Naming the Dead — Confronting the Realities of Rapid Identification of Degraded Skeletal Remains

S. M. Edson, J. P. Ross, M. D. Coble, T. J. Parsons, and **S. M. Barritt**
Armed Forces DNA Identification Laboratory
Rockville, Maryland
United States of America

TABLE OF CONTENTS

1042-7201/16-01/Jan. 2004/63–90/$14.00 •

Naming the Dead — Confronting the Realities of Rapid Identification of Degraded Skeletal Remains[a]

REFERENCES: Edson SM, Ross JP, Coble MD, Parson TJ, Barritt SM: Naming the dead — Confronting the realities of rapid identification of degraded skeletal remains; *Forensic Sci Rev* 16:63; 2004.

ABSTRACT: The Armed Forces DNA Identification Laboratory (AFDIL) is one of the leading laboratories in the world for the processing of degraded skeletal remains. Extended efforts have been made to develop protocols and standards that will hold up to the intense scrutiny of both the scientific world and the U.S. legal system. Presented in this paper are the specifics of the in-house systems and procedures that have allowed AFDIL to streamline the processing of degraded skeletal remains and family references for mitochondrial DNA (mtDNA) analysis. These include the development of our in-house bioinformatics systems by which every package and sample that passes through the laboratory is tracked; protocols designed specifically for both questioned and reference samples; and the difficulties inherent in this type of organization. Two case studies presented involve one of ancient remains and one on the recent event of September 11, 2001. Finally, future directions available to both AFDIL and the DNA analysis community as a whole are discussed.

KEY WORDS: Bioinformatics, degraded skeletal remains, family references, high-throughput, mitochondrial DNA, mtDNA, phylogeography, standard operating procedures, SNPs.

INTRODUCTION

For 10 years, the Armed Forces DNA Identification Laboratory (AFDIL) has handled and processed skeletal fragments subjected to a variety of environments, from the benign to the harsh. Anthropologists from the U.S. Army Central Identification Laboratory, Hawaii (CILHI), have recovered samples from around the globe, from Europe to Asia and the South Pacific, to the northern coast of Africa, to the Himalayan Mountains. AFDIL has received cases where a handful of teeth and bone fragments represented an entire 11-person crew of an aircraft. We have received samples that have been variously left to the elements on mountaintops or in the jungle, submerged in saltwater and jet fuel, frozen, or even stored in conventional coffins. It is this gamut of environmental stressors that have allowed us to hone our protocols to most effectively and efficiently acquire mitochondrial (mtDNA) from such samples and bring these missing home.

The primary mission at AFDIL is to work with CILHI and the Armed Forces Medical Examiner System (AFMES) to identify the remains of service members and civilians missing in current and past military conflicts extending back to World War II. We also assist other outside organizations in historical and current death cases. In our role as the DNA laboratory for the AFMES, the contract laboratory for the National Transportation Safety Board (NTSB), and partner with the Disaster Mortuary Operations Response Team (DMORT), we worked to identify the victims and the terrorists from the September 11, 2001 attacks on the Pentagon and at Shanksville, PA. During this mission, the personnel and laboratory structure of AFDIL were completely reorganized for six weeks, and the entire staff worked seven days a week to lay these victims of terrorism to rest. As a result, remains were identified quickly and efficiently and returned to the families in as short a time as possible.

This type of outside casework permits staff to learn more about different types of samples and thus to develop better procedures for processing degraded specimens. Our current standard operating protocols detail what we have seen to be highly effective methods for extracting and analyzing mtDNA from skeletal remains subjected to environmental stress. Despite the difficulties that such samples present, we have managed to develop a "rapid-throughput" method for processing that maintains strict forensic and scientific criteria.

The analysis of DNA using PCR is a tool that allows for the identification of the missing that was not possible before the late 1980s. Since then, the advent of more and more sophisticated technologies have allowed for the analysis of low-copy number nuclear DNA (lcnDNA) and mtDNA. The tools provided by science have allowed optimization of identification using DNA, but most

[a]The views presented in this paper are those of the authors and are not necessarily those of the United States Department of the Army, the United States Department of Defense, the Armed Forces Institute of Pathology, the American Registry of Pathology, or the Department of Justice. Mention of specific brand names within this paper does not necessarily indicate an endorsement of said product by AFDIL, but is merely a statement reflecting our current usage.

particularly mtDNA. In this manner, we are able to bring closure to families who may have been waiting decades for answers as to the disappearance of their missing. Many of the relatives of the missing are aged, particularly where World War II deaths are concerned. There is a sense of urgency to solve these cases before parents, siblings, or children of the missing are deceased.

It is this immediacy that gives the staff at AFDIL a personal stake in the accurate completion of the cases. It is the feelings of the families that are foremost in our minds. All of our cases are processed in anonymity until completion, thus preventing even the appearance of potential bias. This paper will describe our current protocols and procedures and give the reader insights into how AFDIL functions as a whole.

I. SAMPLE COLLECTION AND MANAGEMENT

A. Specimen Types

1. Unknown Samples

The collection of unknown samples is a varied task. At AFDIL, the primary duty of the mtDNA Section is to identify the remains of service members and civilians missing from the Vietnam War, Korean War, and World War II. Therefore, the mtDNA casework samples that we handle on a daily basis are dried skeletal remains. Anthropologists at CILHI remove these samples from whole or partial skeletons. Samples are sent to AFDIL to be processed in the blind, using unique identifiers assigned upon arrival.

Samples from mass fatalities are collected by local or military medical examiners, DMORT personnel, or, on occasion, members of the AFDIL staff. These samples are generally highly fragmented or subjected to extreme environmental pressures such as burning or immersion in jet fuel. A group at AFDIL (AFDIL^{CS}: AFDIL Consultative Services) within the Nuclear DNA Section handles most current death cases involving nonmilitary personnel. Current military death cases, such as Operation Iraqi Freedom, are controlled by the Nuclear DNA Section.

2. Known Samples

Family reference information and samples are collected by outside organizations, primarily the United States Army Casualty and Memorial Affairs Operation Center (CMAOC). Reference samples arrive at AFDIL in the form of whole, uncoagulated blood in 5-cc potassium-EDTA-treated purple-topped tubes (BD Vacutainer™; BD: Franklin Lakes, NJ) or buccal swabs. The family history of the missing person has been traced by military genealogists or by the service branch casualty officers to ensure the suitability of the reference. The volume of family references received each month was difficult to process in a manual fashion, resulting in a backlog of over 2000 specimens, which has since been eliminated through the implementation of a high-throughput system.

In the event that a family reference is unavailable for mtDNA, direct references such as hair, stamps, or biopsy samples are submitted. Stamps and envelopes are analyzed and reported with caution since the source cannot be verified and a pre-extraction cleaning procedure is not feasible. In current death cases, direct references are more easily obtained and can be more varied. Toothbrushes, razors, shoes, clothes, and hats have been submitted for direct references. As with stamps, the results must be interpreted with caution as other people may have used or worn the samples in question. All samples received by AFDIL, either reference or specimen, are assigned a unique identifier.

B. Data Management

The processing of a large amount of samples for both high-throughput family references and degraded samples would be next to impossible if done entirely by manual processing. Below, we describe the implementation of instrumentation into our protocols to facilitate the processing of both types of samples within the context of this paper. However, we also have a computerized system that manages the actual lab processes being undertaken on a particular sample. In this section we will describe the system in more detail and describe how it has allowed us to increase capacity by removing the manual tracking of samples, which can be tedious and time-consuming.

1. Laboratory Information Management System

For large-scale forensic operations, a laboratory information management system (LIMS) must be used to facilitate the management of cases. Over the last five years, AFDIL has developed a comprehensive, enterprise-wide LIMS. AFDIL's current system, the Laboratory Information Systems Application (LISA), was developed in conjunction with a team from Future Technologies Inc. (FTI: Washington, DC). This comprehensive system manages virtually all aspects of laboratory operations in accordance with AFDIL's validated procedures and policies. User rights and privileges are assigned in LISA via a role-based access system. Within the security application, users are assigned to different groups or roles with the system, thus limiting or granting access to functionality and/or applications contained within the enterprise suite. As in all aspects of forensics, the development of a LIMS is a dynamic process, and LISA is continuing to evolve. A team of programmers from FTI are available full-time on-site and perform continual

updates to the system as our needs change. Specific applications for LISA will be discussed throughout the paper.

2. Case and Sample Tracking

Virtually all aspects of forensic science are bound by the common principles of chain of custody. For this reason, it is imperative that any LIMS be able to manage and track evidentiary materials throughout their lifecycle. LISA tracks packages, AFDIL cases, and evidentiary specimens, as well as evidence retrieval and storage information throughout the testing process. LISA Case Accessioning (CA) defines package, case, and evidence as such:

a. Package: Any specimen or group of specimens received into the laboratory. Receipt of the package could be from a commercial carrier (FedEx®, UPS®, etc.) or hand-delivered.
b. Case: A group of specimens received for a common investigation or incident. Cases are assigned an AFDIL case number for tracking (Year_One-letter case type designator_Four-digit counter, e.g. 2000H0001).
c. Specimen: An individual piece of evidentiary material assigned to an AFDIL case. The evidence is assigned a five-digit alphanumeric designator that is appended to the AFDIL Case Number to which it is assigned (e.g., 2000H0001-0001A).

3. Custody of Evidence

LISA handles all sample tracking and chain of custody (COC) through the Case Accessioning (CA) application. Within LISA-CA the tracking is initiated with the creation of the package. From this point on, all information pertaining to the package location and custody is tracked electronically within the system. The role-based access privileges LISA-CA application allows users to have varying levels of functionality assigned. That is, within the system some users can only accept packages and transfer custody (receiving department), while others can also create new cases, log evidence, and assign analysts to work on the case (evidence custodian, casework analysts). The flow of specimens, cases, and packages are tracked using a password authentication and verification system. In order to transfer custody of a package to another individual within the system you must first have custody of the given specimen. Custody is verified via the system password identification.

4. LISA Mobile

A core mission of AFDIL is to assist the AFMES in human identification casework. This mission often involves lab personnel being deployed to a disaster or incident site to collect specimens for identification purposes. In order to streamline the collection process, ensure data integrity and maintain proper chain of custody of specimens collected in the field, AFDIL developed a mobile version of the in-house LISA application. LISA Mobile is a condensed version of LISA-CA, which has been configured into a self-contained mobile "kit" (**Figure 1**). Each LISA Mobile kit is configured with the supplies needed to collect and document approximately 1000 DNA specimens. This scaled-down version of LISA is configured to work on a touch-screen, rugged mobile laptop from which the user can document evidence, print barcode sample labels, and initiate COC. Features of the system include automated sample number generation, eliminating the possibility of duplicate specimen numbers, and automatic correlation of any contributor number (DMORT, local ME, etc.) with the AFDIL case and specimen numbers. The system also automatically checks the contributor numbers against the existing database to eliminate redundant sampling of

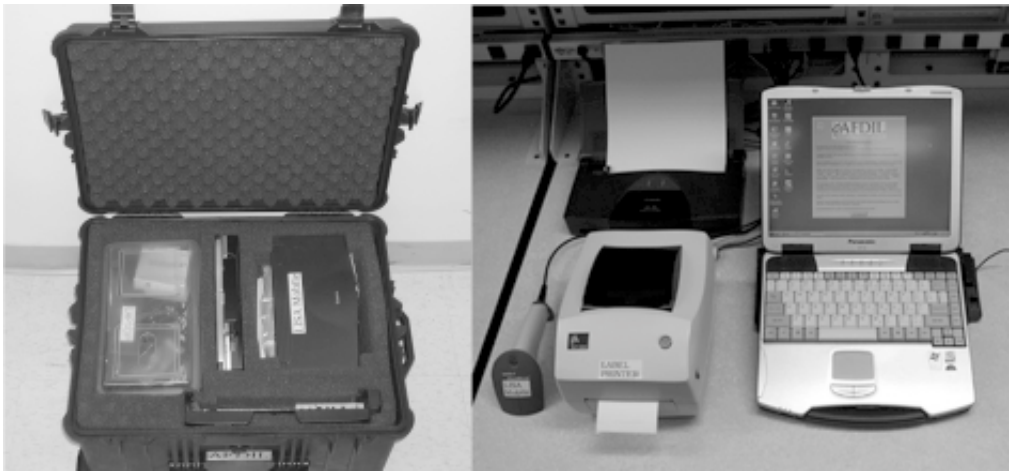

Figure 1. LISA Mobile. On the left is the "Go-Kit" contained within the travel kit. Each piece of the kit has a cushioned slot. Each LISA Mobile kit contains supplies to collect approximately 1000 specimens. On the right is LISA Mobile in operational mode.

specimens during large-scale processing (e.g., September 11, 2001 attacks, Space Shuttle Columbia disaster, and Operation Iraqi Freedom).

Once samples have been collected and are ready for shipment or transport to the testing laboratory, the LISA Mobile system will create a package and produce the proper evidence voucher. During this process, the system creates an export file, which can be transferred to removable media (e.g., floppy disk, CD-R). This file contains the specimen information listed on the COC. Once the "package" (specimens, signed COC and export file) is received at AFDIL, the exported COC information from LISA Mobile can be automatically imported into the LISA production database, thus eliminating data transfer errors and streamlining the start of sample processing.

5. DNA2000

Samples logged into LISA-CA are now ready for laboratory processing. LISA Laboratory Processing (LISA-LP) is designed to automate and track laboratory procedures and contains all of AFDIL's validated forms and protocols for both mitochondrial and nuclear DNA. Users are allowed to retrieve evidence for processing only if it is currently in their custody. Once evidence has been "checked out", it is available in the laboratory processing module. LISA-LP is a protocol/form-driven system. The procedure chosen by the analyst dictates the forms presented to the analyst for use. Procedures predicated on the completion of a previous step are not allowed until completion of the first (i.e., samples are not available for amplification until extraction for those samples has been completed). As in many areas of LISA, manual entry of data is avoided as much as possible. For this reason, LISA-LP is designed to interface with instrumentation where possible. For example, sample sheet and plate record information are exported from LISA to the ABI PRISM® 377 DNA Sequencer and 3100 Genetic Analyzer, respectively, eliminating typographical errors.

II. PROTOCOLS

A variety of techniques is used at AFDIL to effectively and efficiently pursue the high throughput of family references and the rapid throughput of degraded and other samples. Full STR profiles are generated from such samples as buccal swabs, bloodstains, and fresh tissue or bone. In the case of a mass disaster, these are the types of samples most frequently processed. However, the principal mission of the AFDIL mtDNA Section is to process low-copy-number DNA samples where nuclear DNA is not the primary target. In this section, we will explore some of these protocols in greater depth and present some of the strategies currently in use.

A. Handling of Evidence

Upon arrival at AFDIL, a sample is accessioned by our evidence custodians and assigned a case number. Resubmissions are labeled with the appropriate case number and the assigned scientist notified that additional evidence has been received. Evidence is entered into our LISA system, and an electronic chain of custody is initiated. Bone samples received from CILHI are stored at –80 °C until the samples can be extracted. Whole blood is aliquotted onto bloodstain cards (Fitzco: Spring Park, MN), allowed to air dry, vacuum-sealed into individual foil envelopes, and stored at –80 °C until extraction.

Testing usually destroys the entirety of a bone sample. If they are not consumed, the remaining fragments are returned to CILHI upon the closing of the case by CILHI. Family reference bloodstain cards are stored in-house indefinitely.

B. Extraction

The type of extraction used varies with the type of sample being extracted and the purpose of the extraction. Protocols for the extraction of high-copy-number DNA samples varies significantly from that of degraded DNA samples. As discussed in the sections on Laboratory Design and Personnel (Sections III.C.3 and II.H, respectively), the laboratories are separated from each other to prevent the risk of contamination. Specialized samples that are not routinely handled by AFDIL, such as urine, stamps, and envelopes, are processed on a case-by-case basis with explicitly designed protocols that will not be discussed here, but examples of which can be found in the following references: urine [37,44], feces [38,69], ancient muscle [43], stamps and envelopes [2], paraffin [40], hair [3,39,72], and finger- or toenails [6].

1. Degraded Skeletal Remains

Optimally, samples are received from CILHI in 2.5–5.0-g fragments. Samples of powdered dentin are also received from CILHI. The dentin is removed from the teeth at the CILHI laboratory by odontologists, who maintain the integrity of the identifying morphological characteristics of the teeth [64]. Samples are not processed if they consist largely of spongy or trabecular bone or have been exposed to excessive amounts of heat or burned to the point of calcination.

Each sample from a case is handled in a standard evidentiary method with no more than one sample being cleaned at any one time. A sample (other than a tooth) is sanded using an aluminum oxide grinding stone attached to a Dremel® tool (Dremel: Racine, WI). The entire exterior of the sample is sanded clean to remove potential

contaminants [35]. Extremely curved or cracked samples are split along the fissure using a 0.25"-thick emery wheel to assist in cleaning the interior. If the final weight of the cleaned sample is more than 2.0 g, the sample is split into 2.0-g fragments or smaller using a mortar and chisel. The desired sample is washed with sterile distilled water and 100% ethanol and air-dried.

Completely dried samples are pulverized using a sterilized Waring® (Waring: Torrington, CT) MC2 blender cup and 1.0/1.2-L laboratory blender motor and transferred to a sterile 15-mL conical polypropylene tube. Three milliliters of extraction buffer (10 mM Tris, pH 8.0; 100 mM NaCl; 50 mM EDTA, pH 8.0; 0.5% SDS) and 100 µL 20 mg/mL Proteinase K are added to the sample. A reagent blank is initiated at this time.

After an overnight incubation at 56 °C with gentle agitation, samples are cleaned using a Phenol/Chloroform/Isoamyl Alcohol (25:24:1) series of washes, followed by a wash with *n*-Butanol. Purification of the extracts takes place using TE buffer (10 mM Tris; 1 mM EDTA, pH 7.5) washes in Centricon-100® centrifugal filter units (Millipore: Billerica, MA). The concentrate is transferred to a 1.7-mL microcentrifuge tube and diluted to a final volume of 50–100 µL with TE buffer.

2. Whole Blood Family References

a. High-Throughput Extractions

In order to handle the volume of family references submitted for analysis at AFDIL, we have moved almost entirely to an automated system. Upon removal from the freezer, bloodstain cards are batched into groups of 92. The sample removed from each card is a 3.2 mm hole-punched directly into a 96-well plate by a Wallac DBS Puncher (PerkinElmer Life and Analytical Sciences: Boston, MA). An optical reader scans the barcode assigned to each card, and the associated software populates the sample sheet automatically. Each card is punched once into an empty well. The puncher head is cleaned between samples. This process is repeated until all the cards in the set have been sampled. The plate is sealed with Parafilm M® (American National Can: Menasha, WI) until ready for extraction. Extraction of plates is done using a BioRobot® 9604 (Qiagen: Venlo, The Netherlands). Plates are placed on the BioRobot® 9604 and extracted independently of human interference following the manufacturer's recommendations and a protocol modified from Qiagen's protocol for the extraction of buccal swabs.

b. Individual Processing of Blood Stains

In some cases it would be preferable to process a sample either individually or in a small group to generate sequence information in a shorter time frame. The use of an automated recording system is removed, as there is no need for the Wallac DBS Puncher; however, the LISA tracking system (LISA-LP) can be used with individual samples. Each bloodstain card has a single one-eighth-inch punch or 3.0–5.0-mm square removed from it using either a pair of scissors or handheld paper punch. The instrument of removal is thoroughly cleaned between samples. Punches are placed directly into individual 1.7-mL microcentrifuge tubes. A reagent blank is initiated at this point. The sample is incubated at room temperature in sterile distilled water for a minimum of 15 min and then pelleted to remove the heme. The supernatant is removed and the pellet resuspended in 5% Chelex-100® resin beads (Bio-Rad: Hercules, CA). The sample is incubated at 56 °C from 30 min to 24 h depending on the substrate upon which the stain was deposited, vortexed, and incubated for 8 min in boiling water. Sample can be stored at 2–8 °C on the Chelex-100® beads for short-term. Long-term storage requires the removal of the sample from the Chelex-100® resin to a sterile tube for freezing at –20 °C [70]. From this point on, individual family references are handled in the same fashion as degraded skeletal remains.

C. Amplification

We have developed a specialized group of primers that allows efficient amplification of mtDNA and ensures that we are generating the authentic sequence. **Figure 2** displays the overall strategy of available primers and some of the corresponding validated primer pairs. The full control region primer pair (F15971/R599) and the hypervariable (HV) region primer pairs (F15971/R16410 and F15/R389) are used only for the amplification of high-quality mtDNA samples. All other primers are used for the amplification of low-quality mtDNA samples, the primer pair chosen depending on the level of degradation of sample and the desired amplicon. **Table 1** lists the sequences of each of the primers.

1. Degraded Skeletal Remains

The amplification of degraded skeletal remains takes place in a laminar flow hood that has been washed with 10% commercial bleach and exposed to ultraviolet radiation prior to use. The optimal amount of DNA for amplification is 10–1000 pg; however, the mtDNA available in an extract is not quantified prior to amplification. Therefore, 1–10 µL of template is used in the initial PCR reaction and adjusted accordingly in subsequent reactions.

A PCR master mix containing 10× PCR Buffer (100 mM Tris-HCl, pH 8.3; 500 mM KCl; 15 mM $MgCl_2$) (Applied Biosystems: Foster City, CA), DNA grade Bovine

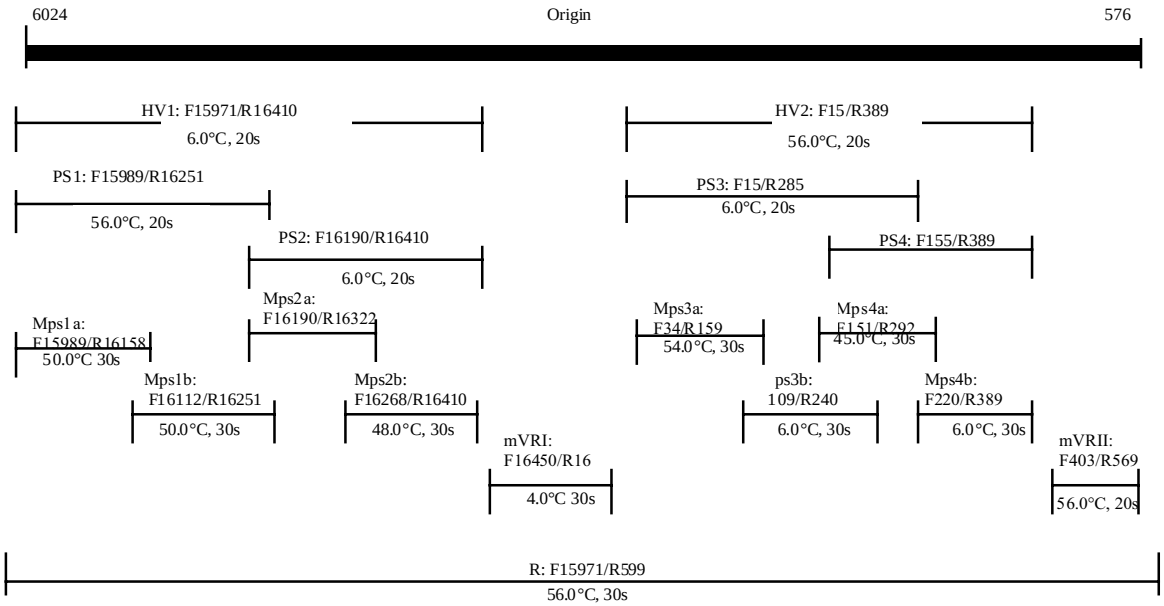

Figure 2. Graphical representation of the possible primer schemes used by AFDIL and the relation of the primers to the mitochondrial DNA control region. Numbering is based on the nucleotide positions within the rCRS [7]. The overlapping strategy of primer design facilitates confirmation of the entire control region. The annealing temperatures and times for the primer pair are listed beneath the indicated primers.

Serum Albumin (BSA, 0.625 µg/µL) (Invitrogen: Carlsbad, CA), deoxynucleoside triphosphate mix (dNTP: 2.5 mM each of dATP, dCTP, dGTP, dTTP) (Applied Biosystems: Foster City, CA), and sterile deionized water. This master mix and reaction tubes for two negative controls, the associated reagent blank, one positive control, and the desired extract are irradiated in an ultraviolet crosslinker with 3 J/cm^2. After irradiation, 10 µM each of the desired primer pair and 5.0–12.5 units/µL AmpliTaq Gold® DNA Polymerase (Applied Biosystems: Foster City, CA) are added to the master mix. The desired amounts of extracted template, reagent blank, and positive control are added to their respective tubes prior to the addition of 40 µL of the master mix to each. Final reaction volume is 50 µL.

AFDIL uses Applied Biosystems (Foster City, CA) GeneAmp® PCR Systems 9600 and 9700 for amplification. Listed below is the program designed for primer sets:

10-min soak at 96.0 ˚C
Followed by 38 cycles of
  20 s at 94.0 ˚C
  20 s at 56.0 ˚C
  30 s at 72.0 ˚C
Final hold of 4 ˚C upon completion

Modifications of this basic program have been designed for the different primer pairs, some of which can be found in Gabriel et al. [25]. Upon completion, samples are quantified upon 2% agarose gels and visualized using 10-mg/ml ethidium bromide (EtBr) and a UV lightbox. Samples and controls are evaluated prior to continuation of amplifications or sequencing.

2. High-Throughput Family References

The 96-well plate strategy for the extraction of blood samples is continued for amplification. From 1 to 10 µL of each template is transferred to a clean and sterilized 96-well plate and brought to a total volume of 10 µL using distilled water. Master mix is made with the same components as for degraded DNA samples without the addition of BSA, which is unnecessary for high template volume samples. Master mix is pre-made in quantities sufficient for an entire plate of samples and stored at –20 ˚C until ready for use. AmpliTaq Gold® DNA Polymerase is added after the samples are thawed and prior to addition to the plate. Final volume in each well of the plate should be 50 µL.

High-throughput samples are primarily amplified using full control region primers (F15971/R599) and the following program:

10-min soak at 96.0 ˚C
Followed by 36 cycles of
  30 s at 94.0 ˚C
  30 s at 56.0 ˚C
  60 s at 72.0 ˚C
Final hold at 4 ˚C upon completion

Table 1. The sequences of the primers used at AFDIL for the amplification and sequencing of both high-quality and degraded samples (Numbering is based on nucleotide positions within the rCRS [6].)

F15971	5'	TTA	ACT	CCA	CCA	TTA	GCA	CC		3'
F15989[a]	5'	CCC	AAA	GCT	AAG	ATT	CTA	AT		3'
F16112	5'	CAC	CAT	GAA	TAT	TGT	ACG	GT		3'
F16144	5'	TGA	CCA	CCT	GTA	GTA	CAT	AA		3'
F16190[a]	5'	CCC	CAT	GCT	TAC	AAG	CAA	GT		3'
F16222	5'	CCT	CAA	CTA	TCA	CAC	ATC			3'
F16268	5'	CAC	TAG	GAT	ACC	AAC	AAA	CC		3'
F16450[a]	5'	GCT	CCG	GGC	CCA	TAA	CAC	TTG		3'
R16158	5'	TAC	TAC	AGG	TGG	TCA	AGT	AT		3'
R16175	5'	TGG	ATT	GGG	TTT	TTA	TGT	A		3'
R16237	5'	TGT	GTG	ATA	GTT	GAG	GGT	TG		3'
R16251	5'	GGA	GTT	GCA	GTT	GAT	GT			3'
R16255	5'	CTT	TGG	AGT	TGC	AGT	TGA	TG		3'
R16258	5'	TGG	CTT	TGG	AGT	TGC	AGT	TG		3'
R16322	5'	TGG	CTT	TAT	GTA	CTA	TGT	AC		3'
R16400	5'	GTC	AAG	GGA	CCC	CTA	TCT	GA		3'
R16410[a]	5'	GAG	GAT	GGT	GGT	CAA	GGG	AC		3'
F15[a]	5'	CAC	CCT	ATT	AAC	CAC	TCA	CG		3'
F29	5'	CTC	ACG	GGA	GCT	CTC	CAT	GC		3'
F34	5'	GGG	AGC	TCT	CCA	TGC	ATT	TGG	TA	3'
F109	5'	GCA	CCC	TAT	GTC	GCA	GTA	TCT	GTC	3'
F140	5'	CCT	GCC	TCA	TCC	TAT	TAT	TTA		3'
F151	5'	CTA	TTA	TTT	ATC	GCA	CCT			3'
F155	5'	TAT	TTA	TCG	CAC	CTA	CGT	TC		3'
F220	5'	TGC	TTG	TAG	GAC	ATA	ATA	AT		3'
F403[a]	5'	TCT	TTT	GGC	GGT	ATG	CAC	TTT		3'
R16	5'	TGA	TAG	ACC	TGT	GAT	CCA	TCG	TGA	3'
R159	5'	AAA	TAA	TAG	GAT	GAG	GCA	GGA	ATC	3'
R240	5'	TAT	TAT	TAT	GTC	CTA	CAA	GCA		3'
R270	5'	TGG	AAA	GTG	GCT	GTG	CAG	AC		3'
R274	5'	TGT	GTG	GAA	AGT	GGC	TGT	GC		3'
R285[a]	5'	GTT	ATG	ATG	TCT	GTG	TGG	AA		3'
R292	5'	ATT	TTT	TGT	TAT	GAT	GTC	T		3'
R377	5'	GTG	TTA	GGG	TTC	TTT	GTT	TT		3'
R381	5'	GCT	GGT	GTT	AGG	GTT	CTT	TG		3'
R389	5'	CTG	GTT	AGG	CTG	GTG	TTA	GG		3'
R484	5'	TGA	GAT	TAG	TAG	TAT	GGG	AG		3'
R569	5'	GGT	GTA	TTT	GGG	GTT	TGG	TTG		3'
R599[a]	5'	TTG	AGG	AGG	TAA	GCT	ACA	TA		3'

[a] These primers are used for the sequencing of the control region in high-throughput family references.

D. Sequencing

The same primers listed in Table 1 are used for the sequencing of mitochondrial DNA. The design of the primers allows the sequencing of overlapping fragments. The overlap induces a level of redundancy that ensures the authenticity of the generated sequence.

1. Degraded Skeletal Remains

After evaluation on a 2% agarose gel, the samples giving a positive result along with the positive control are purified using Centricon-100® centrifugal filter units for hypervariable regions and primer sets and Centricon-30® centrifugal filter units for mini-primer sets. Samples are washed with distilled water and diluted to a final volume of 40–100 µL with distilled water depending on the results determined from the product gel.

The sequencing reactions are set up using two tubes for each sample being sequenced, one tube for each primer. The reaction mixture contains the following: 1.0-µL sequencing primer (10 µM), 1.0–11.0-µL purified PCR product (1.0–20.0 ng), 8.0-µL ABI PRISM® Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (BigDye v. 1.0 or 1.1) (Applied Biosystems: Foster City, CA), and *q.s.* to 20 µL with sterile distilled water. A sequencing control is also initiated at this time using 1.0 µL of –21 M13 primer and 2.0–4.0 µL of PGEM® -3Zf(+) DNA. Applied Biosystems

GeneAmp® PCR Systems 9600 and 9700 are used for sequencing. The following program is used:

25 cycles of
15.0 s at 96.0 °C
5.0 s at 50.0 °C
2 min at 60.0 °C
4.0 °C hold upon completion

The sequencing reactions are purified using Performa® DTR Gel Filtration Cartridges (Edge BioSystems: Gaithersburg, MD) and dried down in an evaporator/concentrator centrifuge.

Sequenced mtDNA from degraded skeletal remains is loaded onto an acrylamide slab gel in an ABI PRISM ® 377 DNA Sequencer. Prior to loading, the samples are resuspended in 4.0 µL of loading buffer (3:1 deionized Formamide/EDTA; 25 mM with 50mg/ml Dextran blue), heated to 96.0 °C for 2.0 min and immediately placed on ice. Gels may run for 1.0–2.5 h depending on the length of the region being sequenced.

2. High-Throughput Family References

The 96-well format is maintained for simplicity. Each plate of amplification reactions is purified using a master mix of 1.5-µL ExoSAP-IT® (USB: Cleveland, OH) and 18.5-µL SAP dilution buffer (50 mM Tris-HCl, pH 8.0) per sample, heated to 37.0 °C for 15 min followed by a second hold at 85.0 °C for 15 min. Samples are then ready to be stored at 4.0 °C or to be sequenced.

The amplification plate is used for storage of the amplified product. From 1.0 to 11.0 µL of each amplicon is transferred to a new 96-well optical plate for sequencing. All negatives, reagent blanks, and positive controls associated with a single amplification are included in each sequencing reaction. One plate is used for each primer to be sequenced. The entirety of the control region can be sequenced by a series of eight primers designated by an indicator in Table 1. A full sequencing reaction using the same parameters as those for degraded samples is performed on each plate.

After sequencing, the samples are purified using 96-well Performa® DTR gel filtration blocks (Edge BioSystems: Gaithersburg, MD). Samples are dried down using an evaporator/concentrator and can be stored at –20 °C for up to one month.

The ABI PRISM 3100® Genetic Analyzer is used for visualizing the sequencing product of high-volume mtDNA samples. Concentrated samples are resuspended in Hi-Di™ Formamide (Applied Biosystems: Foster City, CA) and can be stored at –20 °C after resuspension. Plates of samples are loaded onto the ABI PRISM 3100® Genetic Analyzer and run using a 50- or 36-cm array.

E. Data Analysis

Data from both the ABI PRISM 3100® Genetic Analyzers and the ABI PRISM 377® DNA Sequencers are analyzed by two independent scientists using Sequencher™ Plus, data analysis software developed by Gene Codes Corp. (Ann Arbor, MI). Data are reported only when confirming data exists from either two amplifications or two extractions for a single sample and both scientists agree on the variations from the revised Cambridge reference sequence (rCRS) [5,7]. After a review process involving a minimum of two other independent scientists, the data are entered into our LISA system and made internally available to the entire laboratory.

F. Case Management and mtDNA Sequence Comparisons

AFDIL has worked successfully over the years to increase efficiency and productivity within the laboratory via improved protocols, optimizing workflow, and robotics. However, this increase in productivity has also greatly increased the burden of data management. Analysts are required to manage, review, and interpret greater amounts of data over a wider range of cases. In order to help alleviate some of this burden, a Case Management module was added to the LISA application (LISA-CM). The main challenges in developing LISA-CM was to help streamline the review process and ensure data integrity while still presenting data to the end user in an usable intuitive interface. For mtDNA applications, LISA-CM has four basic submodules: (a) Family Reference Specimen database; (b) Casework profile database; (c) Case Inquiry; and (d) `g:cat`, an mtDNA sequence searching and comparison tool.

LISA-CM was designed to give analysts tools to help manage one of the world's largest missing-persons and disaster-victim databases. Within LISA-CM, the Family Reference Specimen (FRS) database system is used to manage victim and reference demographic and relationship information, case status, mtDNA sequence results, and automated consultation report generation. As in most modules within LISA, a role-based access system is used to grant/restrict user functionality. This security structure is observed predominantly in the online review process where review levels are assigned based on training and experience. That is, some users are restricted to initial review while other members of the senior staff have final review and revision privileges.

As mentioned earlier, some of the design criteria were to streamline data entry and review while maintaining data integrity. In the past, mtDNA analysts would analyze data

using Sequencher™ Plus (GeneCodes Corp.: Ann Arbor, MI), print out the difference list and manually enter that profile into the database. In order to eliminate this time-consuming and error-prone activity for family references, the list of differences from the reference sequence generated by Sequencher™ Plus is imported directly into the LISA application, virtually eradicating transcription errors. Profiles are automatically associated with the correct samples via sample name, which was autogenerated within the LISA DNA2000 module. Consequently, hundreds of profiles can be imported into the system within seconds. However, if needed, sequences may still be entered manually. An added feature of this function is that the total number of sequence bases obtained are automatically calculated as well as the number of matching profiles within the current population database at AFDIL (**Figure 3**). From this point on, the online review process takes over.

Once a profile is imported into the FRS database, it is accessible to all staff for review. The profile and associated information must go through two levels of review before the consultative report generation functionality is enabled. Several features were added to aid the reviewer during the process. All data that appear on the final report are color-coded to indicate the review status (e.g., new, initial, final). Data that have been entered but not reviewed are displayed in red and data that have been 'initial' reviewed are colored blue. This color-coding scheme, which is consistent throughout all systems within LISA-CM, not only instantly directs the user to what needs to be done next but also visually shows the user what data need to be reviewed. Users with the proper access levels "sign off" on each level of review via password authentication. Once data have been final reviewed, the reporting tools are enabled. All changes to the sample data from the point of final review are tracked and audited and can be viewed at any time.

LISA-CM design began with the family reference specimen system, but has since expanded to evidentiary specimen management. The CILHI evidence management

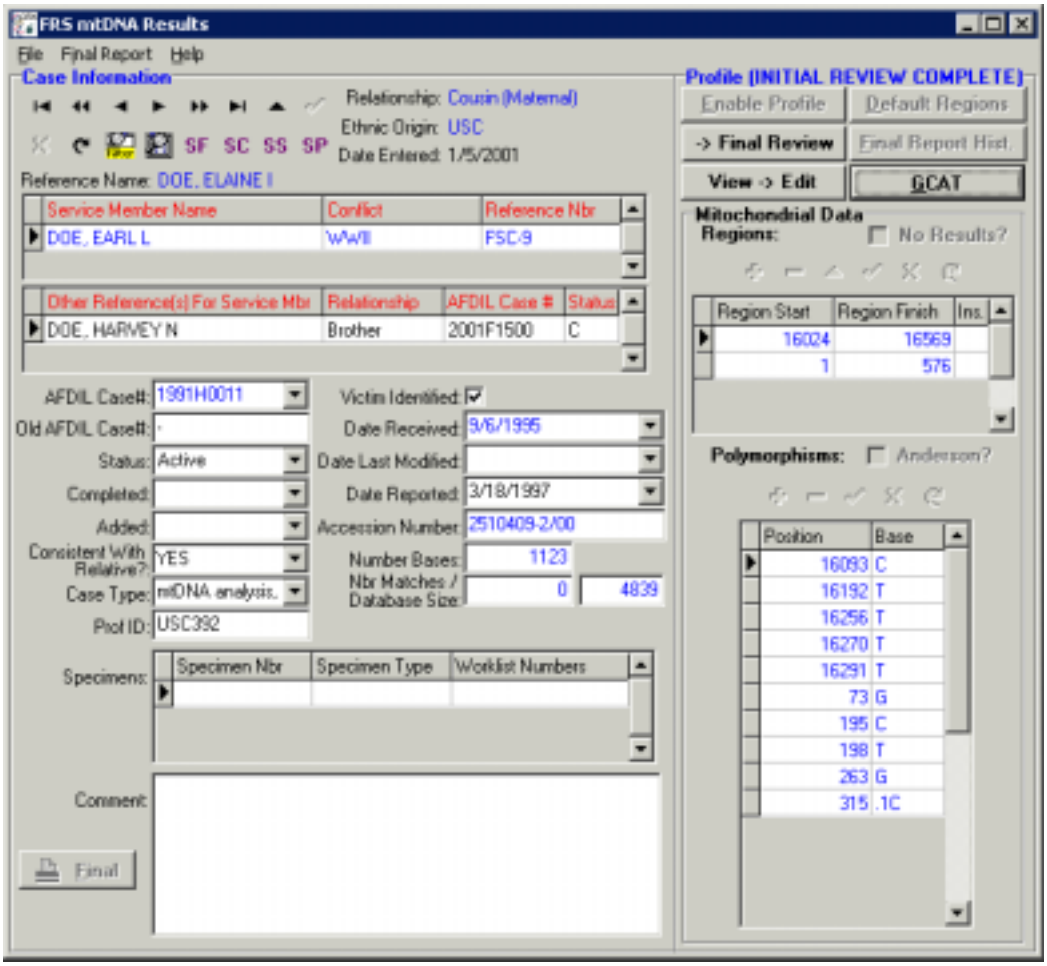

Figure 3. A screen shot of the Family Reference Specimen (FRS) mtDNA results screen during review mode. Information must be reviewed by a second scientist.

module follows the same review and color-coding hierarchy as that of the FRS casework. Sequence data can be either imported or manually entered and reviewed online in a similar fashion to LISA-FRS. It is important to note however, that review credentials for degraded skeletal remains are assigned independently from those granted within the family reference system, as more experience is needed in interpreting this type of data.

Once data have been entered and reviewed in LISA FRS and CILHI modules, they can be compared. Two tools within LISA are used to manage and perform the comparisons. LISA Case Inquiry (LISA-CI) is the tool used to manage the comparison of reference specimens to profiles obtained from unknown material. Within the module, LISA can be used to compare the reference(s) for a specific unknown individual(s) to a case or group of cases. The comparison request is then submitted via email to all scientists involved in the case. The analyst assigned the role of "comparison" is sent a notification to perform the request. The sequence comparisons are processed within LISA using a set of sequence comparison tools called "`g:cat`". `G:cat` allows users to compare individual sequences, groups of sequences, cases, and entire databases in virtually all possible scenarios. In the case of comparison requests, users select the comparison they wish to perform and all specimens are automatically loaded into `g:cat` (**Figure 4**). Within a few clicks of the mouse, the comparison is complete and the report is generated. Using this tool, hundreds of unknown mtDNA profiles can be compared to the population and/or missing-persons databases. **Figure 5** shows the main `g:cat` user interface (UI). From this UI users can manually enter profiles or choose a sample from the database to search. The results of the search are displayed with the differences highlighted and underlined in red.

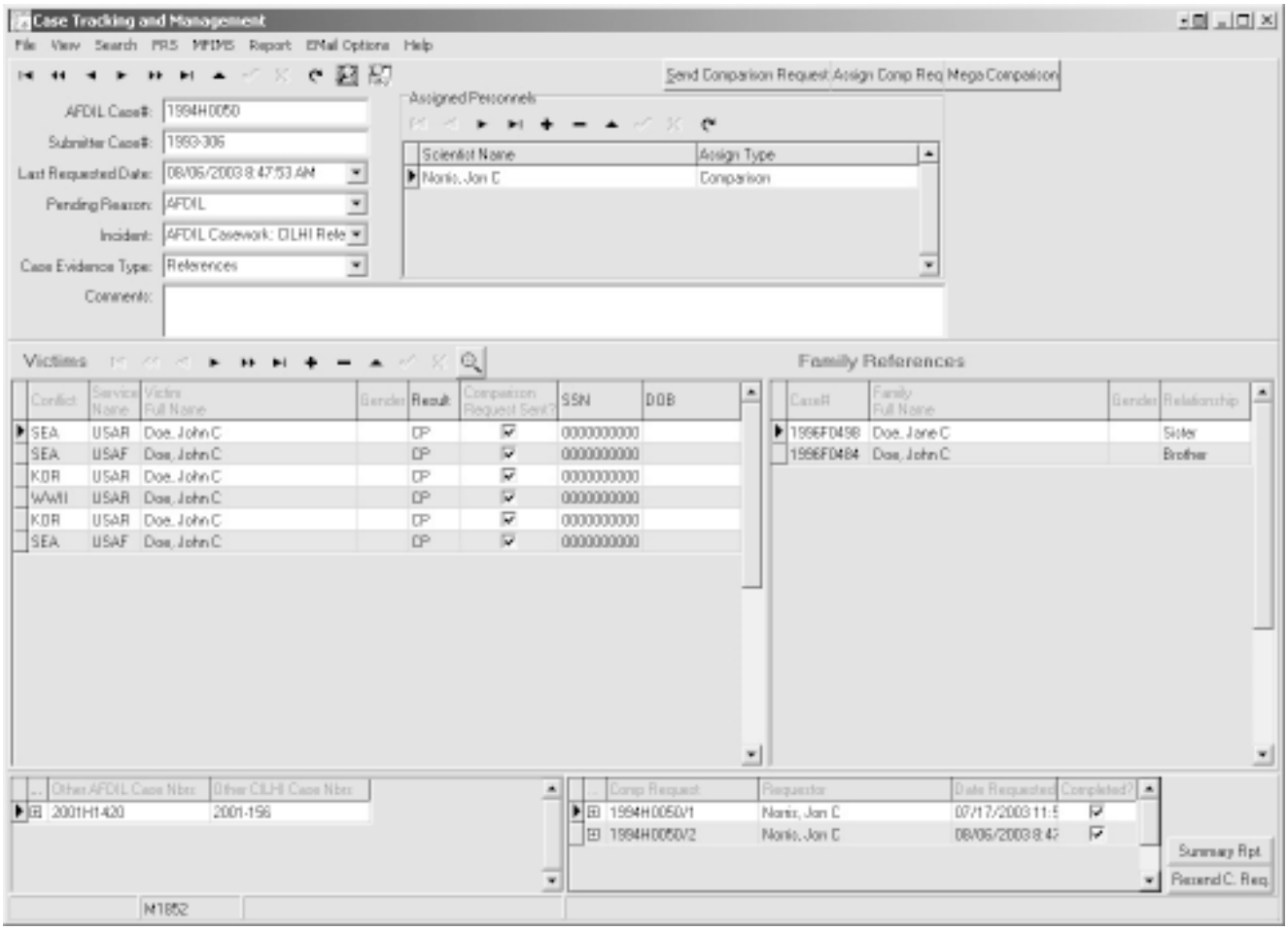

Figure 4. The "Case Tracking and Management" module in LISA. This screen allows the analyst to view the comparisons that have been requested for a particular CILHI case. Underneath the top portion of the screen where the case number and the assigned scientists are listed is the "Victims" section, where the missing persons needed for comparison to a specific CILHI case are listed along with the case numbers of the associated family references. Scientists can also access summaries of all comparisons for a particular case or specimen.

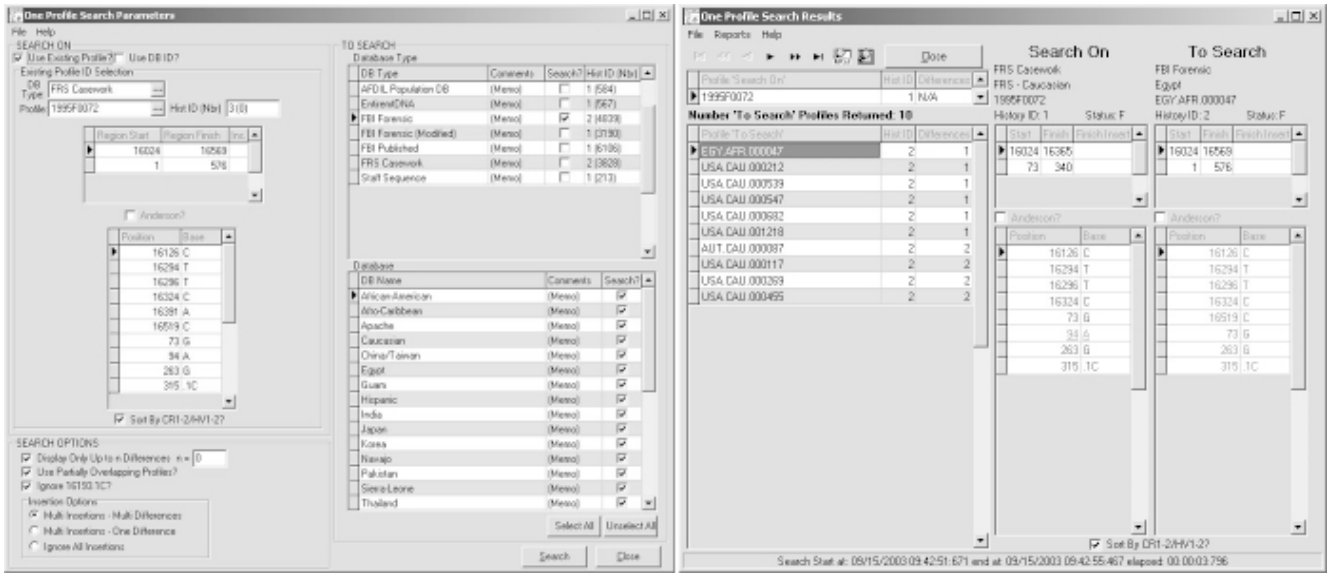

Figure 5. A screen shot of the primary `g:cat` screens in LISA. On the left is a screen shot of the sequence to be compared and the databases to be compared to. On the right is a screen shot of the results. Note that the differences between the searched sequence and the "match" are underlined and highlighted in red.

G. Instrumentation

On their first exposure to robotic systems, forensic scientists are often surprised to observe that the robotic instruments as a rule do not process samples any faster than could a human analyst. The great advantages of well-designed robotic automation, especially when coupled to informatics systems such as LISA, are the consistency of results and the virtual elimination of the potential for sample switching. At the same time that human error is eliminated from routine manipulations, scientists are freed to focus on the critical aspects of data analysis and review, which tend to be the rate-limiting steps in forensic testing. Thus, robotics lend a speed to the process that is greatly in excess of their rather deliberate pace in the actual laboratory manipulations.

While we at AFDIL have managed to streamline our protocols and laboratory design to allow for the rapid manual processing of degraded skeletal remains, it is with the advent of robotics that we believe the processing of low-copy-number mtDNA will truly move into the realm of high-throughput. This can be seen in how rapidly blood family references can be processed. As the Wallac DBS hole-puncher, the Qiagen BioRobot® 9604 (**Figure 6**), and the ABI PRISM 3100® Genetic Analyzer (**Figure 7**) have been implemented in regular usage, we have managed to completely erase the backlog of over 2000 family references. Further implementation of robotics, such as use of the Tecan Genesis RSP 200 (Tecan: Research Triangle Park, NC) for sequencing of high-throughput samples, will increase production. AFDIL's Research Section has been using the Tecan Genesis RSP 200 for cycle sequencing, post-cycle sequencing cleanup, and addition of loading buffer in the processing of total mtDNA genome processing. This has greatly increased the throughput of large-scale research sequencing projects. Not all of these specific instruments are applicable for use in degraded specimen processing; however, the ABI PRISM 3100® Genetic Analyzer is currently being validated for CILHI casework and is anticipated to increase our output of samples.

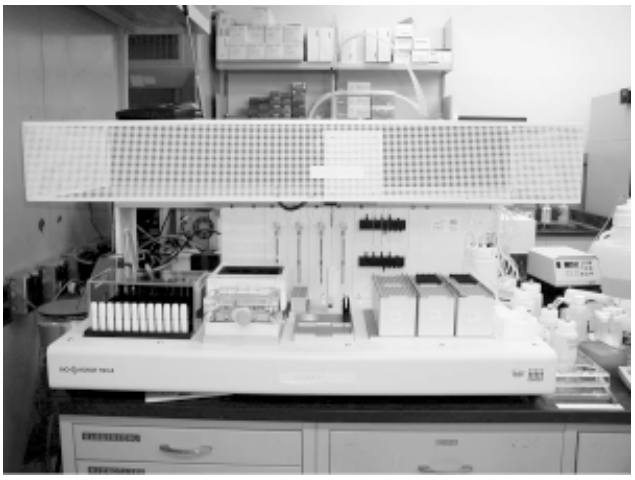

Figure 6. The Qiagen BioRobot® 9604 (Qiagen: Venlo, The Netherlands) used for the high-throughput processing of family reference samples.

Figure 7. The ABI PRISM 3100® Genetic Analyzer used to generate sequence data for high-throughput processing of family references. Validation is in progress for usage on sequencing degraded skeletal remains samples.

H. Personnel Management

With budget cuts and competition among laboratories, it is generally a challenge to retain skilled personnel. With high-throughput testing, this challenge becomes even greater. One approach for large-scale testing that has proved successful for AFDIL is to divide the staff into teams. Tasks can then be compartmentalized so that each team is responsible for a particular part of the processing. **Figure 8** displays a flowchart describing how the staff is organized in addition to their task assignments.

The addition of several supervisory analysts creates an additional level of oversight for the cases. This is useful because much of the time spent processing a case is actually troubleshooting, which includes determination of starting template amounts, proper usage of primer pairs, and data interpretation (i.e., heteroplasmy vs. mixtures vs. background sequence data). The designated sequencing team is responsible for all post-PCR laboratory activity. This allows the analysts and supervisors of the other teams to focus on managing the cases. Separate technicians designated to perform PCR are given assignments from the analysts who are interpreting the generated data. This type of management has also proved to decrease the time required to train and limits the number of personnel in a designated laboratory space, as the technicians are more specialized. With this compartmentalization of tasks also comes the ability to manage a larger staff, increasing efficiency to an even greater level over a more traditional approach.

The addition of a sequencing team and PCR technicians increased the staff by approximately 30%. Subsequently, the output of specimens increased by 300%. **Figure 9** is an evaluation of the number of sequencing reactions completed per month in a six-month time period. One of the largest complaints from analysts working in a high-throughput environment is the inability to follow a case from start to finish. With this compartmentalization of tasks, the analysts can spend an increased amount of time trouble-shooting the cases. This in turn gives a sense of "ownership" and pride toward the case, thereby increasing both the quality of the reported data and employee satisfaction.

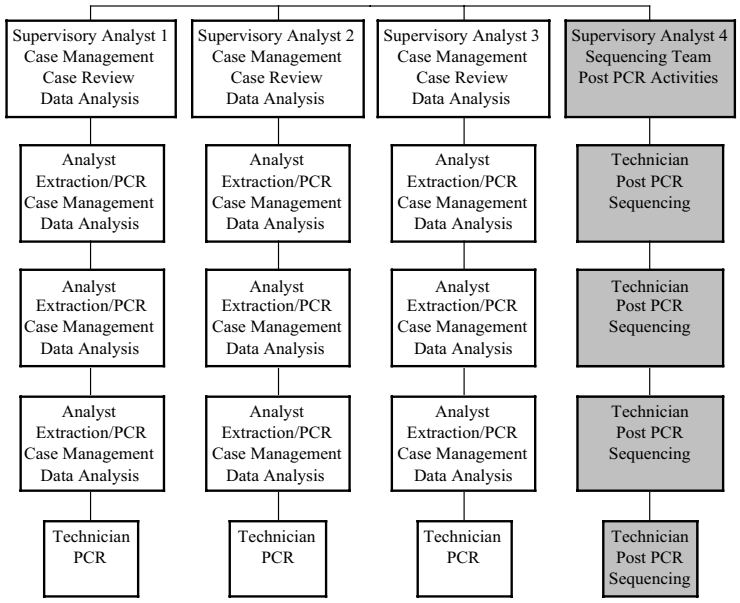

Figure 8. Organizational chart representing the division of tasks among the AFDIL staff.

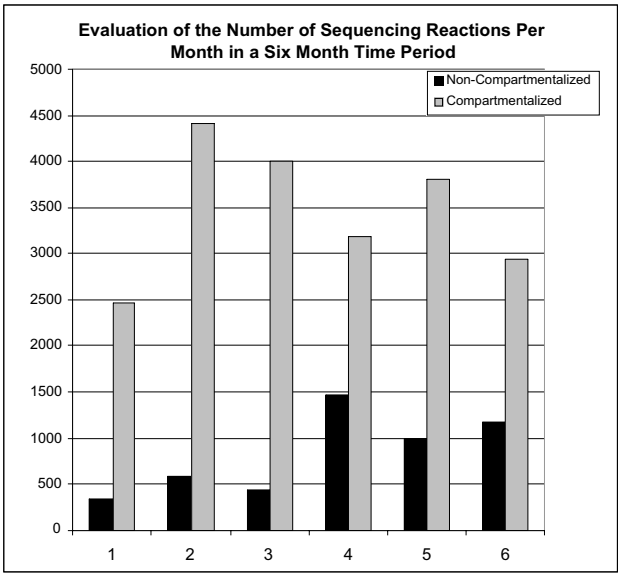

Figure 9. Graph depicting the change of productivity when staff was increased by 30% and reorganized to a high-throughput style of management.

III. METHODOLOGICAL CHALLENGES

A. Sample Type

Since 1994, AFDIL has processed more than 2000 ancient skeletal remains samples as part of the nation's efforts to identify missing service members from World War II, the Korean War, and the Vietnam War. Over 1000 remains submitted by CILHI processed during a three-year time period were examined to determine if there was a general trend among success rates versus specimen types. The empirical data suggest that when possible, prudent sample selection may save time and resources by decreasing the need for retesting/resampling. These samples were processed using techniques as outlined in Section II. Bones received only as "long bones" implied that the bone was too fragmented to make a determination of specimen type and were excluded from the study. Overall, the study showed that 80% of the specimens evaluated produced reportable sequence data. For the purpose of this study, "reportable" is defined as any amount of base pairs obtained that was reproducible either by multiple amplifications or multiple extractions. If a mixture was obtained, the specimen was reported as inconclusive and thus considered not successful. These results are displayed in **Table 2** and **Figure 10**.

Of the long bones, the weight-bearing bones, such as femora and tibiae, were the best specimen types. Metatarsals are also weight-bearing bones, but at initial glance they appear to be inadequate specimens by size alone. However, approximately 80% of the metatarsals

Table 2. Summary of the success rates of each skeletal element tested at AFDIL during a three-year period

Specimen type	No. tested	No. reportable	Success (%)
Cranial	76	39	51.32
Os Coxa	19	12	63.16
Ulna	54	36	66.67
Fibula	21	14	66.67
Radius	37	26	70.27
Mandible	34	26	76.47
Clavicle	30	23	76.67
Scapula	19	15	78.95
Humerus	149	118	79.19
Tooth	184	147	79.89
Metatarsal	21	17	80.95
Vertebra	14	12	85.71
Tibia	145	129	88.97
Femur	192	182	94.79
Rib	26	25	96.15
Total	1021	821	80.41

tested produced reportable results. Ribs were also highly successful, although a larger number of specimens is needed in order to make a suitable determination on their overall success rate as most of the rib samples used for this study were from a single CILHI case. The spiny processes of the vertebra (i.e., spinous processes and transverse processes) that provide an area for muscle attachment are

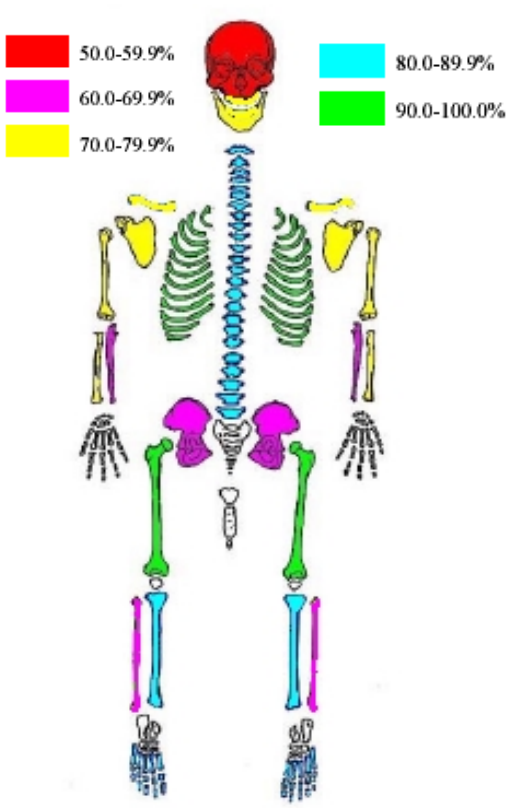

Figure 10. Illustrative representation of the data presented in Table 2.

better than the porous body. Cranial fragments are the most difficult of the samples tested from which to obtain quality sequence data. Cranial fragments have large porous areas of diploic bone that are exposed to environmental elements and modern contaminants. These crevices cannot be cleaned during sanding unless the porous portions are removed altogether, leaving only thin plates of cortical bone. Larger cranial fragments (over 2.0 g), or those cut from more compact portions of the skull may prove to be more favorable specimen types. In general, when sampling ancient skeletal remains, sections that appear to have the densest cortical bone should be sampled, whenever possible.

B. Degraded Specimens

Although most samples processed for mtDNA analysis at AFDIL are considered degraded, some are clearly more challenging than others. The majority of remains submitted to AFDIL for testing have been subjected to harsh environmental elements for at least 30 years. Some individuals were involved in aircraft incidents, which contributed to the highly fragmented nature of the remains. Some remains have been subjected to years of burial in highly acidic or basic soils or trapped in pockets of jet fuel or saltwater. The amplicon size of the standard primer sets (PS) as shown in Figure 2 is approximately 300 base pairs. In contrast, the mini-primer sets (mPS) are approximately 150 base pairs in amplicon size. Mini-primer sets were developed in order to increase the success rate of obtaining reportable data [25]. Since mini-primer sets have been implemented, over 50% of the cases have required the use of at least one of these primer pairs. Theoretically, as the amplicon size decreases, the number of available authentic degraded mtDNA fragments increases. On the other hand, the number of available modern mtDNA contaminants of the same fragment size should remain constant [25,29,55]. However, as the sensitivity of the reaction increases, caution must be taken when interpreting these results. Reproducibility is critical to the analysis of all ancient DNA specimens, and results should be supported by either multiple amplifications, primer set overlap, or multiple extractions [71].

C. Contamination

Contamination is a constant consideration when processing ancient remains for mtDNA analysis. However, the use of multiple controls, appropriate lab design, control-tracking systems, protocols that require reproducibility, and validated procedures designed to remove modern contaminants are all tools to enable the laboratory to produce reliable, authentic results [14,36].

1. Controls

Because of the complexities of ancient mtDNA processing, proper interpretation of extraction and PCR reagent controls is of the utmost importance. A laboratory should determine its own contamination threshold and develop guidelines for interpreting data obtained from these reagent controls. The reagent controls and the samples should be treated alike for all processing steps. Data generated for controls cannot be interpreted if the sequence quality is not at or above the standard set for sample quality.

AFDIL uses a method of detecting contamination by visualizing amplification product on a 2% agarose gel stained with EtBr. Visible bands are carried on to sequencing. Although the possibility exists that sequencing results could on occasion be obtained from a sample or control below this visibility threshold, these results are considered unreliable due to the known presence of sporadic contamination in ancient DNA testing. Authenticity of a sequence generated from a nonvisible band cannot be verified since, in our experience, such a result is not reliably reproducible. With the influx of changing technology and increased sensitivity for sequencing mtDNA, we felt it prudent to revisit the current contamination threshold to determine if adjustments were needed. A two-month timeframe was evaluated where controls that fell below the threshold were carried through to sequencing and data analysis, to determine the effect this would have on the interpretation of controls and the production of reportable sample data. This procedural change increased AFDIL's workload by approximately 45% at a cost of $37.00 per sample (a total of approximately $77,000), which excludes the cost for increased labor and review time. A total of 2097 controls consisting of hypervariable regions, primer sets, and mini-primer sets were evaluated.

None of the 1306 controls amplified using hyper-variable region or primer sets produced readable sequence when run on an ABI PRISM® 377 DNA Sequencer. A total of 10 controls (1.26%) out of 791 amplified using mini-primer set parameters produced readable sequence (**Table 3**). Of these 10 controls, 9 sequenced in one direction only and 6 were consistent with both the rCRS and the associated sample, which is, incidentally, the most common Caucasian type, 263 G and 315.1 C, seen in approximately 7% of the population [1,13]. More importantly, in no instance was the reporting or interpretation of the case affected by sequences produced by these controls that fell below the visibility threshold. Thus, we have verified that using the same thresholds for processing and interpretation for both samples and controls is sufficiently conservative and appropriate. As a cost/benefit analysis, the outcome of our trial was abundantly

Table 3. The number of negative and reagent blank controls that produced readable sequence data after observing a blank sample on a 2% agarose gel stained with ethidium bromide (EtBr) (Data are summarized from two months of testing the applicability of processing all controls.)

Parameter	Number and percentage	Negatives	Reagent blanks	All controls
HV region primers	Total processed	126	63	189
	# with readable sequence	0	0	0
	% with readable sequence	0.0	0.0	0.0
Primer set	Total processed	762	355	1117
	# with readable sequence	0	0	0
	% with readable sequence	0.0	0.0	0.0
Mini-primer set	Total processed	530	261	791
	# with readable sequence	9	1	10
	% with readable sequence	1.70	0.38	1.26

clear, with the result that, except in criminalistic casework, AFDIL does not sequence controls if they do not produce PCR product visible on agarose gels. For the sole purpose of adhering to SWGDAM guidelines, these controls are still sequenced for criminalistic casework.

2. Staff Databases

A laboratory should have mitochondrial DNA sequences on file for all staff that enters the extraction/amplification laboratories for ancient DNA [14]. All reagent preparation staff should also be included in this database. This will enable the laboratory to better identify a source of contamination. In addition, laboratories should provide limited access to the work areas for nonscientific personnel. Ideally, access to the reagent preparation laboratory should be prohibited to anyone other than the reagent prep staff, as is the AFDIL policy.

3. Lab Design

Separation of pre-PCR and post-PCR laboratory space is essential [14]. PCR product contaminants from post-PCR laboratories can easily be carried on the clothes or person of a scientist, in reagents, or on equipment, and contaminate pre-PCR areas, potentially causing the generation of false sequences. The division of tasks among the laboratory staff limits the travel of individuals between each type of space. Physical separation of the processing of samples possessing low and high quantities of DNA is also important. At AFDIL, the laboratories for processing family references or fresh tissue are kept completely separate from those processing degraded skeletal remains. Not only are the physical spaces kept separate, but separate reagents are dedicated to each type of DNA analysis.

4. Contamination Tracking

A tracking system can be incorporated in order to determine the level of contamination in the laboratory. AFDIL uses the LISA system to track contamination and determine when corrective action is needed. For example, if random contamination increases in frequency, a complete cleaning of the pre-PCR laboratories with 10% bleach can be implemented ahead of the regular biannual schedule. **Figure 11** displays the various tracking options our managers have available to them through LISA.

As expected, the more sensitive the system, the more likely a control will produce a positive result. Reproduction of the sample results is a powerful tool for ensuring the authenticity of the data when combined with conservative treatment of samples.

IV. POPULATION RESOLUTION

A. Control Region Analysis Versus Hypervariable Regions I and II

The circular genome of the mitochondria consists of approximately 16569 base pairs. Of particular interest to forensic identification work is the ~1100 bp of the noncoding control region. At AFDIL, the entirety of the control region is sequenced only for family references or population database samples. This is done with a single

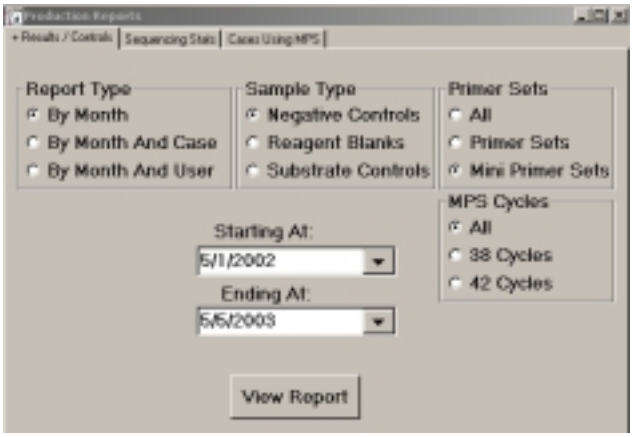

Figure 11. Screen shot from the tracking system built into LISA.

initial amplification of each of two extracts using the primers indicated in Figure 2. Sequencing involves a series of primers designated with an indicator in Table 1. These primers produce overlapping fragments of target mtDNA. We find this to be an efficient manner with which to generate the desired mtDNA type of the individual.

Degraded skeletal remains, however, are not treated in the same fashion, as already indicated in Section II. Standard procedure is to only generate sequence for the Hypervariable Regions I and II (HVI and HVII) as this is where the greatest amount of variation from the rCRS lies. Di Renzo and Wilson [21] examined Middle Eastern and Sardinian populations and found that the majority of the variation in the Control Region lies within the first 400 bp of the region studied, which roughly equates to HVI. We have found a similar result when examining 4021 unrelated individuals in our family reference database, with 41.1% of the variation from the rCRS found in HVI and 33.7% in HVII (**Figures 12–14**). This is generally sufficient variation to differentiate between individuals in a closed population.

In instances where the sequence found in HVI and HVII for an individual is not adequate to differentiate individuals in a group, other areas in the control region are examined. For AFDIL, this entails sequencing mini-variable regions I and II (mVRI and mVRII). These designators mVRI and mVRII are internal classifications for the regions between np16471–np16561 and np424–np548, respectively. MVRII falls within the ranges of what has been designated as HVIII in the literature [48], whereas mVRI has no alternate designator as it lacks a significant amount of variation other than at np16519. However, the additional information afforded by these regions often provides a greater power of exclusion.

There are particular polymorphisms that allow for the division of individuals into defined mtDNA haplogroups. The study of variation in the mtDNA coding region is generally necessary for unambiguous haplogroup determination. In the mtDNA Section at AFDIL, we are not concerned with haplogroups on a regular basis; however, there are circumstances where the classification of such can be advantageous. A number of researchers, i.e., [33,34,46,50,54], have detailed mtDNA polymorphisms characteristic of every haplogroup, most of which occur outside of the control region. **Table 4** presents publications that describe the mtDNA haplogroup characteristics of indicated populations. Section V, Phylogenetics and Forensics, will describe in more detail a specific case instance of the usefulness of haplogroups.

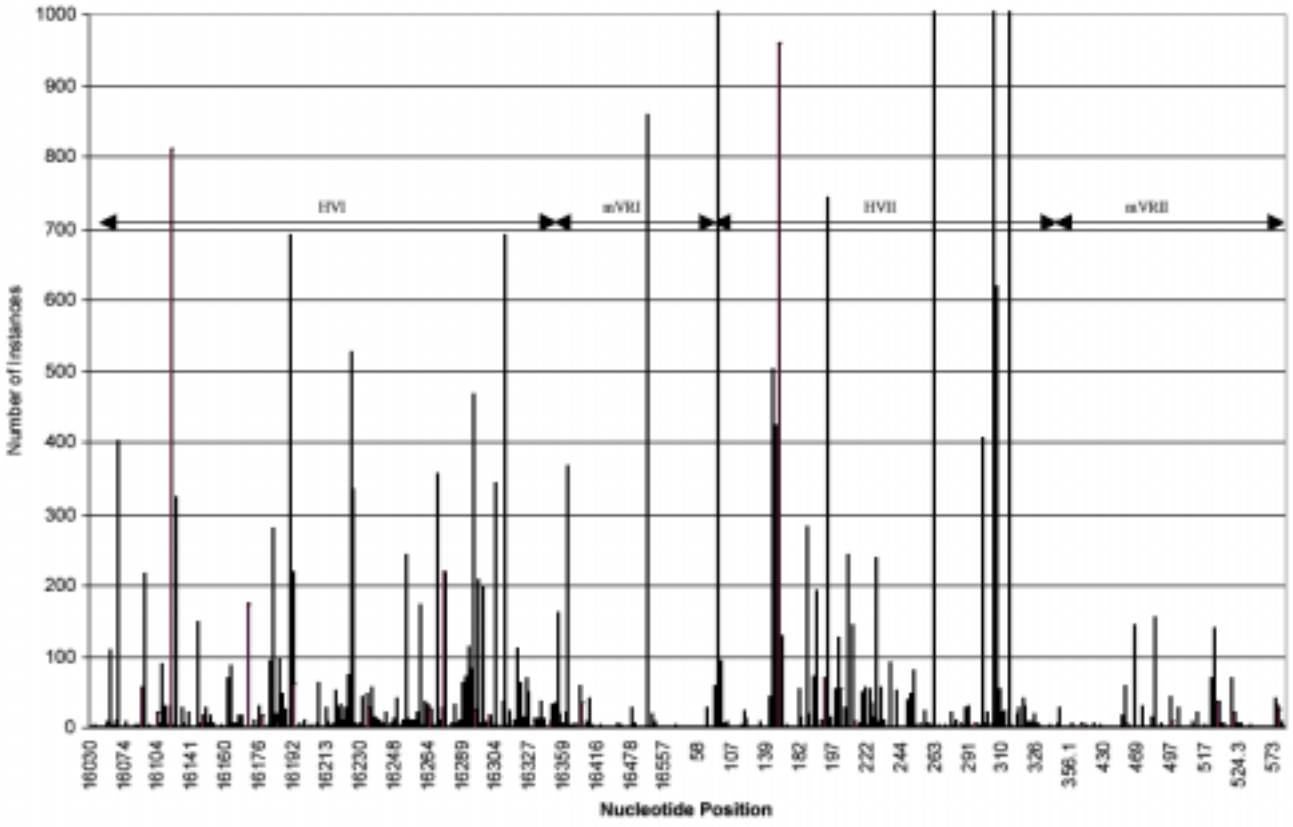

Figure 12. Distribution of the polymorphic positions observed in 4021 unrelated individuals in the AFDIL `g:cat`. Positions 73, 263, 309.1, and 315.1 have values greater than the scale of the chart with 2367, 3987, 2121, and 4007 instances, respectively. Numbering follows the rCRS [7].

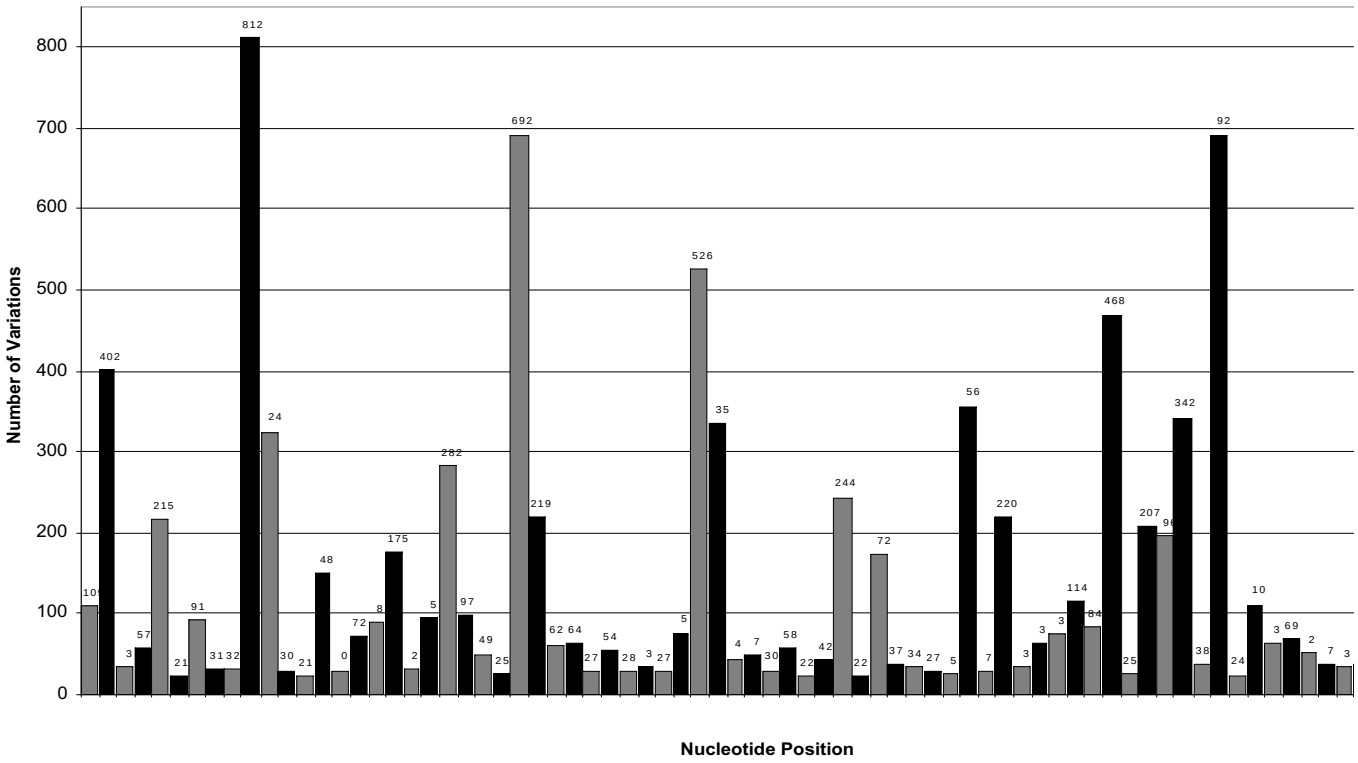

Figure 13. Nucleotide positions in Hypervariable Region I (np16024–np16390) that exhibit variation from the rCRS for more than 20 instances in the AFDIL family reference database of 4021 unrelated individuals.

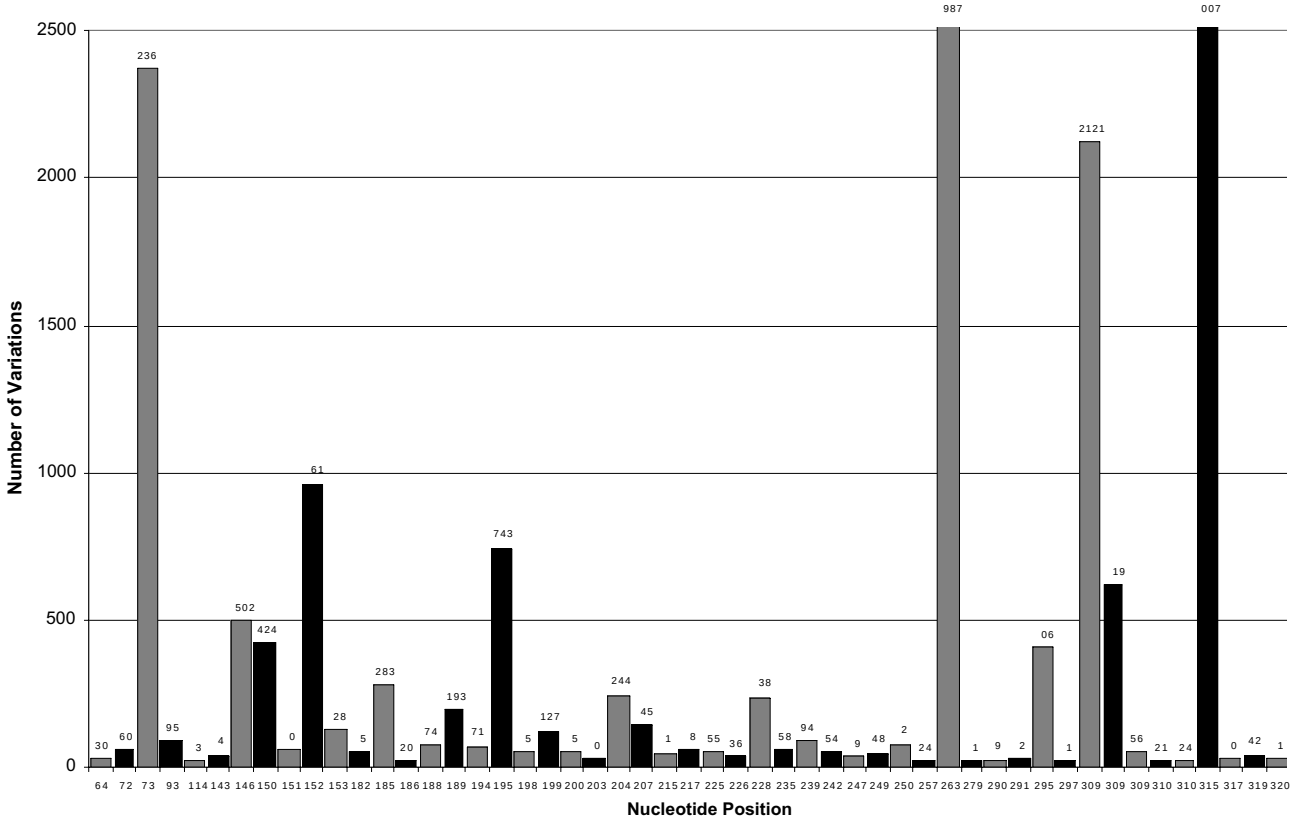

Figure 14. Nucleotide positions in Hypervariable Region II (np35–np369) that exhibit variation from the rCRS for more than 20 instances in the AFDIL family reference database of 4021 unrelated individuals.

Table 4. Publications that describe the mtDNA haplogroup characteristics of indicated populations

Region	Country/Area	Ref.
Africa and Middle East	Northwest Africa (Senegalese)	[59]
	Senegalese Mandenka	[26]
	Sub-Sahara	[65]
	Yemen Hadramawt	[62]
Asia	Aleutians	[19]
	Japan	[41]
	Korea	[45]
	Siberia	[18]
South America	Brazil	[4]
	Columbia	[16]
	Columbia and Costa Rica	[15]
	Ecuador	[63]
	North-Central Mexico	[27]
Europe	Austria	[56]
	Finland	[52]
	Germany	[58]
	Iceland	[31,32]
	Italy	[24]
	Near East and Caucasus	[50]
	Norse and Gaelic	[30]
	Roma (Gypsies)	[28]
	Spain	[68]
	Switzerland	[20]
	Western Europe	[61]
North America	Native Americans	[11,12,47]

B. Databases

The Federal Bureau of Investigation (FBI) distributes a searchable mtDNA population database (CODIS^{MT}) that allows researchers and forensic scientists to search a generated sequence of polymorphic differences from the rCRS against a database of forensic and published sequences to determine the rarity of that profile [51]. The forensic database of CODIS^{MT} includes only those sequences generated by laboratories that meet the standards of the FBI and sequences that are physically found to be on file in those labs. The published database of CODIS^{MT} contains sequences from scientific literature. The latter contains vast amounts of mtDNA sequence data derived from all types of studies, but is of limited utility for forensic purposes. Moreover, the published sequences have no uniform standards regarding sequence quality or the range of reported sequence. Most of these sequences consist of just HVI data. For these reasons, AFDIL refers only to the forensic databases for inferring the significance of mtDNA matches.

AFDIL previously performed its population database searches with CODIS^{MT}, but recently, as described in Section II, we have moved to an in-house program, `g:cat`, that is an integral part of LISA. With this program, databases that include AFDIL generated family references and population sequences, as well as the forensic and published sections of CODIS^{MT}, can be searched.

C. Closed Populations and Identification

It should be emphasized that mtDNA sequence data alone are not a means of positive identification for an individual. Given that mtDNA is a maternally inherited genome, all descendents of a particular maternal line will have the same mtDNA type, barring mutation. CILHI uses the mtDNA sequences generated at AFDIL in conjunction with anthropological and circumstantial evidence to make the overall identification. As samples are processed in the blind, scientists at AFDIL do not know to whom a specific case is related. Once a sequence has been generated, CILHI is notified. With information such as casualty data, site reports and anthropological analysis of the remains, a list of missing persons potentially associated with the case is assembled and returned to AFDIL for comparison to family reference sequences, thereby making mtDNA an integral part of a complete investigational package.

Oftentimes, CILHI already has extensive evidence regarding who is represented in a set of remains. This is especially true in the case of aircraft losses where the flight crew is known. However, in the case of battlefield recovery operations, the number of individuals potentially corresponding to a case could number into the hundreds. It is not uncommon that sequences from a case sample match multiple reference individuals for HVI and HVII. In these instances, mVR testing of the sample is undertaken in a further attempt to discriminate among the family references. Based on AFDIL's criteria for exclusion, references that differ at two or more positions from the case sequence may be excluded in most instances. We do not include length heteroplasmies of the HVII C-stretch, point heteroplasmies, or the mutational hotspot of 16093 in these differences [36].

MtDNA sequence data are most effective for identification in the cases of closed populations of individuals; i.e., in cases where there is a set number of missing people involved with a specific recovery operation. This is especially true in the case of aircraft losses, as mentioned above. A case that illustrates how comparisons are made at AFDIL and how sequence data can be interpreted in a closed population is that of the B-52 bomber crew shot down during the Vietnam War.

In the winter of 1972, in a final effort to bring the Vietnamese to peace talks, the US command initiated a campaign of intense bombings designated Operation Linebacker II, or the "Christmas Bombings" of Hanoi. On the night of 20 December and into the early morning of 21

December, several B-52s made heavy bombing runs of the Hanoi region. Four of the B-52s and one A-6A were lost that night over Vietnam, with a total complement of 27 airmen. Twelve men were captured and later returned as POWs. The remains of the other 15 were not recovered at that time.

In 1985, CILHI conducted the first joint recovery with the Socialist Republic of Vietnam (SRV) of a B-52 crash site where four members of an aircrew had perished [17]. In 1986, the SRV repatriated remains presumed to be associated with the earlier recovery. Five samples were cut from the remains recovered and submitted to AFDIL for testing in 1996 and 1997. The samples consisted of portions of a cranial vault, rib, femoral head, calcaneus, and metacarpal. Full HVI/HVII sequence data were generated for all five samples. In addition, the sequences from the samples were all consistent with each other. Family reference blood samples were acquired for three of the airmen presumed to be associated with the case. Unfortunately, no reference was available for the fourth man (Maj. Irwin S. Lerner) except for five stamped letters. (As mentioned in Section II, envelopes and stamps must be interpreted with caution, as the source of the DNA cannot be ascertained.) However, as nothing else was available as a reference for Maj. Lerner, testing was performed on four of the envelopes.

Sequences generated for the envelopes as well as the mother of one of the other missing airmen (Lt. Col. Randolph A. Perry, Jr.) matched the sequences generated for the skeletal remains. The other two crewmen were excluded. The question at this point was whether the sequence from the stamps was an authentic reference for Maj. Lerner, if Lt. Col. Perry had sealed the envelopes, or if the sequence was extraneous. At the time of the testing in 1997, the case was placed into "inconclusive" status due to the fact that we could not differentiate between these two individuals.

In the fall of 2002, AFDIL received a blood reference from a maternal cousin of Maj. Lerner. The blood sample was rapidly tested and found to be consistent with the sequence generated from the envelopes. Therefore, the sequences generated from the self-reference envelopes were authentic references for Maj. Lerner. As the Lerner and Perry mtDNA lineages still could not be differentiated, mini-variable region testing was undertaken on both the remains and the family references. Each referencwas identical in the mVRI region. However the reference sequences for one of the airmen had a transversion at np512, a relatively rare polymorphism that was lacking in the other family reference. The sets of bones showed two different sequences, one that matched the self and family references for Maj. Lerner, and the other that matched the family reference for Lt. Col. Perry. Nonetheless, due to the single base difference between reference and sample, the case still required further consideration.

It is well established in the forensic mtDNA literature that exclusions are generally not made on the basis of a single base-pair difference between a reference and evidentiary sample, when these are from maternal relatives separated by one or more generations. This is because there is a genetic bottleneck between generations that can permit differential segregation of even low-level heteroplasmy from mother to child. Rarely, this can manifest itself as an apparently complete "mutation" from one generation to the next (the details of intra- and intergenerational mtDNA mutation and heteroplasmy segregation are beyond the scope of this paper; for a more complete discussion relating to forensics, *see* Holland and Parsons [36], and references within). In this case, however, we were dealing with a "closed system" where we seek to ascribe each distinct skeletal sequence to one of two families, all other families having been excluded. Under these conditions, the evidence associating each airman with his respective family (i.e., the family with an exact match, rather than the one differing at position 512) is greatly strengthened. This is because the alternate hypothesis, where the skeletal remains sequences are actually from the maternal lineages that mismatch at position 512, would require a minimum of two independent mutations at that position. This argument reasonably assumes that, due to the absence of a genetic bottleneck, there has not been a fixed mutation between skeletal elements from the same individual. This requirement for two independent mutations is substantively the same as having sequences that differ at two positions, meeting AFDIL's requirement for exclusion. For position 512, this evaluation is all the more conservative because it involves an A-C transversion that is not observed within our control region database of 849 Caucasian individuals (although this transversion is listed in a compendium of polymorphisms maintained at www.mitomap.org), and therefore clearly not prone to mutation.

The rarity of the mtDNA sequence and the collective anthropological and circumstantial evidence permitted a segregation of the skeletal elements into those of Maj. Lerner and Lt. Col. Perry.

V. MITOCHONDRIAL DNA PHYLOGENETICS AND FORENSICS — SEPTEMBER 11, 2001 AS A CASE STUDY

The typical mitochondrial case does not end with the generation of a sequence. The questioned sample is usually compared to a reference sequence (presumably, a maternal

relative) and, to evaluate the significance of the evidence, a relevant population database [36]. Although sequences from all over the world have been added to the growing forensic databases, some geographical regions are greatly underrepresented. This was the case for mitochondrial analyses performed by AFDIL after the terrorist attacks of September 11, 2001.

Nuclear DNA testing (along with dental records and fingerprints) of the remains from the victims aboard American Airline (AA) Flight 77 and within the Pentagon was useful for identifying 178 of the 183 victims. Five missing individuals (four within the Pentagon and one aboard the airplane) could not be identified due to lack of biological material from the crash. Five remaining nuclear STR profiles were obtained from the crash site that did not match any references for the victims. These profiles were thought to represent the terrorists aboard the flight. The 40 victims aboard the United Airline (UA) Flight 93 that crashed near Shanksville, PA, were also identified by nuclear DNA testing, dental records, and fingerprinting. Four nonmatching nuclear DNA profiles were also obtained from the crash site and again tentatively ascribed to the terrorists.

The DNA results strengthened the hypothesis that two of the terrorists were brothers, as indicated by other evidence. Two of the terrorist STR profiles aboard the AA Flight 77 gave a sibling index greater than 500. To further test the hypothesis of maternal relatedness, AFDIL sequenced the HVI and HVII regions of mtDNA for these individuals. The sequences generated did match in HVI and HVII, which is consistent with a maternal relationship between the two men.

To determine the significance of this mtDNA evidence, the sequence was compared to sequences in a global forensic database of 4142 individuals [51]. A total of 6 individuals matched the same sequence of the putative brothers in the database (a frequency of 0.14%). Specifically, the 6 matches occurred within the 1773 Caucasian component of the database (0.34%). Sequencing of the mtDNA of the remaining three terrorists was also performed. None of these sequences matched any of the 4142 individuals in the global forensic database (**Table 5**). However, this database contained only a few Middle Eastern sequences, and therefore was of little use in determining the frequency in the population from which the terrorists were thought to originate.

A recent paper by Richards et al. [60] examined the contribution of Near Eastern mtDNA to the European founder types. This paper lists mtDNA HVI sequence data from 1088 individuals of Near Eastern origin. (The data used in this paper are available on the Web for download at http://www.stats.ox.ac.uk/~macaulay/founder2000/index.html.) A search of this Near Eastern database for the terrorist sequences revealed that 25 individuals matched the putative terrorist brothers (2.3%). The remaining putative terrorist sequences from the AA Flight that crashed into the Pentagon were also represented in the Near Eastern database (Table 5).

Table 5. Number of hypervariable I region matches of the putative terrorist sequences from the American Airline (AA) Flight 77 (Pentagon) and United Airline (UA) Flight 93 (Shanksville, PA) disaster sites from the September 11, 2001 event (The global, forensic database of 4142 sequences and a Near Eastern database from 1088 published sequences were used in this study.)

		Database	
Event	**Sequence**	**Global (n = 4142)**	**Near Eastern (n = 1088)**
AA Flight 77	Sequence 01/02	6	25
	Sequence 03	0	2
	Sequence 04	0	1
	Sequence 05	0	1
UA Flight 93	Sequence 01	0	0
	Sequence 02	0	0
	Sequence 03	0	0
	Sequence 04	0	0

Mitochondrial DNA testing was also performed on the four putative terrorists from the UA flight 93/ Shanksville, PA crash. None of the terrorist sequences matched any sequences in either the global forensic database or the Near Eastern database (Table 5). Lacking matching sequences in any population database, we sought to access the phylogenetic information inherent in the mtDNA sequences to determine if that could suggest additional information as to their origin.

We used the tools of phylogenetic analysis to examine the relationships of the four "unique" terrorist sequences. Phylogenetics is a method of classification that utilizes hypotheses of character changes in the evolution of a biological system (in this case, mtDNA sequence data) to group "taxa" (in this case, individuals) into a hierarchy of nested sets grouped by the presence of shared-derived characters (e.g., Swofford et al. [67]). The interpretation of these relationships is depicted as a phylogenetic tree. These trees are commonly generated by software packages, such as PAUP* version 4.0 [66] or Winclada version 1.0 [53].

Our basic experiment used the sequences of the four putative terrorists from UA Flight 93 together with a restricted sample of published Near Eastern sequences [60] to include representatives of all major Near Eastern haplogroups. For example, haplogroup U1b is associated with the HVI sequence motif 16249 T-C and 16327 C-T. This data matrix of putative terrorist sequences and

sequences from Near Eastern mtDNA haplogroups was used to build a parsimony tree using Winclada [53]. In evaluating a phylogenetic tree, one is guided by the principle that sequences closely related to each other tend to group together on the same branch. We then compared the frequency of that haplogroup in the European Caucasian samples as opposed to the Near Eastern samples (using the data of Richards et al. [60]).

A summary of the phylogenetic data is found in **Table 6**. Two of the sequences occurred more frequently in Near Eastern haplogroups, while the remaining two occurred in haplogroups with near equal frequency in European and Near Eastern populations. It is important to note that building a phylogenetic tree such as this does not definitively tell the researcher where the individual lived. It merely reflects the past evolutionary origin of the mtDNA molecule.

Overall, our results were consistent with a Near Eastern origin of the mtDNA sequences from the putative terrorists. In this case, though, several large caveats must be applied. Firstly, only HVI sequences were used, so the resulting phylogenetic trees must be interpreted with caution as to reliability. Secondly, the relative frequencies of the ascribed haplogroups in the European versus Near Eastern databases, while suggestive, were far from conclusive. Lastly, this type of analysis can only provide information regarding the biogeographic ancestry of the mtDNA molecule itself. Because of the possibility of admixture, both recent and historical, the ancestral origin of the mtDNA that an individual inherited from his maternal ancestor may or may not relate to his actual population affiliation.

With these points acknowledged, this exercise was useful for highlighting two important objectives for future forensic mtDNA research. First, more sequences are needed for the global mtDNA database. Many regions of the world are underrepresented in the current forensic databases. Since most academic laboratories analyzing mtDNA have focused on HVI only, forensic databases (with proper quality control procedures for sequencing) should continue to build global databases of control region sequences. Secondly, the utilization of phylogenetic trees can be a beneficial tool for the field of forensic science. Several groups have made recent use of phylogenetic tools for error detection in forensic databases [8], characterization of Caucasians in forensic databases [1], and identifying unknown plant material [9]. Most recently, a collaborative effort of the University of Innsbruck and AFDIL has been published detailing an optimized molecular assay for a series of single nucleotide polymorphism (SNP) sites from around the entire mtDNA genome that permit classification of European Caucasian mtDNA according to haplogroup [10]. These sites were selected from the phylogenetic literature, and involved evolutionarily stable polymorphisms that avoid the uncertainties described above for HVI sequences. The 20-plus years of population genetics research into mtDNA offers the field of forensics a wealth of information, especially as the use of coding region single nucleotide polymorphisms (SNPs) is introduced to mtDNA analysis [57].

Table 6. Phylogenetic associations and frequencies of the putative terrorist sequences from the United Airline Flight 93 (Shanksville, PA) from a published data set (Haplogroup associations were determined by adding the putative terrorist sequences to representative haplogroup sequences common in European and Near Eastern populations; the databases and frequencies were determined from Richards et al. [61].)

Sequence	Haplogroup	Euro. Caucasian (%) (*n* = 2804)	Near Eastern (%) (*n* = 1088)
Sequence 01	N1b	8 (0.28%)	19 (1.7%)
Sequence 02	U1a	12 (0.43%)	29 (2.7%)
Sequence 03	J1	75 (2.70%)	52 (4.8%)
Sequence 04	U4	84 (3.00%)	21 (1.9%)

VI. FUTURE ADVANCES

AFDIL is active in the pursuit of continued improvements to mtDNA forensic testing and the development of new mtDNA testing tools. Two predominant areas of focus are the characterization and utilization of variation over the entire mtDNA genome for increasing forensic discrimination and determination of biogeographic ancestry, and greatly expanded databases of mtDNA control region sequences and entire mtDNA genome SNPs.

For the last three years, the National Institute of Justice (NIJ) has funded research at AFDIL designed to increase forensic discrimination in the case of common mtDNA types. This work, recently submitted for publication, has identified 59 carefully selected SNP sites over the entire mtDNA genome that greatly enhance the power of discrimination of mtDNA testing when common types in the Caucasian population are encountered. The entire mtDNA genome was sequenced from multiple individuals matching 18 common HVI/HVII types that together comprise 18% of the Caucasian population. The battery of 59 SNP sites resolves 241 individuals with these 18 types into 106 different haplotypes, 56 of which have been seen only once. The sites are organized into eight separate panels to permit facile development of multiplex SNP assays. AFDIL and colleagues at the National Institute of Standards and Technology (NIST) have developed and optimized an 11-plex primer extension assay that

powerfully discriminates among individuals matching the most common mtDNA HVI/HVII type in the Caucasian population (present at ~7%). The assay is demonstrated to be highly effective in typing forensic samples typical of mtDNA casework. Development continues for additional multiplexes targeting other common Caucasian types, and work in the near future will target other population groups for entire mtDNA genome sequencing and development of SNP assays for increased forensic discrimination.

There is a vast literature on human mtDNA variation that comes from evolutionary anthropological studies, recently including a substantial number of sequences of the entire mtDNA genome [22,33,42,49,68]. If sufficient sequence or SNP information is known, mtDNA types can be unambiguously ascribed to an mtDNA haplogroup. As mentioned above, this has been done for Caucasian haplogroups in the forensic context [10]. It is a straightforward extension to broaden assay development so that unknown mtDNA samples could be traced to their ancestral biogeographic origin on a global basis. We believe that such information, if properly interpreted, has a useful role to play in certain types of investigations. It should be noted, however, that the biogeographical determination of a mtDNA sample is a reflection only of the maternal lineage and does not indicate phenotypical racial characteristics such as hair or skin color. One obvious application would be assisting in assessing the potential of skeletal remains recovered from Southeast Asia to represent, say, Caucasian U.S. service personnel as opposed to individuals of Southeast Asian descent. While there is potential for introgression (Asian mtDNAs within "Caucasian" individuals, and vice versa), it is nonetheless clear that evidence of mtDNA ancestry could be helpful as an investigation unfolds. AFDIL and others are currently developing SNP markers for major global haplogroups.

MtDNA testing requires reference to an appropriate population, and the strength of mtDNA evidence is in many instances limited by database size. In populations encountered to date, the majority of mtDNA types are rare (although some are common) and only with large databases can we reveal the full evidentiary significance of these types [36]. Therefore, the most important means for increasing the global utility of mtDNA testing are quite straightforward: increase the size of databases and greatly expand the global population representation. For forensic purposes, this requires standardization of the DNA regions targeted, and levels of quality assurance that consider not only DNA sequencing, but also the integrity of data manipulation and archiving (for issues of quality in mtDNA databases *see* Forster [23]). As mentioned above, the AFDIL mtDNA Family Reference Section has generated sequences for 4021 unrelated family references, and for some time has been targeting the entire control region for its database. The AFDIL Research Section has also completed large-scale sequencing projects, including some 285 complete mtDNA genome sequences (the equivalent of ~4300 control region sequences), as well as population databases for Japanese [41], Austrian [56], Nairobi (in preparation), and Uzbekistan (in preparation) populations, over 1300 control region sequences in total. As a result of these efforts, AFDIL has a robustly established robotic system for high-throughput mtDNA population databasing, including DNA extraction, cycle sequencing, and cycle sequencing cleanup. Of possibly greater importance is a well-developed bioinformatics system whereby consensus sequences are automatically exported to a database, and, after quality control confirmation, locked to avoid subsequent corruption.

A principal future goal of AFDIL is to aggressively contribute to the growth of high-quality global mtDNA sequence databases. This needs to be activity pursued for and by the entire forensic community. We are presently in partnership with the European DNA Profile Group mtDNA Population Database (EMPOP: http://www.empop.org), administered by the Institute of Legal Medicine, University of Innsbruck, for curation and distribution of population databases generated at AFDIL. A goal of EMPOP is to link raw data files in perpetuity with a searchable database of consensus sequences, thereby establishing a permanent documentation of database quality. The database information will be maintained openly on the Internet for use of the entire mtDNA community. In the AFDIL/EMPOP partnership, a highly beneficial feature is that databases are independently generated from raw data at both laboratories, compared, and confirmed not only at the level of individual sequences, but at the level of the archived database itself. Quality is controlled at the level of basic sequencing accuracy and for bioinformatics manipulations as well, the latter being the more difficult and pervasive source of error in published databases. We would like to suggest that forensic laboratories worldwide should cooperate to make DNA samples (such as those used, for example, to establish STR allele frequency databases) available for the establishment of mtDNA population databases of uniform quality and unrestricted access.

CONCLUSIONS

The technologies available to molecular forensics are growing by leaps and bounds. AFDIL remains on the leading edge of these changes. The end of 2003 will see the implementation of ABI PRISM 3100® Genetic Analyzer for the processing of degraded skeletal remains and the

Tecan Genesis RSP 200 for the sequencing of family references. We plan to augment our scientific staff in accordance with our increased capacity for processing. We have gone to great lengths to provide quality DNA processing and analysis, and intend to maintain this level of standard as our abilities continue to expand.

The need for large global databases of mtDNA genotypes will become more apparent as mtDNA is increasingly used in international cases of missing persons identification, mass disasters, and criminal investigations. AFDIL hopes to be able to provide some of this data as the dedicated family reference section completes the backlog of reference bloodstains and must content itself with approximately 100 references a month in a high-throughput system or put that system to good use. Because most mtDNA types are unique even in the largest population databases, increasing the size of databases is important for realizing the full strength of mtDNA testing as a general forensic tool. Of even greater importance is to expand the global populations that are represented by appropriate forensic databases.

As we increase our abilities to process greater numbers of samples and to do additional types of testing for other projects, we will be careful not to lose sight of our goals. Thousands of families have lost siblings, parents, cousins, and children to the uncertainties of war and have yet to see their loved ones returned. At AFDIL it is our continuing focus to work with our partners at CILHI and AFMES and to use the tools of science and technology to bring these missing heroes home.

ACKNOWLEDGMENTS

The authors would like to thank Brion Smith, James Canik, Demris Lee, and the entire staff of AFDIL for their hard work and dedication to the mission. The authors would also like to thank Timothy McMahon, Heather Thew, and Jodi Irwin; Victor Weedn, Mitch Holland, Rhonda Roby, Ed Huffine, and John Ryan, all of whom had an indelible impact on mtDNA testing at AFDIL, and we are likewise indebted to many other past employees and staff of AFDIL, without whom our current level of skill would be impossible; Walther Parson (University of Innsbruck) and John Butler (NIST) and others of their laboratories for ongoing productive research collaborations; the National Institute of Justice, Office of Justice Programs, Department of Justice for research funding (2000-IJ-CX-K010); the American Registry of Pathology for administrative support; The Armed Forces Institute of Pathology; and Future Technologies, Inc., for the development and continuing improvements to LISA. Lastly we would like to thank Tom Holland, Mark Leney, and the staff of the United States Army Central Identification Laboratory, Hawaii, who often risk life and limb to acquire remains from our missing service members.

REFERENCES

1. Allard MW, Miller K, Wilson M, Monson K, Budowle B: Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA dataset for 1771 human control region sequences; *J Forensic Sci* 47:1215; 2002.
2. Allen M, Saldeen T, Gyllensten U: PCR-based DNA typing of saliva on stamps and envelopes; *Biotechniques* 17:546; 1994.
3. Allen M, Engstrom AS, Meyers S, Handt O, Saldeen T, von Haeseler A, Pääbo S, Gyllensten U: Mitochondrial DNA sequencing on shed hairs and saliva on robbery caps: sensitivity and matching probabilities; *J Forensic Sci* 43:453; 1998.
4. Alves-Silva J, da Silva Santos M, Guimarães PEM, Ferreira ACS, Bandelt H-J, Pena SDJ, Prado VF: The ancestry of Brazilian mtDNA lineages; *Am J Hum Genet* 67:444; 2000.
5. Anderson S, Bankier AT, Barrell GB, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG: Sequence and organization of the human mitochondrial genome; *Nature* 290:457; 1981.
6. Anderson TD, Ross JP, Roby RK, Lee DA, Holland MM: A validation study for the extraction and analysis of DNA from human nail material and its application to forensic casework; *J Forensic Sci* 44:1053; 1999.
7. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N: Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA; *Nature Genet* 23:147; 1999.
8. Bandelt HJ, Lahermo P, Richards M, Macaulay V: Detecting errors in mtDNA data by phylogenetic analysis; *Int J Legal Med* 115:64; 2001.
9. Bever R, Golenberg E, Barnes L, Brinkac L, Jones E, Yoshida K: Molecular analysis of botanical trace evidence: development of techniques; *Proceedings — 2000 Annual Meeting of the American Academy of Forensic Sciences*; Reno, Nevada; pp 23–24; 2000.
10. Brandstätter A, Parson W, Parsons TJ: Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups; *Int J Legal Med*; 117:291; 2003.
11. Brown MD, Hosseini SH, Torroni A, Bandelt H-J, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC: mtDNA haplogroup X: An ancient link between Europe/ western Asia and North America? *Am J Hum Genet* 63:1852; 1998.
12. Budowle B, Allard MW, Fisher CL, Isenberg AR, Monson KL, Stewart JEB, Wilson MR, Miller KWP: HVI and HVII mitochondrial DNA data in Apaches and Navajos; *Int J Legal Med* 116:212; 2002.
13. Budowle B, Wilson M, DiZinno J, Stauffer C, Fasano M, Holland M, Monson K: Mitochondrial DNA regions HVI and HVII population data; *Forensic Sci Int* 103:23; 1999.
14. Carracedo A, Bär W, Lincoln P, Mayr W, Morling N, Olais B, Schneider P, Budowle B, Brinkmann B, Gill P, Holland M, Tully G, Wilson M: DNA Commission of the Interna-

tional Society for Forensic Genetics: Guidelines for mitochondrial DNA typing; *Forensic Sci Int* 110:79; 2000.
15. Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, Leon P, Ospina J, Bedoya G, Freimer N, Ruiz-Linares A: Genetic demography of Antioquia (Columbia) and the central valley of Costa Rica; *Hum Genet* 112:534; 2003.
16. Carvajal-Carmona LG, Soto ID, Peneda N, Ortíz-Barrientos D, Duque C, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G, Ruiz-Linares A: Strong Amerind/White sex bias and a possible sephardic contribution among the founders of a population in northwest Columbia; *Am J Hum Genet* 67:1287; 2000.
17. CILHI Case Report: Randolph A. Perry, Irwin S. Lerner: United States Army, Central Identification Laboratory, Hawaii; 2003.
18. Derbeneva OA, Starikovskaya EB, Wallace DC, Sukernik RI: Traces of early Eurasians in the Mansi of Northwest Siberia revealed by mitochondrial DNA analysis; *Am J Hum Genet* 70:1009; 2002.
19. Derbeneva OA, Sukernik RI, Volodko NV, Hosseini SH, Lott MT, Wallace DC: Analysis of mitochondrial DNA diversity in the Aleuts of the Commander Islands and its implications for the genetic history of Beringia; *Am J Hum Genet* 71:415; 2002.
20. Dimo-Simonin N, Grange F, Taroni F, Brandt-Casadevall C, Mangin P: Forensic evaluation of mtDNA in a population of from southwest Switzerland; *Int J Legal Med* 113:89; 2000.
21. Di Rienzo A, Wilson AC: Branching pattern in the evolutionary tree for human mitochondrial DNA; *Proc Natl Acad Sci USA* 88:1597; 1991.
22. Finnilä A, Lehtonen MS, Majamaa K: Phylogenetic network for European mtDNA; *Am J Hum Genet* 68:1475; 2001.
23. Forster P: To err is human; *Ann Hum Genet* 67:2; 2003.
24. Francalacci P, Bertranpetit J, Calafell F, Underhill PA: Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe; *Am J Phys Anthrop* 58:127; 1996.
25. Gabriel MN, Huffine EF, Ryan JH, Holland MM, Parsons TJ: Improved mtDNA sequence analysis of forensic remains using a "mini-primer" set amplification strategy; *J Forensic Sci* 46:247; 2001.
26. Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L: Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample; *Mol Biol Evol* 12:334; 1995.
27. Green LD, Derr JN, Knight A: mtDNA affinities of the peoples of North-Central Mexico; *Am J Hum Genet* 66:989; 2000.
28. Gresham D, Morar B, Underhill PA, Passarino G, Lin AA, Wise C, Angelicheva D, Calafell F, Oefner P, Shen P, Tournev I, de Pablo R, Kuĉinskas V, Perez-Lezaun A, Marushiakova E, Popov V, Kalaydjieva L: Origins and divergence of the Roma (Gypsies); *Am J Hum Genet* 69:1314; 2001.
29. Handt O, Richards M, Trommsdorff M, Kilger C, Simanainen J, Georgiev O, Bauer K, Stone A, Hedges R, Schaffner W, Utermann G, Sykes B, Pääbo S: Molecular genetic analysis of the Tyrolean Ice Man; *Science* 264:1775; 1994.
30. Helgason A, Hickey E, Goodacre S, Bosnes V, Stefansson K, Ward R, Sykes B: mtDNA and the islands of the North Atlantic: Estimating the proportions of Norse and Gaelic ancestry; *Am J Hum Genet* 68:723; 2001.
31. Helgason A, Hrankelsson B, Gulcher JR, Ward R, Stefánsson K: A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: Evidence for a faster evolutionary rate of mtDNA lineages than Y-chromosomes; *Am J Hum Genet* 76:1370; 2003.
32. Helgason A, Sigurðardóttir S, Gulcher JR, Ward R, Stefánsson K: mtDNA and the origin of the Icelanders: deciphering signals of recent population history; *Am J Hum Genet* 66:999; 2000.
33. Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N: Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups; *Am J Hum Genet* 70:1152; 2002.
34. Hofmann S, Jaksch M, Bezold R, Mertens S, Aholt S, Paprotta A, Gerbitz KD: Population genetics and disease susceptibility: Characterization of central European haplogroups by mtDNA gene mutations, correlations with D-loop variants and association with disease; *Hum Mol Genet* 6:1835; 1997.
35. Holland MM, Fisher DL, Mitchell LG, Rodriguez WC, Canik JJ, Merril CR, Weedn VW: Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War; *J Forensic Sci* 38:542; 1995.
36. Holland MM, Parsons TJ: Mitochondrial DNA sequence analysis — Validation and use for forensic casework; *Forensic Sci Rev* 11:22; 1999.
37. Holland MM, Roy R, Fraser MD, Liu RH: Application of serological and DNA methods for identification of urine specimen donors; *Forensic Sci Rev* 5:1; 1993.
38. Hopwood AJ, Mannucci A, Sullivan KM: DNA typing from human fæces; *Int J Legal Med* 108:237; 1996.
39. Houck MM, Budowle B: Correlation of microscopic and mitochondrial DNA hair comparisons; *J Forensic Sci* 47:964; 2002.
40. Howe JR, Klimstra DS, Cordon-Cardo C: DNA extraction from paraffin-embedded tissues using a salting-out procedure: A reliable method for PCR amplification of archival material; *Histol Histopath* 12:595; 97.
41. Imaizumi K, Parsons TJ, Yoshino M, Holland MM: A new database of mitochondrial DNA hypervariable regions I and II sequences from 162 Japanese individuals; *Int J Legal Med* 116:68; 2002.
42. Ingmann M, Kaessman H, Pääbo S, Gyllensten U: Mitochondrial genome variation and the origin of modern humans; *Nature* 408:708; 2000.
43. Jehaes E, Decorte R, Peneau A, Petrie JH, Boiry PA, Gilissen A, Moisan JP, Van den Berghe H, Pascal O, Cassiman JJ: Mitochondrial DNA analysis on remains of a putative son of Louis XVI, King of France, and Marie-Antoinette; *Eur J Hum Genet* 6:383; 1998.
44. Junge A, Steevens M, Madea B: Successful DNA typing of a urine sample in a doping control case using human mitochondrial DNA analysis; *J Forensic Sci* 47:1002; 2002.

45. Lee SD, Lee YS, Lee JB: Polymorphisms in the mitochondrial cytochrome B gene in Koreans: An additional marker for individual identification; *Int J Legal Med* 116:74; 2002.
46. Lindholm E, Cavelier L, Howell WM, Eriksson I, Jalonen P, Adolfsson R, Blackwood DHR, Muir WJ, Brookes AJ, Gyllensten U, Jazin EE: Mitochondrial sequence variants in patients with schizophrenia; *Eur J Hum Genet* 5:406; 1997.
47. Lorenz JG, Smith DG: Distribution of four founding mtDNA haplogroups among native North Americans; *Am J Phys Anthrop* 101:307; 1996.
48. Lutz S, Weisser H-J, Heizmann J, Pollak S; Location and frequency of polymorphic positions in the mtDNA control region of individuals from Germany; *Int J Legal Med* 111:67; 1998.
49. Maca-Meyer N, González AM, Larruga JM, Flores C, Cabrera VM: Major genomic mitochondrial lineages delineate early human expansions; *BMC Genetics* 2:13; 2001.
50. Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonné-Tamir B, Sykes B, Torroni A: The emerging tree of West Eurasian mtDNAs: A synthesis of control-region sequences and RFLPs; *Am J Hum Genet* 64:232; 1999.
51. Monson KL, Miller KWP, Wilson MR, DiZinno JA, Budowle B: The mtDNA population database: An integrated software and database resource for forensic comparison; *Forensic Sciences Communication* 4(2); 2002 (http://www.fbi.gov/hq/lab/fsc/backissu/april2002/index.htm).
52. Niemi A-K, Hervonene A, Hurme M, Karhunen PJ, Jylhä M, Majamaa K: Mitochondrial DNA polymorphisms associated with longevity in a Finnish population; *Hum Genet* 112:29; 2003.
53. Nixon KC: WinClada, version 1.00.08; Ithaca, NY; 2002.
54. Ozawa T, Tanaka M, Ino H, Ohno K, Sano T, Wada Y, Yoneda M, Miyatake T, Tanaka T, Itoyama S, Ikebe S, Hattori N, Mizuno Y: Distinct clustering of point mutations in mitochondrial DNA among patients with mitochondrial encephalomyopathies and with Parkinson's disease; *Biochem Biophys Res Comm* 176:938; 1991.
55. Pääbo S: Ancient DNA — Extraction, characterization, molecular cloning, and enzymatic amplification; *Proc Natl Acad Sci USA* 86:1939; 1989.
56. Parson W, Parsons TJ, Scheithauer R, Holland MM: Population data for 101 Austrian Caucasian mitochondrial DNA d-loop sequences: Application of mtDNA sequence analysis to a forensic case; *Int J Legal Med* 111:124; 1998.
57. Parsons TJ, Coble MD: Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome; *Croat Med J* 42:304; 2001.
58. Pfeiffer H, Forster P, Ortmann C, Brinkmann B: The results of an mtDNA study of 1200 inhabitants of a German village in comparison to other Caucasian databases and its relevance for forensic casework; *Int J Legal Med* 114:169; 2001.
59. Rando JC, Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM, Bandelt H-J: Mitochondrial DNA analysis of northwest African populations reveals genetic exchanges with European, Near Eastern, and Sub-Saharan populations; *Ann Hum Genet* 62:531; 1998.
60. Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Golge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Demaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Di Rienzo A, Novelletto A, Oppenheim A, Norby S, Al-Zaheri N, Santachiara-Benerecetti S, Scozari R, Torroni A, Bandelt HJ: Tracing European founder lineages in the Near Eastern mtDNA pool; *Am J Hum Genet* 67:1251; 2000.
61. Richards MB, Macaulay VA, Bandelt H-J, Sykes BC: Phyologeography of mitochondrial DNA in Western Europe; *Ann Hum Genet* 62:241; 1998.
62. Richards M, Rengo C, Cruciani F, Gratrix F, Wilson JF, Scozzari R, Macaulay V, Torroni A: Extensive female-mediated gene flow from Sub-Saharan Africa into Near Eastern Arab populations; *Am J Hum Genet* 72:1058; 2003.
63. Rickards O, Martinex-Labarga C, Lum JK, DeStefano GF, Cann RL: mtDNA history of the Cayapa Amerinds of Ecuador: Detection of additional founding lineages for the North American populations; *Am J Hum Genet* 65:519; 1999.
64. Smith BC, Fisher DL, Weedn VW, Warnock GR, Holland MM: A systematic approach to the sampling of dental DNA; *J Forensic Sci* 38:1194: 1993.
65. Soodyall H, Vigilant L, Hill AV, Stoneking M, Jenkins T: mtDNA control-region sequence variation suggests multiple independent origins of an "Asian-specific" 9-bp deletion in sub-Saharan Africans; *Am J Hum Genet* 58:595; 1996.
66. Swofford DL: *PAUP® Phylogenetic Analysis Using Parsimony (and Other Methods)*, Version 4; Sinauer Associates: Sunderland, MA; 2002.
67. Swofford DL, Olsen GJ, Waddell PJ, Hillis DM: Phylogenetic inference; In Hillis DM, Mortiz C, Mable BK (Eds): *Molecular Systematics*; Sinaur Associates: Sunderland, MA; p 407; 1996.
68. Torroni A, Campos Y, Rengo C, Sellitto D, Achilli A, Magri C, Semino O, García A, Jara P, Arenas J: Mitochondrial DNA haplogroups do not play a role in the variable phenotypic presentation of the A3243G mutation; *Am J Hum Genet* 72:1005; 2003.
69. Vandenberg N, van Oorschot RAH: Extraction of human nuclear DNA from feces samples using the QIAmp DNA stool mini kit; *J Forensic Sci* 47:993; 2002.
70. Walsh PS, Metzger DA, Higuchi D: Chelex®100 as a medium for simple extraction of DNA for PCR-based typing from forensic material; *Biotechniques* 10:506; 1991.
71. Wilson MR, Holland MM, Stoneking M, DiZinno JA, Budowle B: Guidelines for the use of mitochondrial DNA sequencing in forensic science; *Crime Laboratory Digest* 20:68; 1993.
72. Wilson MR, Polansky D, Butler J, DiZinno JA, Replogle J, Budowle B: Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts; *Biotechniques* 18:662; 1995.

ABOUT THE AUTHORS

S.M. Edson; J.P. Ross; M.D. Coble; T.J. Parsons; S.M. Barritt

Suni M. Edson graduated with a B.S. degree in biology from Texas Tech University (Lubbock, TX) and an M.S. degree in environmental and forest biology from the State University of New York College of Environmental Science and Forestry (Syracuse, NY). She is currently a supervisory DNA analyst at the Armed Forces DNA Identification Laboratory (AFDIL) in Rockville, MD.

While pursuing her undergraduate degree, Ms. Edson was employed as a research technician in a campus laboratory that focused on population level characterizations of ribosomal DNA and "jumping genes" in numerous species including rodents, horses, and cotton. As a graduate student, Ms. Edson turned her attention to the study of the introgressive hybridization of duck species in the Adirondack Park, NY. Prior to coming to AFDIL in 1999, Ms. Edson was employed as a research technician at the Lombardi Cancer Research Center, Georgetown University (Washington, DC). While there, she worked with a team of scientists to characterize novel HLA alleles in humans.

In her current role as a supervisor, Ms. Edson oversees a team of DNA analysts and technicians working with the Central Identification Laboratory, Hawaii (CILHI) casework. During her tenure at AFDIL, she has also worked as a technician in the family reference group, an analyst in the mtDNA section, and participated in the DNA evidence collection team at Dover Air Force Base in response to the events of September 11, 2001.

James P. Ross graduated with a B.S. degree in biology from Gettysburg College (Gettysburg, PA) in 1993. Mr. Ross has been the chief information officer (CIO) of the Department of Defense DNA Registry since 2000.

Before switching to information management and bioinformatics, he was a DNA analyst for AFDIL Nuclear DNA Section for seven years. In his current role as CIO, Mr. Ross's responsibilities include laboratory automation, laboratory information system (LIMS) development, and bioinformatics solutions for both AFDIL and the Armed Forces DNA Specimen Repository. Throughout his career at the DNA Registry, he has been actively involved in data management, specimen collection, and the identification of remains from mass disasters, including Branch Davidians (Waco, TX), TWA Flight 800, Egypt Air Flight 990, Alaska Air Flight 77, the Tri-State Crematory Incident (Noble, GA), and the terrorists attacks of September 11, 2001 at the Pentagon and Shanksville, PA.

Michael D. Coble received a B.S. degree in biology from Appalachian State University (Boone, NC) and a master of forensic science degree from George Washington University (Washington, DC). He is currently a research technologist in AFDIL's research section.

Mr. Coble is currently finishing his doctoral degree in genetics from George Washington University. His dissertation and research focus involves the identification of informative single nucletotide polymorphisms (SNPs) within the entire mitochondrial genome, and the characterization of mutation rate variation in mtDNA. The intent of his work is to increase the forensic discrimination of common mtDNA haplotypes through the use of SNPs.

Mr. Coble is a member of the American Society of Human Genetics and the Mid-Atlantic Association of Forensic Scientists.

Thomas J. Parsons received his undergraduate degree in physics from the University of Chicago (Chicago, IL) and a Ph.D. degree in biochemistry from the University of Washington (Seattle, WA) in 1989. Dr. Parsons is currently the chief scientist of AFDIL and an adjunct faculty member in both the Departments of Genetics and Forensic Sciences at George Washington University (Washington, DC).

As a postdoctoral fellow at the Smithsonian Institution, he focused on molecular evolution and phylogenetics, as well as mtDNA biogeography and avian speciation. He continued to research molecular systematics and population genetics during a research faculty appointment at the University of Nebraska. Dr. Parsons joined AFDIL in 1994. One of Parsons' primary current roles is to direct the AFDIL Research Section. Particular areas of emphasis include: (a) Development of high-throughput robotic systems for mtDNA sequencing; (b) Mitochondrial DNA genomics for increased forensic discrimination; (c) Statistical interpretation of forensic data; (d) Mitochondrial DNA mutation rate and evolution; (e) Improved techniques for recovery of DNA from highly degraded sources; and (f) Bioinformatics. His work that has received particular attention relates to an unexpectedly high mutation rate in human mtDNA and the identification of Tsar Nicholas II and his family.

Dr. Parsons was a finalist for the 2001 Berry Prize in Federal Medicine. He serves on the Scientific Advisory Board of the International Commission on Missing Persons (ICMP), as well as an expert advisory panel (KADAP) for data interpretation issues for the World Trade Center DNA identification efforts.

Suzanne M. Barritt received her B.S. degree in biology from Pennsylvania State University (University Park, PA) and her M.S. degree in administration of justice with a concentration in forensic science from Virginia Commonwealth University (Richmond, VA). Ms. Barritt is currently the technical leader and section chief of the mtDNA section at AFDIL.

As the technical leader, Ms. Barritt is responsible for overseeing the processing of thousands of specimens for mtDNA analysis and the largest forensic mtDNA testing staff in the world. Ms. Barritt began her employment at AFDIL as a DNA analyst in 1994. Of particular interest is her participation with the identification of Anna Anderson, the woman claiming to be the Grand Duchess Anastasia, daughter of Tsar Nicholas II. Details of her involvement in this case are described in Robert K. Massie's *The Romanovs: The Final Chapter*. While pursuing her education, Ms. Barritt began her career in forensic science at the Virginia Division of Forensic Science, where she played a large role in the Virginia convicted offender DNA testing program. Her work contributed to the first database hit in the state. She has also participated in the collection and identification of remains from the attacks on the Pentagon and UA Flight 93 on September, 2001. Recently, Ms. Barritt served on the Scientific Advisory Board of the ICMP.

In 2002, Ms. Barritt was awarded a Special Recognition, DNA Laboratory Award, from the Armed Forces Institute of Pathology and the American Registry of Pathology.