
UNIVERSITY OF NORTH TEXAS
HEALTH SCIENCE CENTER AT FORT WORTH
1970

Ariyo, Bolanle. Amplification of Mitochondrial DNA Regions HVI and HVII in its Entirety and Reducing Cycle Sequencing Reactions. Master of Science (Forensic Genetics), August 2004, 46 pages, 10 figures, 7 tables, 18 references.

Mitochondrial DNA is widely used in the forensic community because its of high copy number in cells, location, and mode of inheritance. Yet this method of analysis is expensive, time consuming, and labor intensive, therefore labs should take steps to improve the procedure of mtDNA analysis. This study is performed to validate the use of amplifying HVI and HVII region in its entirety (2 primer sets) for use in reference samples. Amplification performed using primers F15989-R16410 (HVI) and F73- R340 (HVII). The current method of amplification is 4 primer sets at full cycle sequencing reactions. The cost of Cycle Sequencing Kit is also expensive, therefore performing half and quarter reactions would be beneficial in reducing the amount of kit consumed. To validate the use of reducing cycle sequencing reactions, half and quarter cycle reactions were performed using 2 and 4 primer sets. Results demonstrate that sequence data for reducing cycle sequence data is consistent with the sequence data using the current method. Results also show that sequence data obtained using two primer sets was consistent with sequence data amplified by the current method with the exception of two samples at length heteroplasmy polycytosine regions.

AMPLIFICATION OF MITOCHONDRIAL DNA REGIONS HVI AND HVII IN ITS ENTIRETY AND REDUCING CYCLE SEQUENCING

Bolanle Ariyo, B.S.

Major Professor

Committee Member

Committee Member

University Member

Department Chair

Graduate Dean

AMPLIFICATION OF MITOCHONDRIAL DNA REGIONS HVI AND HVII IN ITS ENTIRETY AND REDUCING CYCLE SEQUENCING

INTERNSHIP PRACTICUM REPORT

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

MASTER OF SCIENCE

By

Bolanle Ariyo, B.S.

Fort Worth, Texas

August 2004

ACKOWLEDGEMENTS

I would like to thank Orchid Cellmark, Dallas, TX for providing the opportunity to work in their laboratory and on this project. I would like to thank Kristina Paulette, Cassie Johnson, and Joe Warren for all of their time, guidance, and patience on this project. I learned a great deal from my experience at Orchid Cellmark, Dallas, TX.

I would also like to thank my committee members, Dr. John Planz and Dr. Arthur Eisenberg for their help and supervision. Lastly, I would like to thank Dr. Joseph Warren, my major professor, for his time, help, and guidance through this journey.

For all the people that showed support, friends, family , and associates, I owe my appreciation. Last, but not least, I would like to thank my Lord and Savior, Jesus Christ , because without him all of this would not have been possible.

TABLE OF CONTENTS

List of Tables

Page

List of Figures

Page

CHAPTER 1

Introduction

Mitochondrial DNA analysis provides an alternative method to analyze samples that are not suitable for nuclear DNA analysis such as hairs, degraded tissue, and skeletal remains. Analyzing sequence variation in the hypervariable regions of the human mitochondrial DNA (mtDNA) genome is widely used in the forensic community because of high copy number in cells, location, and mode of inheritance of the mitochondria. MtDNA has been used in other fields such population genetics for over twenty years, pathology, medical and anthropological studies. Currently, the FBI lists seven laboratories in the US as being online for mtDNA analysis for forensic testing (9). This number is steadily growing and does not include private labs that perform this type of testing. A draw back of mtDNA analysis is that the procedure is time-consuming, labor-intensive and expensive process. For this reason, measures must be taken to reduce these limitations and improve mtDNA analysis. To increase through-put time and reduce costs, the mitochondrial DNA Lab at Orchid Cellmark, Dallas, TX seeks to validate two protocol changes; the first is to amplify the HVI and HVII mtDNA region in its entirety using 2 primer sets F15989/R16410 (HVI) and F15/R389 (HVII) for reference samples. The second validation study is the use of half and quarter cycle sequencing reaction mix. The cost of Cycle Sequencing Kit is expensive. Therefore, performing half and quarter

reactions would be valuable in reducing the amount of kits used. The mitochondrial DNA analysis section currently consumes one kit per month.

National Standards set by DNA Advisory Board (DAB) requires that forensic laboratories validate new procedures, equipment and reagents before its use in the laboratory. In 1998 these standards stated that "Material modifications made to analytical procedures shall be documented and subject to validation testing" (8.1.3.4) (1). After the new method has been validated in the scientific community, then an internal validation must take place within the laboratory utilizing the new method. The internal validation should be performed to determine the "reliability of the procedure in-house", to include testing the system with known samples, comparing the procedure to the original procedure if significant changes have been made, ensuring the testing procedure does not introduce contamination (6).

Current Amplification Method

The mitochondrial DNA Lab at Orchid Cellmark, Dallas, TX currently amplifies reference and evidentiary samples using 4 primer sets F15989/R16251, F16190/R16410, F15/R285, and F155/R389. These samples are cycle sequenced using full cycle sequence reactions. Each sample is sequenced twice according to their guidelines that there must be two concordant sequences to utilize the sequence.

Background

Mitochondrion is a double membrane organelle in the cytoplasm of the cell. This organelle is the site where oxidative phosphorylation occurs. Mitochondrion is often called the energy powerhouse of the cell. Inside the matrixes of the mitochondria are structures called "nucleoids", this is where the mitochondrial DNA resides. Mitochondrial DNA is circular double-stranded DNA that encodes mitochondrial transfer RNAs, ribosomal RNAs, and proteins. Human mitochondrial DNA genome is 16,569 bp that encodes 37 genes (6). Twenty-two genes code for transfer RNAs, two encode ribosomal RNA, and 13 encode for protein enzymes. Several copies of the mitochondrial genome are found in each organelle. Each nucleoid may contain 4-5 copies of the mitochondrial DNA (mtDNA).

There are 2 regions of the mtDNA, the coding region and non-coding region. The region that is of interest in forensics is the control region or D-Loop (this name is for the structure this region makes during replication); this section is non-coding. The control region is highly variable in comparison to the coded region, and is 1125 base pairs in length. Two hypervariable regions, HVI and HVII are located within the control region. It is these regions that are primarily analyzed in forensic testing. The HVI region extends from base16024-16365 and HVII region extends from the bases 73-340. This numbering system is derived from the original human mtDNA genome sequence obtained by Anderson et al in 1981(2).

The mode of inheritance is maternal without recombination because sperm contributes essentially no mtDNA to the fertilized egg (16). This means an offspring will have the same sequence

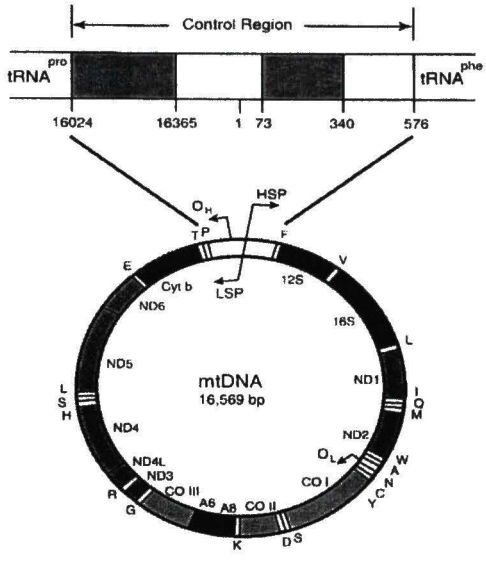


Figure 1 Map of the human mitochondrial genome and expanded diagram of non-coding control region (6).

as the mother with the exception of mutation. Consequently, mtDNA can not be used to uniquely identify an individual because individuals that share the same maternal lineage will have the same the mtDNA sequence.

Several steps are involved before a sequence can be analyzed. They include DNA extraction, PCR amplification, quantification of the amplified DNA, purification of the amplified DNA, cycle sequencing, and data analysis (8). DNA is first extracted from

sample with the proper protocol required for the sample to remove biological molecules such a proteins and cellular material. DNA extract is then amplified using a process called Polymerase Chain Reaction (PCR). PCR produces millions of copies of specific regions on the DNA that are of interest. To quantify the DNA, the amplified product is examined on a gel or capillary electrophoresis. A comparison is made between the amplified DNA and a known DNA standard (ladder) to determine the concentration of the amplified product. After the DNA quantity has been determined, it is purified to get rid unincorporated dNTP's, primers, and excess reagents. The samples are then cycle sequenced using the Sanger method, first devised in 1975. The Sanger method, dideoxy terminator method, is similar to PCR, but uses different reagents. In the presence of salts, buffer, deoxynucleotides, and dideoxynucleotides the primer is extended by DNA polymerase (13). This technique utilizes 2', 3'-dideoxynucleotide triphospates (ddNTPs), molecules that differ from deoxynucleotides by the having a hydrogen atom attached to the 3' carbon rather than an OH group. These molecules terminate DNA chain elongation because they cannot form a phosphodiester bond with the next deoxynucleotide. The dideoxynucleotides carry a fluorescent dye that is detected by an automated instrument and compete with the deoxynucleotides to be incorporated into the growing strand. This results in DNA fragments that differ in size by one base that has a fluorescent label. The different dye colors identify the base present at the end of each strand, which provides a method for determining the sequence of the entire strand. Once the cycle sequencing procedure is completed, then the samples are purified again to remove extraneous reagents. The samples are then placed on an automated fluorescence based sequencer to

obtain sequence data and analyzed by editing software such as Base Spectrum™ and Sequencher™. The samples are compared to a standard reference sequence to observe polymorphisms. In forensic analysis the mtDNA sequences between the

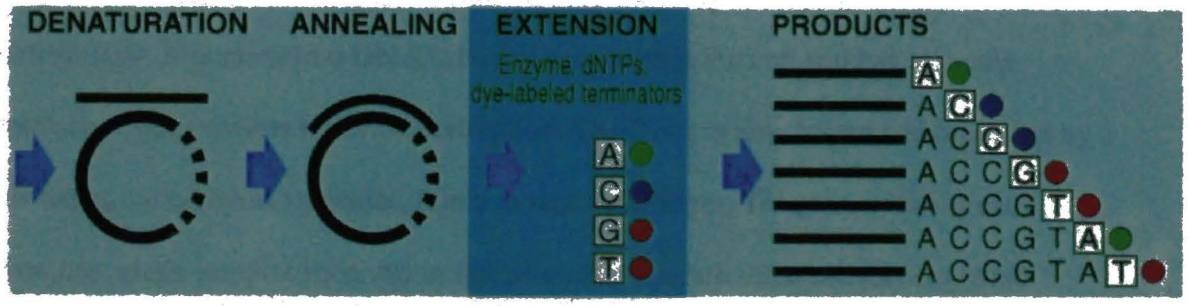


Figure 2 Dye Terminator Cycle Sequencing (Planz J, Forensic mitochondrial sequence analysis Module 3.1)

Reference and evidence samples are compared in order to discern any differences or similarities between them.

The primary sequence of the entire mtDNA genome was sequenced in 1981 by Anderson et al and is often referred to as the Anderson or Cambridge sequence. It is used as a standard reference sequence, by which other human sequences are compared. When there is a difference between an individual's sequence and the reference it sometimes called a "polymorphism with respect to the Anderson sequence" (13). The problem with the Anderson sequence is that when formulated it consisted of Caucasian samples. Therefore people from European descent will not have many differences from the Anderson sequence while others may. A new standard reference sequence may be desirable to be representative of the current population.

MtDNA has advantages over nuclear DNA for certain forensic samples. There may be hundreds of copies of mtDNA in each cell while nuclear DNA contains two copies of DNA. Therefore, mtDNA has a greater copy number than nuclear DNA. The location and structure of mtDNA protect it from degradation when exposed to the environment. Mitochondrial DNA is buried deep within the cell and has a circular structure which protects it from deterioration. The DNA is also bound and protected by a substance, called hydroxyapatite, found in teeth and bones. This serves as a great advantage since samples received in forensic casework are often old and degraded therefore full profiles can not be produced from nuclear DNA. It is most often used where the only available specimens are bone, teeth, hair shafts or decomposed tissue. Since mtDNA has a high copy number of DNA there is a better chance of locating and amplifying a piece of DNA that is not degraded in a sample. Secondly, because of the maternal inheritance it is beneficial in missing person cases, in which where reference samples are only available from maternal relatives. By DNA fingerprinting standards, the identification power of mtDNA is moderate: on average there is approximately a one percent chance that two random individuals will match in mtDNA sequence (6). One disadvantage that mtDNA has compared to nuclear DNA is that the genome is less polymorphic because over ninety percent of the genome is coding, although the non-coding region is highly polymorphic.

Currently this procedure is used in forensic casework, paternity, and missing person cases, along with the other fields mentioned. Common samples used are blood, bone, teeth, hair and ancient samples. MtDNA has been accepted in Federal court as well

as twenty-six state courts and the number for admissibility hearings has diminished substantially (11).

MtDNA analysis has been utilized in many fields other than forensic genetics. The identification of ancient remains, population genetics, evolutionary research, and medical studies dealing with diseases are some fields in the scientific community that deal with mtDNA. Mitochondrial DNA has been used to identify the remains of Tsar Nicolas and the Romanov family, identify skeletal remains in mass graves, and the bones of soldiers from the Vietnam and Korean Wars (10).

Heteroplasmy

Heteroplasmy is the condition where more than a single mtDNA sequence is present in an individual. Generally, heteroplasmic mtDNA sequences differ at a single site within an individual. Heteroplasmy at two or more sites is expected to occur at much lower levels. Heteroplasmy may be observed in several ways including but not limited to: 1) individuals may have more than one mtDNA type in a single tissue; 2) individuals may exhibit one mtDNA type in one tissue and a different type in another tissue; and 3) individuals may be heteroplasmic in one tissue sample and homoplasmic in another tissue sample (15). There are two types of heteroplasmy, sequence and length heteroplasmy. Sequence heteroplasmy results in more than one base at a site in mtDNA sequence. The level of heteroplasmy may not be the same in different tissues (3). Length heteroplasmy is observed when two or more length variants are found within an individual. Length

heteroplasmy is most often seen as multiple lengths of homopolymeric tracts at a particular site. Within and between cells or tissues, homopolymeric cytosine stretches (C-stretches) of different lengths can occur. Length heteroplasmy arises when the two sequences differ in the number of bases in a homopolymeric region and yields a characteristic out-of-phase pattern downstream from the point of heteroplasmy (11). Common areas where length heteroplasmy occur are in the HVI (16184-16193) and HV2 (303-310) regions. In HVI (16184-16193) heteroplasmy occurs when a transition occurs at position 16189 (T-C). This mechanism suggested for generating length heteroplasmy is replication slippage (17). Heteroplasmy can complicate the interpretation of mtDNA results by misinterpretation of the data by the analyst. Heteroplasmy can also be beneficial, by its presence at identical sites which can improve the probability of a match.

Limitations of mtDNA analysis

The main concern when dealing with mitochondrial DNA is contamination because of the sensitivity of mtDNA amplification. There are many steps in mtDNA procedure, as stated previously, which leaves room for contamination to ensue. Steps are taken to ensure that samples are free of extraneous DNA, such as a separate pre-PCR and post-PCR lab, reagent blanks, negative controls, disposable gloves, lab coats and 10 % bleach to decontaminate work surfaces. Separate blower hoods are used when extracting samples and performing PCR. Gloved hands are cleaned with bleach before opening each tube. Lab design consists of separate pre-PCR and post-PCR labs; therefore no amplified

DNA is in contact with non-amplified DNA. Controls are used to monitor the system. Reagent blank controls are samples that are prepared containing the same reagents as evidentiary samples except no biological material is added. This control is used to show that reagents are performing optimally and are not contaminated with human DNA. Contamination if present would be observed in the quantification step or during data analysis. If this control (reagent blank) is contaminated the samples must be re-extracted, if sufficient sample is available. Negative control is processed during the PCR amplification step to observe if any contamination has taken place during the amplification procedure. Once again contamination would be observed during the quantification or when analyzing sequence data. Positive controls are used in amplification and sequencing step to monitor the success of amplification and sequencing (12). The samples are invalid if the positive control does not perform successfully. Cross linking water and tubes before use and using ten percent bleach to wipe gloved hands after opening each tube reduces the risk of contamination. While taking these precautions, contamination may still be prevalent because of the high sensitivity, therefore good lab practices must be maintained.

CHAPTER 2

Research Design and Methodology

There are 6 objectives to this study.

Study 1

DNA extracted from reference samples (analysts in the lab). Using these samples amplify HVI and HVII region in its entirety using primers F15989/16410, and F15/R389. Samples used are sequences that have been determined previously in the lab to compare if this amplification procedure produces the same profile as when amplified full primer sets and full cycle sequencing reaction. Full reactions were used during cycle sequencing.

Study 2

To validate the use of half cycle sequencing reactions, samples were amplified using two primer sets F15989/16410, and F15/R389. Samples used are sequences that have been previously sequenced in the mtDNA lab using full primer sets and full cycle sequencing reaction.

Study 3

To validate the use of quarter cycle sequencing reactions, samples were amplified using two primers sets F15989/16410, and F15/R389. Samples used are sequences that have been previously sequenced in the mtDNA lab using full primer sets and full cycle sequencing reaction.

Study 4

To validate the use of half cycle sequencing reactions for mtDNA casework on samples amplified by using four primer sets, F15989/R16251, F16190/R1640, F15/R285, and F155/R389. Samples used are sequences that have been previously sequenced in the mtDNA lab using full primer sets and full cycle sequencing reaction.

Study 5

To validate the use of quarter cycle sequencing reactions for mtDNA casework on samples amplified by using four primer sets, F15989/R16251, F16190/R1640, F15/R285, and F155/R389. Samples used should have been previously sequenced in the mtDNA lab using full primer sets and full cycle sequencing reaction.

Instrumentation and Experimental Design

DNA Extraction

The Qiagen Dneasy Tissue Kit™(Cat. No. 69506) uses a silica- gel membrane column with selective DNA binding properties. Cells are initially lysed with proteinase K and the products are loaded onto a Qiagen column. DNA adsorbs to the silica membrane in the

presence of high salt while contaminants are passed through the column (12). The centrifugation step removes contaminants and enzyme inhibitors while DNA is selectively bound to the column matrix. The bound DNA is eluted in buffer.

DNA Quanitation

Invitrogen E-Gel® Agarose gels(Cat. No. G7008-02) are self-contained, pre-cast agarose, ethidium bromide, and electrodes packaged inside a dry, disposable, UV-transparent cassette (12). Each E-gel® 96 contains 96 samples lanes and 8 markers lanes. The concentration of ethidium bromide in each gel ranges from 0.1 to .5μg/ml.

Post- Amplification PCR Purification

Princeton Separation PSΨ Clone PCR 96 kit (Cat. No. PC-501) designed for purification of amplification products from PCR reactions. Protocols are fast and result in high yield, high purity DNA (12). DNA fragments bind to the membrane in the plate in the kit. Subsequent wash and centrifugation steps remove residual primers, nucleotides, enzymes, and salts. DNA is eluted in a low salt buffer at high concentrations (12).

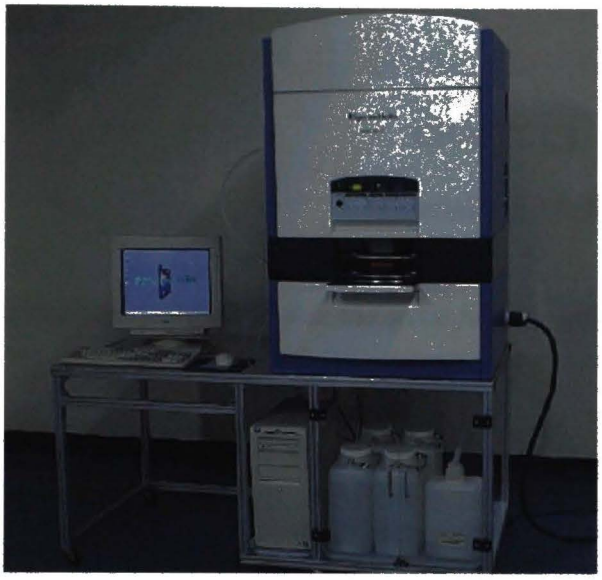


Fig 3 Spectrumedix 9610 Genetic Analyzer (Orchid Cellmark)

Spectrumedix 9610 Genetic Analyzer (Fig 3)

This instrument utilizes an array of 96 capillaries to detect fluorescently labeled DNA. The capillaries are filled with a polymer matrix that allows electrophoresis of DNA fragments. DNA is injected into the capillaries by means of electrokinetic force (12). Fluorescent dye molecules are excited by a laser at the detection window during cycle sequencing (12). Fluorescent emissions from the 96 capillaries are simultaneously collected and stored for analysis.

CHAPTER 3

Material and Methods

* *All samples were treated in the same fashion unless otherwise noted*

1. *DNA Sample Collection*

   Samples were collected from personnel at Orchid Cellmark in Dallas, TX. Samples used in study were samples that were sequenced previously using 4 primer sets at full cycle sequencing reactions. A total of seven known reference samples were used in this study. Samples were obtained by swabbing the inside of the cheek with a two sterile cotton swabs.

2. *DNA extraction*

   DNA was extracted using the Qiagen DNeasy Tissue Kit™ (Cat. No. 69506). The Qiagen Extraction of DNA from buccal swabs procedure was utilized. Twenty-five percent of the swab was cut and placed in 1.7 ml microcentrifuge tube. 200 µl of PBS, 20µl of proteinase K (20mg/ml), and 200µl Buffer AL are added to sample tube. Vortexed and spun briefly. The samples were incubated for 10 minutes at 56°C. Samples removed from water bath and 200 µl ethanol added. They are vortexed for 10-15 seconds and pulse-spun for 60 seconds at 8K rpm. 700 µl of mixture added to DNeasy spin column. Centrifuged for 1 minute. Column placed in a new labeled 2 ml collection tube. 500 µl of AW1 Buffer added to column and spun for 1 minute. 500 µl of AW2 Buffer added to column

and spun for 3 minutes. 100 µl of AE buffer added to column and centrifuged for 1 minute to collect extracted DNA. The final volume of extracted samples was 100 µl.

3. *PCR Amplification*

PCR Amplification using 2 primer sets- mtDNA HVI and HVII region in its entirety. Amplification mix consisted of ABI Taq Gold DNA™ polymerase (Applied Biosystems, Foster City, CA), 2.5mM dNTP mix (Applied Biosystems, Foster City, CA), 10X PCR buffer(Applied Biosystems, Foster City, CA), 15mM $MgCl_2$(Applied Biosystems, Foster City, CA), BSA (Bovine Serum Albumin) 4mg/ml, sterile diH_20. The primers used are custom-made from Applied Biosystems, Foster City, CA. The two primer sets utilized were F15989 and R16410 for HVI and F15 and R389.

Table 1. Amplification Primers for Polymerase Chain Reaction (12).

HVI	**Primer Sequences**
F15989	5'-CCC AAA GCT AAG ATT CTA AT-3'
R16190	5'-CCC CAT GCT TAC AAG CAA GT-3'
F16251	5'-CGA GTT GCA GTT GAT GT-3'
R16410	5'-GAG GAT GGT GGT CAA GGG AC-3'

HVII	**Primer Sequences**
F15	5'-CAC CCT ATT AAC CAC TCA CG-3'
F155	5'-TAT TTA TCG CAC CTA CGT TC-3'
R285	5'-GTT ATG ATG TCT GTG TGG AA-3'
R389	5'-CTG GTT AGG CTG GTG TTA CG-3'

10ul of each sample was added to 20ul master mix in designated PCR reaction tube. Samples were placed on ABI 9700 thermocycler. The thermocycling parameters were as follows:

Pre-Heat at:	96°C for 10 minute and
	94° C for 30 seconds
38 cycles at:	94°C for 20 second (denaturation)
	56°C for 10 seconds (annealing)
	72°C for 30 seconds (extension)
	4°C hold

PCR Amplification using 4 primer sets

Amplification mix consisted of ABI Taq Gold DNA polymerase (Applied Biosystems, Foster City, CA), 2.5mM dNTP mix (Applied Biosystems, Foster City, CA), 10X PCR buffer, 15mM $MgCl_2$ (Applied Biosystems, Foster City, CA), BSA (Bovine Serum Albumin) 4mg/ml, sterile diH_20 for the master mix. The 10 uM amplification primers used are customized from ABI. The 4 primer sets utilized were F15989/R16251, F16190/R16410 for HVI, F15/R285 and F155/R389 for HVII. 10ul of each sample was added to 20ul master mix in a PCR reaction tube. Samples were placed on ABI 9700 thermocycler. The thermocycling parameters were as follows:

Pre-Heat at:	96°C for 10 minute and
	94°C for 30 seconds

38 cycles at: 94°C for 20 second (denaturation)

56°C for 10 seconds (annealing)

72°C for 30 seconds (extension)

4°C hold

4. *Quanitation*

Samples were quantified using Invitrogen E-Gel® 96 2 %Agarose Gel (Cat No. G7008-02). Each E-gel® 96 contains 96 samples lanes and 8 markers lanes. The concentration of ethidium bromide in each gel ranges from 0.1 to .5μg/ml. 15 μl of .5X loading dye (blue juice) and 5μl of amplified product are loaded into each designated well. 5μl of E-Gel® Low Range DNA marker (Invitrogen Cat. No. 12369-013) and 15 μl of .5X loading dye are loaded onto E-gel® 96 marker well. E-Gel transparent cassette placed on E-Gel® 96 mother base (Cat. No. G7100-01). The mother base contains an electrical plug that can be connected directly to an electrical outlet and is used for electrophoresis (12). The default run time for the E-Gel ®96 is 12 minutes. Gels are visualized on an ultraviolet transilluminator for proper staining, resolution, and migration bands (12). Bands are compared to the molecular weight ladder for the concentration. A digitized image of gel is processed by computer to generate a gel image of amplification products (12).

5. *Post- Amplification Purification*

Samples were purified using 96-well Princeton Separation Plates™ (Cat. No. PC-501). 1 volume (25ul) of Binding Buffer added to PCR reaction tubes. Mixed well. Mixture loaded to the plate and centrifuged for 3 minutes. Plate washed by adding 200ul of diluted Wash Buffer into each well. Centrifuged for 3 minutes. Wash step repeated 2 times. Last wash centrifuge for 13 minutes to dry the plate. The purified DNA was eluted with 75µl of elution buffer into the PCR 96 reaction plate.

6. *Cycle Sequencing of Purified Product*

Cycle Sequencing is performed using ABI BigDye Terminator Cycle Sequencing Reaction Kit™ v1.1 (Applied Biosystems, Foster City, CA). Dideoxynucleotide triphospates, which the fluorescent dye molecules have been covalently bound, serve as both chain terminators for the reaction and color indicators of the specific ddNTPs incorporated (12). Cycle sequencing master mix is prepared according to Table 2. Master Mix includes BigDye Sequencing Ready Reaction Mix, 5x Sequencing Buffer, Sterile Distilled water, and 3.2uM Sequencing Primer. Primers used for polymerase chain reaction are used for cycle sequencing (Table 1).

Primers for Sequencing 2 primer sets- F15989, R16410, F15, R389

Primers for Sequencing 4 primer sets- F15989, R16251, F16190, R16410, F15, R285, F155, R389

Reduction of Cycle Sequencing Reactions- Volumes were adjusted according to manufacturer's protocol for full, half and quarter reactions.

Table 2

Example of Master Mix cycle sequencing volumes for 10 samples.

Primer Sets	Reagents	Full	Half	Quarter
4 primer sets	BigDye Sequencing Ready Reaction Mix	768µl	384 µl	192 µl
	5x Sequencing Buffer	0	192 µl	288 µl
	Sterile Distilled water	672 µl	864 µl	960 µl
2 primer sets	BigDye Sequencing Ready Reaction Mix	384 µl	192 µl	96 µl
	5x Sequencing Buffer	0	96 µl	144 µl
	Sterile Distilled water	336 µl	432 µl	480 µl

16ul of primer/master mix and 4 ul (.15pmol) of purified amplified DNA product added to designated well on the 96 well reaction plate. Each PCR product sequenced twice and the plate placed in the ABI 9700 thermocycler under following cycling conditions:

Pre-Heat at: 96°C for 1 minute, then

25 cycles at: 96°C for 10 seconds

50°C for 5 seconds

60°C for 4 minutes

4°C hold

7. *Purification of Cycle Sequencing Products*

Cycle Sequence samples need to be purified of unincorporated ddNTPs and primers. Samples purified with Edge Biosystems, Performa DTR (Dye Terminator Removal) 96-Well Short Plate™ kit (Cat. No.99016). Performa DTR 96- Well Short Plate kit is a gel filtration system that consists of plates with 350-µl volume columns in a standardized array, packed with a gel matrix optimized to effectively remove dye terminators, dNTPs, salts and other low molecular weight materials from sequencing reactions (5). Centrifuge plate for 3 minutes to dispose of buffer in the wells. Cycle sequence samples added to the plate and centrifuged for 3 minutes. Samples collected in new 96 well reaction plate and dried for two and half hours. Stored at 4° C

8. *Sequencing mtDNA*

The purified sequencing products are then prepared to be placed on SpectruMedix 9610 Genetic Analyzer .Samples are resuspended in 18ul of HiDi Formamide (Applied Biosystems, Foster City, CA). The samples are then denatured for 3 min at 95°C on the ABI 9700 thermocycler. Next, they are snap-cooled in 20°C ethanol for 3 minutes. After this step, the sample plate is prepared to be placed on the SpectruMedix 9610 Genetic Analyzer. Data is analyzed using the Spectrumedix BaseSpectrum Software which permits the analysis of data collected by the 9610 sequencer. The software displays captured fluorescence dye emissions as colored peaks on a lone graph (12). To sequence the data provided

by BaseSpectrum analysis, the Sequencher software program is employed. This software displays captured fluorescence dye emissions as colored peaks on a line graph (12). Sequence data is reported with respect to the Revised Cambridge Reference Sequence (rCRS)

CHAPTER 4

Results

Study 1

Amplification with 2 primer sets at full reaction

Five samples were amplified using 2 primer sets F15989/R16410 and F15/R389 at full cycle sequence reaction. Sequences were consistent with known reference samples with the exception of sample KP and RH at polycytosine heteroplasmic region. Sample KP at the 303- 315 region amplified at with the current method produced a 309.1 C, 309.2 C, and 315.1 C. In the same region amplified with 2 primer sets at full reaction was inclusive at 303-315. Sample RH amplified with the current method produced 16189 C while it was inconclusive at 16183-16193 when amplified with 2 primer sets at full reaction. All other samples were concordant with previously recorded known reference samples. Reagent blank, negative control and positive control performed successfully.

Tables show each sample compared to the sample analyzed with the Current method. Both sequences are compared to the revised Cambridge sequence; differences from this sequence are noted.

TABLE 3
2 Primer Sets/Full Reaction
5 SAMPLES

HVI 16024-16365

SITE	16069	16193	16278	16362
SAMPLE # **HL60**	T	T	T	C
CM	T	T	T	C
rCRS	C	C	C	T

HVII 73-340

SITE	73	150	152	263	295	315.1
SAMPLE # **HL60**	G	T	C	G	T	C
CM	G	T	C	G	T	C
rCRS	A	C	T	A	C	-*

HVI 16024-16365

SITE	16093	16192	16207	16256	16270	16291	16324
SAMPLE #**CJ**	C	T	G	T	T	T	C
CM	C	T	G	T	T	T	C
rCRS	T	C	A	C	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **CJ**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16256	16270	16362
SAMPLE # **JF**	T	T	C
CM	T	T	C
rCRS	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **JF**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16093	16224	16311
SAMPLE # **JW**	C	C	C
CM	C	C	C
rCRS	T	T	T

HVII 73-340

SITE	73	114	263	315.1
SAMPLE # **JW**	G	T	G	C
CM	G	T	G	C
rCRS	A	C	A	-*

HVI 16024-16365

SITE	16129
SAMPLE # **KP**	A
CM	A
rCRS	G

HVII 73-340

SITE	263	309.1	309.2	315.1
SAMPLE # **KP**	G	INC		
CM	G	C	C	C
rCRS	A	-*	-*	-*

HVI 16024-16365

SITE	16111		16217
SAMPLE # **RH**	T	16183-16193 INC	C
CM	T	16189-C	C
rCRS	C		T

HVII 73-340

SITE	73	151	263	
SAMPLE # **RH**	G	T	G	303-315 INC
CM	G	T	G	303-315 INC
rCRS	A	C	A	

*- indicates there is no base at this site

Study 2

Amplification with 2 primer sets at half reaction

Four samples were amplified using 2 primer sets F15989/R16410 and F15/R389 at half cycle sequencing reactions. Samples were consistent with previously recorded known reference samples amplified with the current method, with the exception of samples KP and RH as stated in the first study. Reagent blank, negative control and positive control performed successfully.

Tables show each sample compared to the sample analyzed with the Current method. Both sequences are compared to the revised Cambridge sequence; differences from this sequence are noted.

TABLE 4
2 Primer Sets/Half Reaction
4 SAMPLES

HVI 16024-16365

SITE	16069	16193	16278	16362
POSITIVE CONTROL	T	T	T	C
CM	T	T	T	C
rCRS	C	C	C	T

HVII 73-340

SITE	73	150	152	263	295	315.1
POSITIVE CONTROL	T	T	T	C	T	C
CM	T	T	T	C	T	C
rCRS	A	C	T	A	C	-*

HVI 16024-16365

SITE	16256	16270	16362
SAMPLE # **JF**	T	T	C
CM	T	T	C
rCRS	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **JF**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16093	16224	16311
SAMPLE # **JW**	C	C	C
CM	C	C	C
rCRS	T	T	T

HVII 73-340

SITE	73	114	263	315.1
SAMPLE # **JW**	G	T	G	C
CM	G	T	G	C
rCRS	A	C	A	-*

HVI 16024-16365

SITE	16129
SAMPLE # **KP**	A
CM	A
rCRS	G

HVII 73-340

SITE	263	309.1	309.2	315.1
SAMPLE # **KP**	G	INC		
CM	G	C	C	C
rCRS	A	-*	-*	-*

HVI 16024-16365

SITE	16111		16217
SAMPLE # **RH**	T	16183-16193 INC	C
CM	T	16189-C	C
rCRS	C		T

HVII 73-340

SITE	73	151	263	303-315
SAMPLE # **RH**	G	T	G	INC
CM	G	T	G	INC
rCRS	A	C	A	-*

*- indicates there is no base at this site

Study 3

Amplification with 2 primer sets at quarter reaction

Four samples were amplified using 2 primer sets F15989/R16410 and F15/R389 at quarter cycle sequencing reactions. Samples were consistent with previously recorded known reference samples amplified with the current method, with the exception of samples KP and RH as stated in the first study. Reagent blank, negative control and positive control performed successfully.

Tables show each sample compared to the sample analyzed with the Current method. Both sequences are compared to the revised Cambridge sequence; differences from this sequence are noted.

TABLE 5
2 Primer Sets/Quarter Reactions
4 SAMPLES

HVI 16024-16365

SITE	16069	16193	16278	16362
POSITIVE CONTROL	T	T	T	C
CM	T	T	T	C
rCRS	C	C	C	T

HVII 73-340

SITE	73	150	152	263	295	315.1
POSITIVE CONTROL	G	T	C	G	T	C
CM	G	T	C	G	T	C
rCRS	A	C	T	A	C	-*

HVI 16024-16365

SITE	16256	16270	16362
SAMPLE # **JF**	T	T	C
CM	T	T	C
rCRS	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **JF**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16093	16224	16311
SAMPLE # **JW**	C	C	C
CM	C	C	C
rCRS	T	T	T

HVII 73-340

SITE	73	114	263	315.1
SAMPLE # **JW**	G	T	G	C
CM	G	T	G	C
rCRS	A	C	A	-*

HVI 16024-16365

SITE	16129
SAMPLE # **KP**	A
CM	A
rCRS	G

HVII 73-340

SITE	263	309.1	309.2	315.1
SAMPLE # **KP**	G	INC		
CM	G	C	C	C
rCRS	A	-*	-*	-*

HVI 16024-16365

SITE	16111		16217
SAMPLE # **RH**	T	16183-16193 INC	C
CM	T	16189-C	C
rCRS	C		T

HVII 73-340

SITE	73	151	263	
SAMPLE # **RH**	G	T	G	303-315 INC
CM	G	T	G	303-315 INC
rCRS	A	C	A	

*- indicates there is no base at this site

Study 4

Amplification with 4 primer sets at half reaction

Four samples were amplified with 4 primer sets F15989/R16251, F16190/R16410 for HVI, F15/R285 and F155/R389 for HVII at half cycle sequencing reactions. Samples were consistent with previously recorded known reference samples amplified with the

current method. Reagent blank, negative control and positive control performed successfully.

Tables show each sample compared to the sample analyzed with the Current method. Both sequences are compared to the revised Cambridge sequence; differences from this sequence are noted.

TABLE 6
4 Primer Sets/Half Reaction
4 SAMPLES

HVI 16024-16365

SITE	16069	16193	16278	16362
POSITIVE CONTROL	T	T	T	C
CM	T	T	T	C
rCRS	C	C	C	T

HVII 73-340

SITE	73	150	152	263	295	315.1
POSITIVE CONTROL	G	T	C	G	T	C
CM	G	T	C	G	T	C
rCRS	A	C	T	A	C	-*

HVI 16024-16365

SITE	16042	16192	16224	16261	16311
SAMPLE # **AM**	A	T	C	T	C
CM	A	T	C	T	C
rCRS	G	C	T	C	T

HVII 73-340

SITE	73	146	152	263	315.1
SAMPLE # **AM**	G	C	C	G	C
CM	G	C	C	G	C
rCRS	A	T	T	A	-*

HVI 16024-16365

SITE	16256	16270	16362
SAMPLE # **JF**	T	T	C
CM	T	T	C
rCRS	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **JF**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16093	16224	16311
SAMPLE # **JW**	C	C	C
CM	C	C	C
rCRS	T	T	T

HVII 73-340

SITE	73	114	263	315.1
SAMPLE # **JW**	G	T	G	C
CM	G	T	G	C
rCRS	A	C	A	-*

HVI 16024-16365

SITE	16192
SAMPLE # **CD**	T
CM	T
rCRS	C

HVII 73-340

SITE	186	263	315.1
SAMPLE # **CD**	A	G	C
CM	A	G	C
rCRS	G	A	-*

*- indicates there is no base at this site

Study 5

Amplification with 4 primer sets at quarter reaction

Four samples were amplified with 4 primer sets F15989/R16251, F16190/R16410 for HVI, F15/R285 and F155/R389 for HVII at quarter cycle sequencing reactions. Samples were consistent with previously recorded known reference samples amplified with the current method with the exception of sample KP. Sample KP with current method sequenced 309.1, 309.2, and a 315.1 with this proposed method, results where 303-315 INC (inconclusive). Reagent blank, negative control and positive control performed successfully.

Tables show each sample compared to the sample analyzed with the Current method. Both sequences are compared to the revised Cambridge sequence; differences from this sequence are noted.

TABLE 7
4 Primer Sets/Quarter Reaction
4 SAMPLES

HVI 16024-16365

SITE	16069	16193	16278	16362
POSITIVE CONTROL	T	T	T	C
CM	T	T	T	C
rCRS	C	C	C	T

HVII 73-340

SITE	73	150	152	263	295	315.1
POSITIVE CONTROL	G	T	C	G	T	C
CM	G	T	C	G	T	C
rCRS	A	C	T	A	C	-*

HVI 16024-16365

SITE	16042	16192	16224	16261	16311
SAMPLE # **AM**	A	T	C	T	C
CM	A	T	C	T	C
rCRS	G	C	T	C	T

HVII 73-340

SITE	73	146	152	263	315.1
SAMPLE # **AM**	G	C	C	G	C
CM	G	C	C	G	C
rCRS	A	T	T	A	-*

HVI 16024-16365

SITE	16093	16192	16207	16256	16270	16291	16324
SAMPLE # **CJ**	C	T	G	T	T	T	C
CM	C	T	G	T	T	T	C
rCRS	T	C	A	C	C	C	T

HVII 73-340

SITE	73	263	315.1
SAMPLE # **CJ**	G	G	C
CM	G	G	C
rCRS	A	A	-*

HVI 16024-16365

SITE	16093	16224	16311
SAMPLE # **JW**	C	C	C
CM	C	C	C
rCRS	T	T	T

HVII 73-340

SITE	73	114	263	315.1
SAMPLE # **JW**	G	T	G	C
CM	G	T	G	C
rCRS	A	C	A	-*

HVI 16024-16365

SITE	16129
SAMPLE # **KP**	A
CM	A
rCRS	G

HVII 73-340

SITE	263	309.1	309.2	315.1
SAMPLE # **KP**	G	INC		
CM	G	C	C	C
rCRS	A	-*	-*	-*

*- indicates there is no base at this site.

CHAPTER 5

Discussion

As stated previously, any new method, instrument, or reagents being used in a laboratory it must be validated as stated Standard 8.1 by the DAB. In this study an improvised method is being introduced. The validity of this method must be assessed in order to observe the results as opposed to the current method utilized. The new method must be equal or better than the previous method. The incorporation of the new method into the lab may prove to be beneficial. This study sought to reduce the amount reagent used i.e. reduction of cost and reduction of time and labor.

MtDNA analysis is laborious and time consuming process. As shown in chapter three there are eight steps in this process before reaching a final result. These steps include sample preparation, DNA extraction, amplification, quanitation, purification of PCR product, cycle sequencing, purification of product, sequencing and analyzing sequence data. Therefore the analyst must practice good laboratory techniques to prevent contamination. Contamination, the presence of extraneous DNA, is a major problem in mtDNA analysis because it is highly sensitive. If contamination is present in the reagent blank or negative control, one these many steps may have to be repeated which wastes time, money and reagents. In this study, contamination was observed and therefore samples were re-isolated or re-amplified which consumed excess reagents and time.

The current amplification process uses 4 primer sets and each sample is amplified twice during cycle sequencing according to interpretation guidelines. Two

concordant sequences are needed to confirm the sequence. When amplifying using 4 primer sets there will at least be two corresponding samples that will match (Fig 4). This is not the case when amplifying using 2 primer sets (Fig 5). This reduces the number of sequences from eight to four. If a sample does not sequence on the forward or reverse reaction there will not be two sequences to confirm the sequence per the guidelines. Fig 6 shows sequence data for 2 primer sets, at site 16360, it is inclusive. Even though the R16410 strand was sequenced twice the bottom sequence could not be called therefore it is labeled N, the other strand is a C for this site but it will labeled inconclusive because there the other strand at that site is not concordant. Using the four primer set leaves an opportunity for error if a sample not amplifying completely, there will always be two at least sequences to compare. Observe Fig 4 site 16260 although the 2 sequences at this site where designated N there where still 2 other sequences to compare.

Figure 4 Sequence data from amplification with 4 primer sets- 8 Sequences

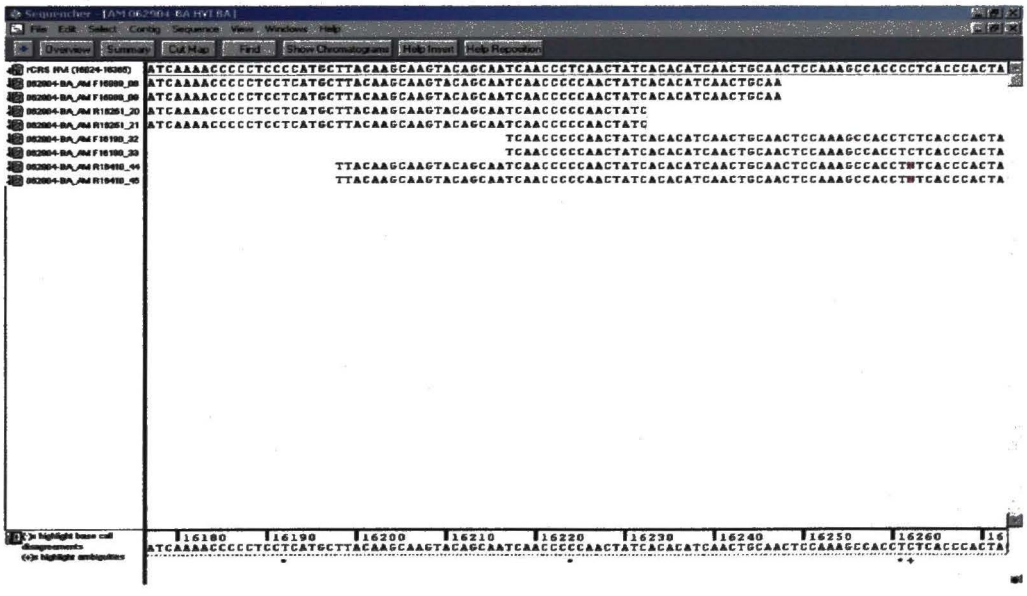


Figure 5 Sequence data from amplification with 2 primer sets – 4 Sequences

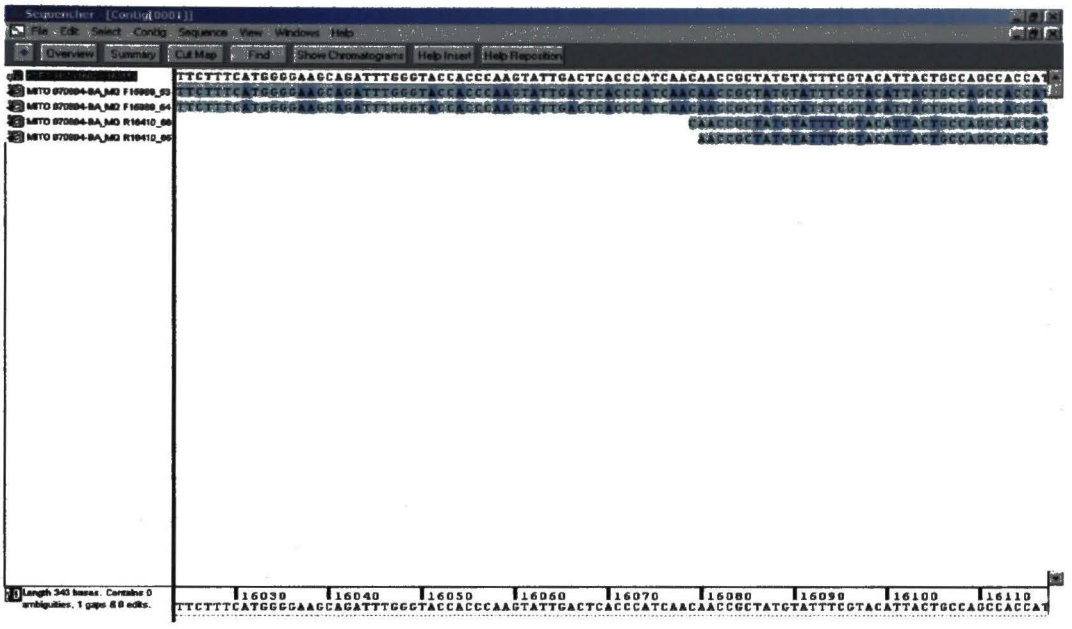
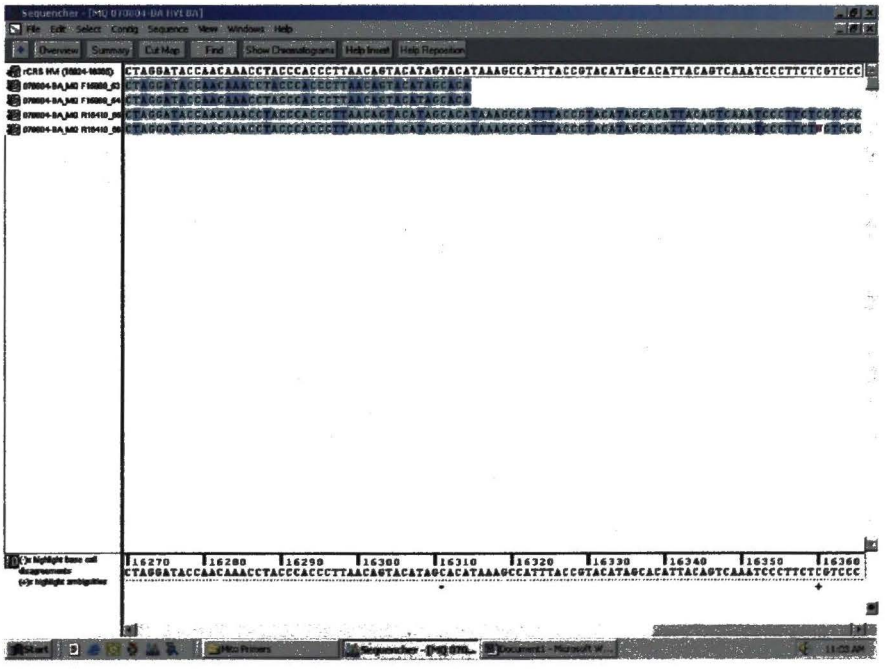


Figure 6 - 2 Primer Set- Inconclusive at site16360



The experimental samples were consistent with known reference sample sequences in all the studies with the exception of the 16183-16193 HVI region and 303- 315 HVII region for two samples RH and KP. This is a drawback with amplifying HVI and HVII in its entirety. Although the two primer method will only be used for reference samples, it does not produce a conclusive haplotype. If the evidentiary samples are amplified using the current method, while reference samples are amplified using the proposed method of two primer sets the sequences will not be concordant in the 303-315 region. For this reason, two primer sets should not be used for evidentiary samples. Evidentiary samples are usually degraded and in low quantity. Smaller primers should be used to amplify these delicate samples.

Another drawback for the amplification of HVI and HVII in its entirety deals with heteroplasmy as discussed earlier. Many sample haplotypes included the common length heteroplasmy at the 16183-16193 (Fig 7) and 303-315 (Fig 8) region. Guidelines for this lab state that for the forward sequence the C-stretch and after must be deleted and for the reverse sequence the C- stretch and sequence before the C-stretch must be deleted. Deleting sequences gives the risk of there not being two concordant sequences. Therefore 2 primer sets should not be used if there are not two concordant sequences to confirm the sequence. If the two primer sets should be used, interpretational guidelines should be revised. It is suggested that if there are not two concordant sequences to confirm a sequence then the sample should be re-amplified using four primer sets.

Figure 7 HVI region 16183-16193 INC length heteroplasmy

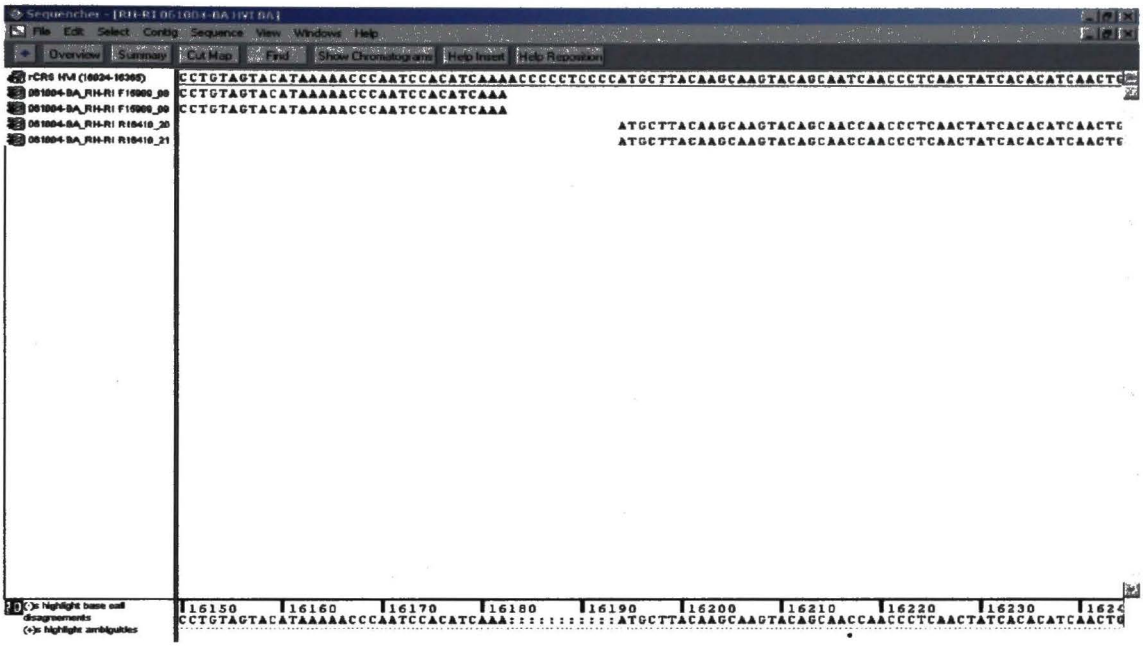


Figure 8 HVII region 303-315 INC length heteroplasmy

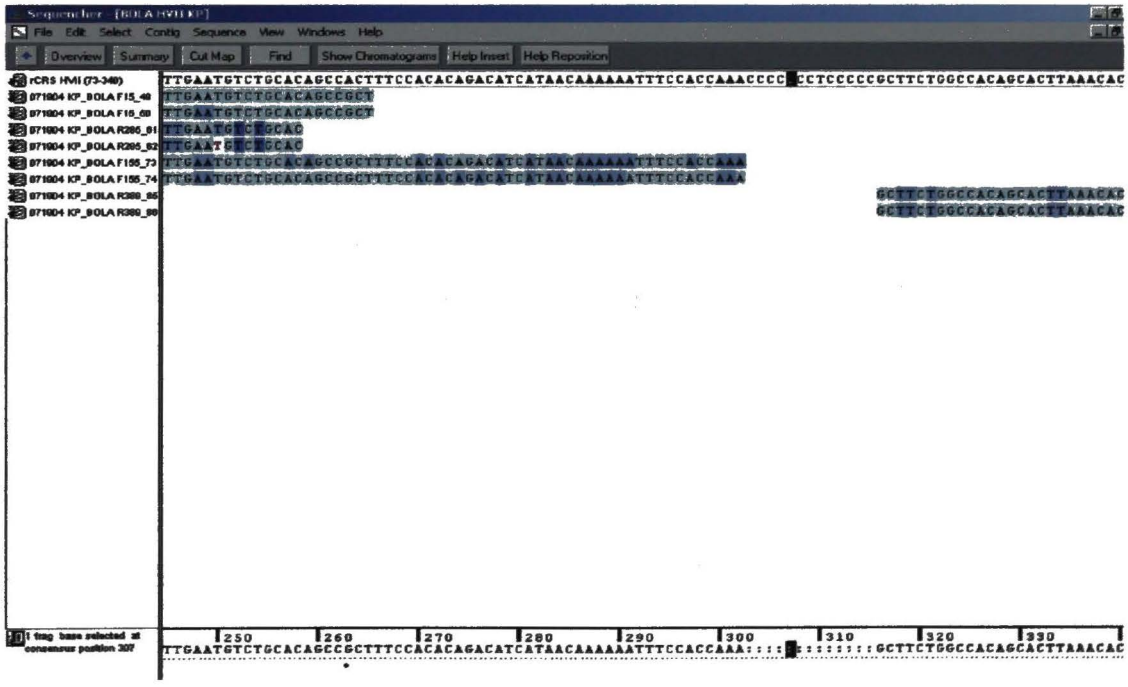


Figure 9 Chromatogram of 16183-16193- replication slippage observed

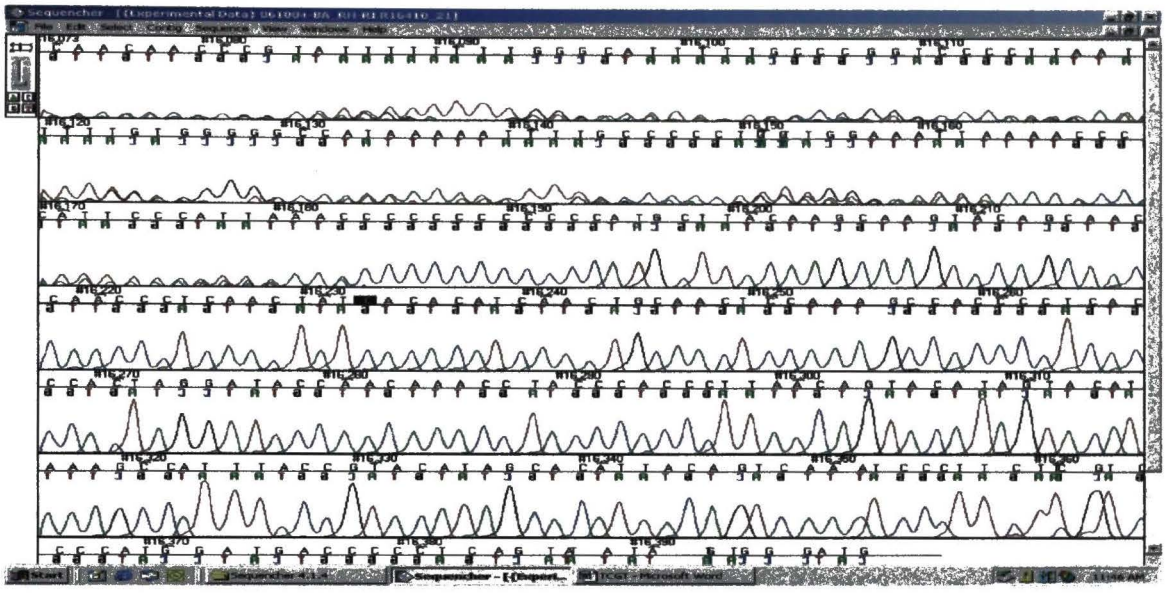
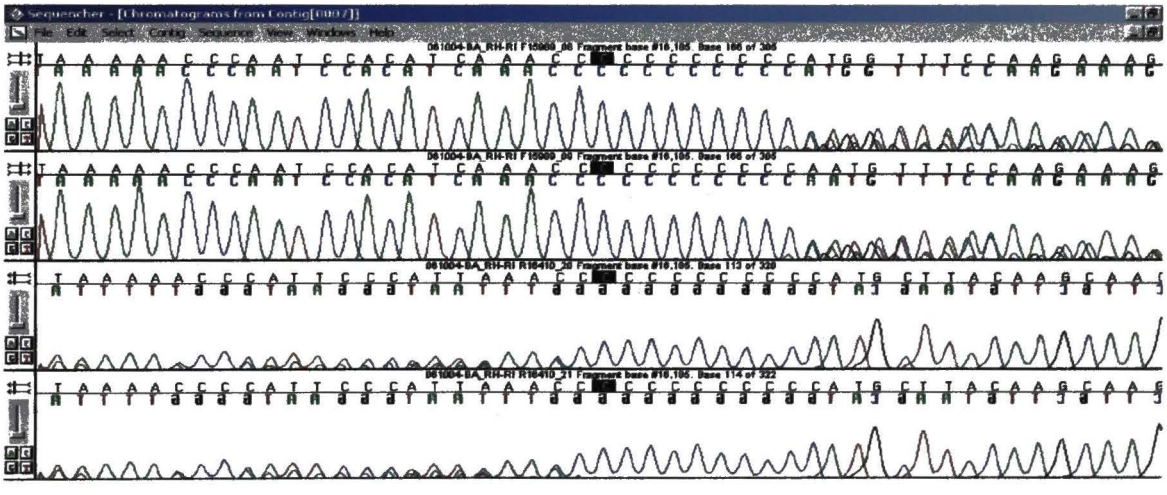


Figure 10 Chromatogram of 303-315- replication slippage observed



The method of amplification with 2 primer sets does have it benefits. Two amplification reaction reduces the amount of tubes, pipetting, and the amount of handling when compared to four amplification reactions. This will in turn reduce the chance of

contamination. Also, analyzing sequence data is not as time consuming with two primer sets. When analyzing four primer sets, primers must be cut, which is not required with two primer sets. The analyst analyzes four sequences as opposed to eight sequences with four primer sets.

Reducing cycle reaction volumes will be beneficial for both commercial and private labs. The current cycle reaction kit used by Orchid Cellmark, ABI BigDye Terminator Cycle Sequencing Reaction Kit v1.1, costs $6,990 per kit. The lab consumes one kit every four weeks. This totals to $ 83, 800 per year. The total would be $ 41, 490 per year using half reactions and $20, 970 per year using quarter reactions. Results demonstrate that sequences cycle sequenced with quarter reaction were concordant with sequences cycle sequenced with full reactions. This proposed method of reducing cycle sequencing reaction volumes proves to be cost effective in labs with a reduction of 40% in costs.

Conclusion

In conclusion, reduction of cycle sequencing would be beneficial for labs in that, it decreases cost by 25-40%. The results are the same as when using full reaction therefore the sequences will not be affected. The other study of this report may not be as advantageous. It is the decision of the lab on whether to incorporate the proposed procedure of amplifying the HVI and HVII region in it entirety because it has advantages and disadvantages. Performing two amplification reactions as opposed to four amplification reactions decreases the amount tubes, handling, and use of primers. Consequently, this will reduce the chance of contamination. Sequencing analysis will be

less time consuming and expensive. However, when analyzing sequences produced by this method they are not as conclusive as sequences amplified with the current method. As illustrated, with sample RH and KP at polycytosine stretches. This a major concern when comparing evidentiary samples amplified with four primer sets and reference samples amplified with two primer sets, if the sequences will not be truly concordant at all polymorphic sites. Therefore, it is the judgment of the lab director to incorporate this method into the protocol based on these factors. It is suggested that the lab should continue use of the four primer sets because of the read-through problems in the polycytosine stretches. If the two primer sets are utilized in the lab and samples demonstrate length heteroplasmy read-through problems, then the questioned sample can be re-amplified using 2 primer sets. Although this proposed method will still be time consuming and expensive to re-test a sample.

These proposed methods are to be used for reference samples. Evidentiary samples are usually in small quantity and degraded therefore it would be favorable to use small primer sets to amplify these small regions of DNA. Reduction of cycle sequencing reactions could be used for both evidentiary and reference cases because results show that sequences with the current method and sequences with reduced reaction volumes were concordant.

Future studies analyzing more samples should be performed with two primer sets since only 2 samples in this study had read-through problems in the c-stretch. Samples should be analyzed using the current method and then compared using two primer sets to observe if the same problem is observed. Furthermore, additional samples should be

tested in reducing cycle sequencing reactions to observe if the sequences will be consistent over time because in this study only four or five samples were tested. There may be problems in the sequence data when producing a cycle sequence master mix for more than twenty samples such as there not being enough labeled dideoxynucleotides.

Until additional studies have been performed analyzing samples with the read through-problems in the c-stretch, amplification using two primer sets should be used with caution. In effort to reduce time, two primer sets may increase the time spent and use of reagents if samples have to be re-amplified using four primer sets because of the read through problems in the c-stretch. In this study reducing cycle sequencing reactions proved to be beneficial but a sample size was used, therefore further testing should be performed with a larger sample size to demonstrate this method is consistent.

CHAPTER 6

References

1. DNA Advisory Board 1998. Quality Assurance Standards for Forensic DNA Testing Laboratories. *For Sci Com* July, 200:2-15

2. Anderson, S., A. T. Bankler, B. G. Barrell et al. (14 co-authors). Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465. (1981)

3. Budowle B., Allard M.W., Wilson, M.R. Critique of interpretation of high levels of heteroplasmy in the human mitochondrial DNA hypervariable region I from hair. *Forensic Science International.* 126:30-33 (2002).

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

5. Edge Biosystems Product Summary. Performa® DTR 96-Well Short Well Plate Kit. www.edgebio.com

6. Holland MM, Parsons TJ. Mitochondrial DNA sequence analysis—validation and use for forensic casework. *Forensic Sci Rev* 11:21–50(1999).

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

8. Isenberg, Alice R .Forensic Mitochondrial DNA Analysis a Different Crime Solving Tool. *FBI Law Enforcement Bulletin.* 16-20 (2002)

9. Isenberg, Alice R, Moore Jodi M. Mitochondrial DNA Analysis at the FBI Laboratory. *Forensic Science Communications.* 1:1-10 (1999)

10. Bicanville, Robert. Mitochondrial DNA Analysis: Admissibility Update. New York

11. Fisher, Constance L., Isenberg, Alice R., Stewart, John E.B., Miller, Kevin W.P., Theisen, Catherine E., Wilson, Mark R., DiZinno, Joseph A.,

Budowle Bruce. Mitochondrial DNA :Today and Tomorrow, Laboratory Division, Federal Bureau of Investigation

12. Orchid Cellmark Forensics Mitochondrial Identification Laboratory Standard Operating Procedures Version 5.2

13. Wilson M.R., Stoneking M. Holland M.M, DiZinno J.A., Budowle B. Guidelines for the use of Mitochondrial DNA Sequencing in Forensic Science. *Crime Laboratory Digest* 20:68-77

14. Wilson M.R, DiZinno J.A., Polanskey D., Replogle, J., and Budowle B. Validation of mitochondrial DNA sequencing for forensic casework analysis. *Int J. Leg. Med.* 105:83-86 (1995)

15. Budowle B., DiZinno J.A, Wilson, M.R. Interpretation Guidelines for Mitochondrial DNA Sequencing. Scientific Analysis Section, Laboratory Division, FBI Academy

16. Budowle et al. DNA Typing Protocols: Molecular Biology and Forensic Analysis. *Eaton* 2000.

17. Bendall KE, Sykes BC. Length Heteroplasmy in the first hypervariable segment of the human mtDNA control region. *AM J. Hum Genet.* 57:248-256 (1995)

18. Butler, John M. Forensic DNA Typing. *Academic Press* 2001.