroplasmy in the Grand Duke of Russia GeorgiJ Romanov Establishes the Authenticity of the Remains of Tsar Nicholas II." Nature Genetics **12**: 417-20.

Jacobs, K. A., R. Rudersfor, S. D. Neill, J. P. Dougherty, E. L. Brown and E. F. Fritsch (1988). "The Thermal Stability of Oligonucleotide Duplexes is Sequence Independent in Tetraalkylammonium Salt Solution: Application to Identifying Recombinant DNA clones." Nucl Acids Res **16**: 4637-50.

Jeffreys, A. J., J. F. Y. Brookfield and R. Semenoff (1986). "Positive Identification of an Immigration Case Using Human DNA Fingerprints." Nature **317**: 818-19.

Jeffreys, A. J., V. Wilson and S. L. Thein (1985). "Hypervariable Minisatellite Regions in Human DNA." Nature **314**(67-73).

Jeffreys, A. J., V. Wilson and S. L. Thein (1985). "Individual Specific 'fingerprints' of Human DNA." Nature **316**: 76-79.

Jenkins, S. (2001). The quest for 6,4347 Identities. Washingtonpost.com. Washington D.C.: www.washingtonpost.com/ac2/wp-dyn/a32203-2001Spet26?language=printer.

Krings, M., A. Stone, W. Schmitz, H. Krainitzki, M. Stoneking and S. Paabo (1997). "Neanderthal DNA Sequences and the Origin of Modern Humans." Cell **90**: 19-30.

Lonsway, K. A. (2000). "Successfully Investigating Acquaintance Sexual Assult: A National Training Manual for Law Enforcement." DNA Evidence and Issues: 5-40.

Luminex Corporation, A. T. (2000). Luminex 100 ™ User's Manual Version 1.7**:** 1-97.

Maniatis, T., E. F. Fritsch and J. Sambrook (1989). "Molecular Cloning: Laboratory Manual." Cold Spring Harbor Laboratory.

Melton, T., C. Ginther, G. Sensabaugh, H. Soodyall and M. Stoneking (1997). "Extent of Heterogeneity in Mitochondrial DNA of Sub-Saharan African Populations." J Forensic Sci **42**: 582-92.

Melton, T. and K. Nelson (2001). "Forensic Mitochondrial DNA Analysis: Two years of Commercial Casework Experience in the United States." Croatian Medical Journal **42**(3): 298-303.

Melton, T. and M. Stoneking (1996). "Extent of Heterogenetiy in Mitochondrial DNA of Ethnic Asian Populations." J Forensic Sci **41**: 591-602.

Melton, T., M. Wilson, M. Batzer and M. Stoneking (1997). "Extent of Heterogenetiy in Mitochondrial DNA of European populations." J Forensic Sci **42**: 437-446.

Michaels, G. S., W. W. Hauswirth and P. J. Laipis (1982). "DNA Copy Number in Bovine Oocytes and Somatic cells." Dev Biol **94**: 246-51.

Momoi, M. (1993). "Mitochondrial DNA Mutations and Disease." Physico-Chemical Biology **37**: 345-51.

Monnat, R. J. and L. A. Loeb (1985). "Nucleotide Sequence Preservation of Human Mitochondrial DNA." Proc Natl Acad Sci USA **82**: 2895-9.

Mullis, K. B., F. Falloona, S. J. Scharf, R. K. Saiki, G. T. Horn and H. Erlich (1986). "Specific Enzymatic Amplification of DNA in vitro: the Polyermase Chain Reaction." Cold Spring Harbor Symp Quant Biol **51**: 263-73.

Nikiforov, T. T., R. B. Rendle, P. Goelet, Y. Rogers, M. L. Kotewicz, S. Anderson, G. L. Trainor and M. R. Knapp (1994). "Genetic Bit Analysis: A Solid Phase Method for Typing Single Nucleotide Polymorphism." Nuc Acids Res **22**: 4167-75.

Parson, W., T. J. Parson, R. Scheithauer and M. M. Holland (1998). "Population Data for 101 Austrian Caucasian Mitochondrial D-loop Sequence: Application of mtDNA Sequence Analysis to a Forensic Case." J Legal Med **112**: 58-61.

Piko, L. and L. Matsumoto (1976). "Number of Mitochondria and some Properites of Mitochondrial DNA in the Mouse Egg." Dev Biol **49**: 1-10.

Planz, J. (1999). Map of Human Mitochondrial Genome.

Planz, J. (2002). Estimation of mtDNA Analyzed per Month Based on 3100 ABI Genetic Analyzer automation.

Planz, J. V., V. M. Fish, J. M. Rader, M. W. DuPont, J. E. Warren, P. Gill and R. C. Giles (1997). "Development and Validation of a GBA™-based mtDNA Sequence Evaluation System for Forensic Testing." Proceedings from the Eight International Sympiosium on Human Identification.: www.promega.com/geneticidproc/ussymp8proc/30.html, 107-111.

Planz, J. V., P. Pogue, A. L. Sumts and M. Castro (1999). "Diversity of Microsatellite Loci and Mitochondrial DNA Haplotypes in a Population from the Peruvian Andes." Proceedings from the Nineth International Sympiosium on Human Identification.: www.promega.com/geneticidproc/ussymp9proc/ab05.pdf.

Reynolds, R., B. Clark, T. Marschak, J. Valaro and S. Rubin (1996). "Mitochondrial DNA Typing Using Sequence-Specific Oligonucletide Probes." Genetics **48**: 137-44.

Robin, E. D. and R. Wong (1988). "Mitochondrial DNA Molecules and Virtual Number of Mitochondrial per Cell in Mammalian Cells." J Cell phys **136**: 507-13.

Sanger, F., S. Nicklen and C. R. Coulson (1977). "DNA Sequencing with Chain-Terminating Inhibitors." Proc Nat Acad Sci USA **74**: 5463-68.

Scheffler, I. E. (1999). Mitochondria, Wiley-Liss, New York.

Scruggs, R., S. Zipperstein, R. Lyon, V. Gonzalez, H. Cousins and R. Beverly (1993). Report to the Deputy Attorney General on the Events at Waco, Texas. February 28 to April 19, 1993, Department of Justice.

Southern, E. M. (1975). "Detection of Specific Sequences Among DNA Fragments Separated by Gel Electrophoresis." J Mol Biol **98**: 503-17.

Stoneking, M., D. Hedgecock, R. G. Higuchi, L. Vigilant and H. Erlich (1991). "Population Variation of Human mtDNA Control Region Sequences Detected by Enzymatic Amplification and Sequence Specific Oligonucleotide Probes." Am J Hum Genet **48**: 370-82.

Sullivan, K., R. Hopgood and P. Gill (1992). "Identification of Human Remains by Amplification and Automated Sequencing of Mitochondrial DNA." Int. J. Legal Med. **105**: 83-86.

Sullivan, K. M., G. Tully, R. Alliston-Greiner and A. Hopwood (1995). A Two Stage Strategy for the Automated Analysis of Mitochondrial DNA. Procedings of the International Society of Forensic Haemogenetics, Berlin, Springer-Verlag.

Tajima, F. (1989). "Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism." Genetics **123**: 585-95.

Taylor, J. D., D. Briley, Q. Nguyen, K. Long, M. A. Iannone, S. Li, F. Ye, A. Afshari, E. Lai, M. Wagner, J. Chen and M. P. Weiner (2001). "Flow Cytometric Platform for High-Throughput Single Nucleotide Polymorphism Analysis." BioTechniques **30**: 661-9.

Tully, G., K. M. Sullivan, P. Nixon, R. E. Stones and P. Gill (1996). "Rapid Detection of Mitchondrial Sequence Polymorphism using Multiplex Solid-phase Fluorescent Minisequencing." Genomics **34**: 107-13.

Venter, J. G. (2001). "The Sequence of the Human Genome." Science **291**: 1304-51.

Wallace, R. B. (1992). "Diseases of the Mitochondrial DNA." Annu Rev Bichem **61**: 1175.

Wallace, R. B., J. Shaffer, R. Murphy, J. Bonner, T. Hirose and K. Itakura (1979). "Hybridization of Synthetic Oligodeoxyribonucleotides to phi chi 174 DNA; the effect of Single Base Pair Mismatch." Nucleic Acids Res. **6**: 3543-57.

Wambaugh, J. (1989). The Blooding, Bantam.

Watson, J. D. and F. H. C. Crick (1953). "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid." Nature **171**: 737-38.

Weedn, V. W. and J. W. Hicks (1997). "The Unrealized Potential of DNA testing." National Institue of Justice Journal: 16-23.

Williams, D. (1998). "The Bombing of the World Trade Center in New York." Inter Crim Pol Rev **1998**: 469-71.

Wilson, M. C., D. J. A., D. Polanskey, J. Replogle and B. Budowle (1995). "Validation of Mitochondrial DNA Sequencing for Forensic Case Work Analysis." Int J Legal Med **108**(2): 68-74.

Wilson, M. R., D. Polanskey, J. M. Butler, J. DiZinno, J. Repogle and B. Budowle (1995). "Extraction, PCR Amplification and Sequencing of Mitochondrial DNA from Human Hair Shafts." Biotechniques **18**: 662-9.

Wood, I. W., J. Gitschier, L. A. Lasky and R. M. Lawn (1985). "Base Composition Independent Hybridization in tetramethylammonium chloride: A method for Oligonucleotide Screening of Highly Complex Gene Libraries." Proc Natl Acad Sci USA **82**: 1585-8.

Wyman, A. R. and R. White (1980). "A Highly Polymorphic Locus in Human DNA." Proc Natl Acad Sci U.S.A. **77**: 6754-58.