IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF TEXAS
SAN ANTONIO DIVISION

JOHN A. PATTERSON, et al.,	)	
	)	
Plaintiffs,	)	
	)	
v.	)	No. 5:17-CV-00467
	)	
DEFENSE POW/MIA ACCOUNTING AGENCY, et al.,	)	
	)	
Defendants.	)	

DECLARATION OF TIMOTHY MCMAHON

I, Timothy P. McMahon, pursuant to 28 U.S.C. § 1746, declare as follows:

1. I am currently the Director of Department of Defense (DoD) DNA Operations for the Armed Forces Medical Examiner System (AFMES). The DoD DNA Operations division is comprised of two sections, the Armed Forces Repository of Specimen Samples for the Identification of Remains (AFRSSIR) and the Armed Forces DNA Identification Laboratory (AFDIL). From October of 2016 until selection for my current position in April of 2017, I served as the Director of AFDIL. From 2012 to 2016, I served as Director of Forensic Services within AFDIL.

2. I am a Forensic Specialist with over 16 years of specialized experience in overseeing receipt, forensic analysis, forensic research and return of physical evidence associated with criminal investigations and I have over 26 years of biology, biochemistry and molecular biology experience.

3. I received a PhD in Biomedical Sciences from the School of Public Health at the

University of Albany, New York in August of 2001. My graduate studies and post-doctoral research were performed in the division of Infectious Disease and Immunology at the New York State Department of Health. From 2002 to 2007, I worked for the American Registry of Pathology as a contractor supporting AFDIL and AFMES. From 2007 to 2012, I worked for Applied Biosystems where I was responsible for developing an organization to help create new DNA forensic laboratories and aided established government crime laboratories implement new automated and manual forensic technologies.

4. The statements contained in this declaration are based on my personal knowledge and AFMES records and information made available to me in my official capacity.

5. I am responsible for managing 150 contract scientist and support staff in meeting our mission requirements of performing DNA sequencing and testing on human remains for the AFMES and the Defense POW/MIA Accounting Agency (DPAA), criminal paternity testing for military investigative organizations, sample switches for military treatment facilities, and other Government, State or Local agencies as dictated. This includes serving as the AFMES subject matter expert on DNA, molecular biology, virology, biochemistry, genetics, DNA forensic applications, emerging technologies & research initiatives, and laboratory design & management. I am also responsible for guiding DoD DNA Operations in its development of new testing technologies, for developing and implementing growth plans to meet AFMES and DPAA needs, and for serving as a DoD subject matter expert in DNA human remains testing.

AFDIL's Mission and Organization

6. The Armed Forces DNA Identification Laboratory (AFDIL) was established in 1991 as the only DoD forensic DNA testing laboratory for the identification of human remains.

7. AFDIL's present day accounting and past accounting sections provide the DoD

and other federal and international agencies with human identification DNA testing support in the areas of personnel accounting, national security, law enforcement, humanitarian missions, and defense. The primary missions of AFDIL are to provide: (1) forensic DNA testing of remains and other biological evidence in support of identification efforts through its past accounting section, which supports the DPAA, as well as its present day accounting section, which supports the Office of the Armed Forces Medical Examiner; (2) to create a conflict-specific mitochondrial DNA (mtDNA), autosomal short tandem repeat (auSTR), and Y chromosome short tandem repeat (Y-STR) family reference database for use in the past accounting identification process; (3) to modify or create new methods to increase the present and past accounting sample success rates; and (4) to provide worldwide consultation, research, and education services in the field of forensic DNA to the DoD and other agencies.

8. In 1998, AFDIL received its American Society of Crime Laboratory Directors - Laboratory Accreditation Board (ASCLD) and Federal Bureau of Investigation-Quality Assurance Standards (FBI-QAS) accreditation in Biology for nuclear and mitochondrial DNA testing and has been accredited continuously since 1998. AFDIL was one of the first laboratories accredited by ASCLD for mtDNA testing. In 2014, AFDIL successfully underwent reaccreditation from the ASCLD-Legacy program to the American Society of Crime Laboratory Directors - Laboratory Accreditation Board (ASCLD-LAB) International Program, which found AFDIL to be in compliance with the International Organization of Standardization (ISO) 17025, ASCLD-LAB Forensic Requirements, and the Federal Bureau of Investigations Quality Assurance Standards (FBI-QAS) for Accreditation. See Exhibit 1. AFDIL has maintained its accreditation through each interim review process. See Exhibits 2, 3. ASCLD-LAB has now been acquired by the ANSI-ASQ National Accreditation Board (ANAB), and AFDIL is in the

process of renewing its accreditation through ANAB.

9. AFDIL is comprised of six sections:

a. Current Day Operations: Works directly with the Office of the Armed Forces Medical Examiner system to assist with human remains DNA testing for the identification of service members killed in current theaters of operation or training mishaps, and to assist military criminal investigative organizations with criminal paternity or kinship analysis cases. The section is comprised of a Technical Leader and Assistant Technical Leader who oversee the day to day operations of 10 qualified DNA analysts.

b. Past Accounting Operations: Works directly with the DPAA and the AFMES to perform mtDNA, auSTR, and Y-STR testing on specimen samples submitted from WWII, Korea, Vietnam, and the Cold War. The section is comprised of one Technical Leader and three Assistant Technical Leaders, who manage eight teams comprised of 60 qualified DNA analysts and technicians. This section utilizes a team approach for efficiency and allows the greatest flexibility to meet changing DPAA requirements.

c. Family Reference Specimen-Laboratory Automation: Works directly with the service causality offices and the past accounting section to process all of the family references that are submitted for inclusion in AFDIL Family Reference Database. This database is an internal database that is protected by HIPAA and only AFDIL scientist have access to. The family reference database is used by the past accounting section to compare the results generated from unknown specimens to the FRS specific conflict or person. This section is comprised of a Technical Leader and Assistant Technical Leader,

who manage 11 qualified DNA analysts and technicians.

d. Emerging Technology: Responsible for developing the new methods that are currently not commercially available but are needed to handle the highly degraded samples processed by AFDIL. These include the demineralization buffer that is used by most laboratories now and most recently the Next Generation mtDNA Capture assay. This section is comprised of a single team headed by a PhD and 6 research scientists.

e. Validation and Quality Control: AFDIL employs both commercially available reagents as well as reagents that are generated in house. All of these reagents must meet defined validated procedures and accreditation requirements. To accomplish this, AFDIL has eight DNA scientist who are responsible for performing all of the validation, performance checks, and quality control of the instruments and reagents used by the casework sections.

f. Quality Management and Training and Education: This is a six person team whose sole responsibility is the training of all scientists to meet accreditation requirements, maintenance of AFDIL's accreditation, and the management of all proficiency tests assigned to qualified scientists.

AFDIL's DNA Testing Capabilities

10. AFDIL uses auSTR and Y-STR tests to analyze nuclear DNA, and mtDNA sequencing (Sanger and Next Generation) to analyze mitochondrial DNA. For the past accounting program, because of the age and degradation of the DNA due to environmental conditions, mitochondrial DNA testing is the most sensitive and is usually the first type of DNA testing used. If the appropriate reference materials are available, autosomal DNA and Y-DNA will be tested as well.

11. Different DNA testing methods have different strengths and weaknesses when testing highly degraded samples and their use in the human remains identification process. The multidisciplinary testing approach is most efficient and allows for the inclusion and exclusion of missing service members that might not be possible on the basis of one method alone.

12. For example, mtDNA testing is highly effective in compromised skeletal cases because of its durability and high-copy number per sample compared to the single copy of nuclear DNA. Additionally, mtDNA is very effective for use in closed-population groups and in situations where auSTR or Y-STR reference samples may be difficult to obtain.

13. AFDIL has a demonstrated record of developing methods to meet the needs of the AFMES and DPAA and is considered a world leader in human remains DNA testing. AFDIL monitors success rates for testing and currently has a greater than 90% success rate for mtDNA Sanger sequencing results of non-chemically treated specimens, a greater than 50% success rate for auSTR and YSTR testing, and about a 45% success rate for mtDNA Next Generation Sequencing results. To consistently achieve these success rates, AFDIL extensively tests samples with in-depth troubleshooting to make them work and to develop innovative solutions.

14. To compare laboratories on the basis of metrics such as success rate, one must ensure that they involve comparable types of samples (e.g., highly degraded samples versus contemporary criminal casework).

15. AFDIL developed and implemented in 1998 mtDNA mini-primer-sets and currently is the only laboratory that uses these. The advantage is that the amplicon size is approximately half the size of mtDNA primer-sets that all other laboratories use. Having the ability to use primer-sets and mini-primer sets increases the chance of success for specimens associated with DPAA. Currently, about 65% of all non-chemically treated DPAA specimens

require mini-primer-sets to obtain results.

16. In 2006, AFDIL developed advances, including the demineralization buffer, which have reduced the needed sample size from 2.5 g of bone to 0.2 g of bone and allowed for the complete digestion of the bone, which made it possible to recover what little nuclear DNA was present and to perform auSTR and Y-STR testing. The previous extraction method used by all forensic laboratories failed to release enough usable nuclear DNA for testing of highly degraded samples. These advances also allowed for submission of smaller bones that could not be visually distinguished as human. AFDIL developed and implemented a 12s rRNA test to determine if a bone extract was human or non-human, which allows AFDIL to stop testing non-human samples and focus on human samples.

17. In 2015, AFDIL developed and forensically validated the Next Generation Sequencing (NGS) mtDNA Capture Assay and custom analysis software for analyzing NGS-derived mtDNA sequencing data. See Exhibit 3, C. Marshall, et al., Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples, 31 Forensic Science Int'l: Genetics 198-206 (2017); Exhibit 4, Kimberly Sturk-Andreaggi, et al., AQME: A forensic mitochondrial DNA analysis tool for next-generation sequencing data, 31 Forensic Science Int'l: Genetics 189-197 (2017). This grew out of AFDIL's longstanding effort to meet the need to identify approximately 850 sets of remains that at the end of the Korean War that were treated with chemical agents (formaldehyde) to preserve the remains (known as the "Korean Punchbowl" remains). Conventional mtDNA Sanger sequencing methods worked less than 5% of the time for these specimens. AFDIL began successfully using this new sequencing method on Korean Punchbowl samples in March of 2016.

18. Currently, AFDIL is the only forensic laboratory with a forensically validated

NGS mtDNA testing method for highly degraded samples. This method was externally reviewed by an audit team in 2016 with no findings of any deficiencies. For disinterments associated with World War II sites like Cabanatuan, where the remains were chemically treated before final burial, and for highly degraded samples from Vietnam where traditional methods do not work, the NGS mtDNA capture assay is frequently the only method that will work.

19. AFDIL, as part of a multi-laboratory team, performed the National DNA Index System (NDIS) testing on commercially available NGS forensic panels (auSTR, Y-STR, Ancestry SNPs, Phenotypic SNPs). The results from this testing were summarized, written up and submitted to the NDIS committee in late 2017, and are currently undergoing the review process. The team observed that the commercially available kits worked well for modern high copy criminal casework samples, but are not optimal for low copy or degraded samples. AFDIL does not find the commercially available kits useful for the past accounting mission.

Family Reference Sample Collection

20. AFDIL maintains a collection of family reference samples to support comparison of DNA testing results from unidentified remains. Collection began in 1991, focused on family members associated with Vietnam losses, and in 1995 expanded to include family members associated with Korean War losses. After Congress provided additional funding to the service causality offices in 2010, DoD has engaged in a substantial push to gather all references for losses associated with World War II, Korean War, Vietnam War, and Cold War. Due to this push for collection of family references, AFDIL currently has 91% coverage for Korean War missing service members; 90% for the Cold War, 86% for the Vietnam War, and 8% for World War II.

21. AFDIL receives and processes all family reference samples and maintains the family reference database. The service casualty offices are responsible for identifying suitable

family references and sending the DNA collection kit to the family. The family receive the kit along with a shipping label for sending it to AFDIL.

22. AFDIL coordinates with respective service casualty offices to generate a list of potential donors to collect at the Family Member Updates. Additionally, DPAA will request family reference samples as needed if suitable ones are not currently in the database and AFDIL will process those references as priority once received.

23. AFDIL retains all family reference samples donated since the onset of collections and keeps the stored in a controlled environment that protects them from degradation. As technologies advance, AFDIL has the ability to retrieve these samples and retest them as needed.

24. Initially, family references from direct family members—to include brothers, sisters, and parents—of missing service members were targeted as a priority. The mothers, brothers and sisters are both good mtDNA references and good auSTR references, while the fathers and brothers are good Y-STR references.

25. Since 2006, AFDIL has broadened the collection to include all suitable Y-STR and auSTR family members. AFDIL is actively asking that at a minimum, two maternal (mtDNA testing), two paternal (Y-STR Testing) and two autosomal (auSTR testing) references be collected when possible. Not all extended relatives will have DNA usable as reference samples. See Exhibit 5, Family Reference Collection Form at 3 (link).

26. AFDIL's family reference database is protected under the Privacy Act and Health Information Portability and Accountability Act (HIPAA). See Exhibit 5, Family Reference Collection Form at 4. The database is only accessible by AFDIL scientists who have been approved to do comparison reports. It is not accessible to any outside individuals, including other DoD components.

AFDIL Past Accounting Program Procedures

27. Once samples are received by AFDIL for processing from the DPAA laboratory, the skeletal elements or biological material is signed over to an evidence custodian who photo-documents the remains and enters the information into the laboratory's information management system.

28. The Technical Leader assigns samples to a team and the evidence custodian signs the specimens over to a DNA analyst for processing. Case samples are processed on a rolling basis, in the order they are received, unless the DPAA laboratory changes the priority of a specific sample. AFDIL has approximately 600 samples in progress at any one time.

29. The samples are cleaned, ground into a powder, and the powder is dissolved, which release nuclear and mitochondrial DNA into a solution.

30. The DNA is then purified, concentrated and analyzed using mtDNA Sanger or NGS and/or Y-STR and/or auSTR testing methods. In 2013, success rates for mtDNA testing were 90%, but for STR testing were about 25% using organic purification methods. AFDIL looked at many different purification methods and identified a post PCR amplification purification kit that was shown to remove downstream inhibitors to sequencing. AFDIL forensically validated this kit for extract purification, which increased STR success rates to over 50%. AFDIL monitors results success rates and the current technological advancements and, through biweekly scientific meetings, establishes the requirements for developing, testing, validating and implementing technologies that will keep success rates high.

31. Each specimen is processed in duplicate, and the final results have to match in order for DNA results to be reported. This is a key aspect of AFDIL's quality assurance measures and was supported by the Defense Science Board 1995. See Exhibit 6, Defense

Science Board, The Use of DNA Technology for Identification of Ancient Remains (July 1995).

32. AFDIL performs two independent DNA analyses from the same skeletal specimens tested, using overlapping sequencing products, and dedicated separate laboratory rooms. When processing specimens in duplicate, each sample is extracted twice and processed to completion with the appropriate testing methods. To report out the duplicate extracts, the results between the individual extracts need to be consistent with one another; if the results are not consistent the samples are reported as "inconclusive." This differs dramatically from how modern criminal casework is processed at commercial, state and local laboratories, where a single extraction and analysis is sufficient to report out a result. Due to the low quality of the samples AFDIL receives, it is very easy to amplify a modern contaminant over the low quality authentic DNA; and it's why reproducibility of results are essential.

33. The average turn-around-time for processing a sample in duplicate (extraction to DNA summary report) is approximately 85 days.

34. The DNA results, when appropriate, are compared to the family reference database and these results are reported back to the DPAA Laboratory.

35. The entire testing procedure is carried out in the "blind"; this means that AFDIL DNA analysts do not know the potential identity of the individual for the specimen being tested. Analysts are informed of the conflict (i.e. Vietnam, Korea, or World War II), and where the remains were found, as environmental conditions specific to loss location will play a role in the extraction and DNA process. AFDIL maintains several validated extraction, purification, and testing procedures, which allows AFDIL to adapt to any sample type submitted for testing.

Challenges in Obtaining Results From Aged Remains

36. The biggest challenge to obtaining results from aged remains is DNA degradation

both from the environment (acidic soil, temperature, humidity) and post mortem effects (fire, chemical treatment, and time). As a result, the samples received by AFDIL's past accounting section that have not been chemically treated have an average mtDNA size between 100 and 300 base pairs and an average nuclear DNA size between 100 and 400 base-pairs. For remains that have been chemically treated, including many remains coming from Manila American Cemetery, the average size is significantly smaller. Modern DNA samples that have not been degraded generally have sizes greater than 400 base pairs.

37. To counteract the effects of degradation laboratories need the flexibility to employ a variety of testing strategies. Several of the strategies AFDIL developed or implemented are discussed above. In addition, in 2007, AFDIL was part of the developmental validation and one of the first laboratories to utilize MiniFiler, the first commercial STR system to target degraded samples. And in 2013, AFDIL forensically validated a low copy Y-STR testing method that increased success rates with degraded samples.

38. An additional challenge is finding appropriate references for the missing service members. Some service members have no apparent living relatives. Many others have no autosomal references (mother, father, brother, sister, children), but do have a maternal or paternal reference. Thus, the use of linage markers (mtDNA and Y-STR) as well as auSTR opens up the number of viable references and increases the chance of success.

39. The farther afield one goes for references, however, the more references may be needed. For example, cousins only share about 12.5% of the DNA with each other, so due to inheritance patterns, it would take more than two references from both the paternal and maternal side to develop a sufficient reference. Locating so many relatives becomes progressively more difficult as time passes.

* * * * *

Pursuant to 28 U.S.C. § 1746, I declare under penalty of perjury that the foregoing is true and correct.

Executed this 18 day of May, 2018.

TIMOTHY P. McMAHON
Director, DoD DNA Operations
Armed Forces Medical Examiner System

EXHIBIT 1

AMERICAN SOCIETY OF CRIME LABORATORY DIRECTORS
LABORATORY ACCREDITATION BOARD

August 27, 2014

Capt. Edward A. Reedy
Armed Forces Medical Examiners System
Armed Forces DNA Identification Laboratory
115 Purple Heart Drive, Dover AFB
Dover, DE 19902-5051

Dear Director Reedy:

On August 12, 2014, the ASCLD/LAB Board of Directors considered the application for ASCLD/LAB-*International* accreditation from the Armed Forces DNA Identification Laboratory. Based upon the documentation provided in the final assessment report and in accordance with the recommendation of Staff Assessor Jo Ann Given, the Board is satisfied that the laboratory meets or exceeds the requirements for accreditation as set forth in *ISO/IEC 17025:2005* and the applicable *ASCLD/LAB-International Supplemental Requirements*. As you know, the assessment and accreditation process also considered conformance with your program's own documented management system.

It is my pleasure to advise you that the Armed Forces DNA Identification Laboratory was accredited on August 12, 2014, in the Field of Forensic Science Testing. The specific scope of accreditation is declared on a Scope of Accreditation document provided as an attachment. The accreditation is for a period of four (4) years, ending on August 11, 2018. An original (full-size) accreditation certificate and original Scope of Accreditation document will be shipped to your attention via the United States Postal Service.

Accreditation is granted only after a thorough evaluation of the laboratory's management system and technical procedures and practices. Accreditation is the result of an extensive commitment of resources and much preparation by the management and personnel of the entire program. I commend the efforts of all employees involved in this achievement.

Accredited laboratories are expected to maintain the high standards which were required to achieve accreditation. In addition to maintaining conformance with accreditation standards, please read and ensure the laboratory's ongoing compliance with the enclosed obligations of ASCLD/LAB accredited laboratories.

The Armed Forces DNA Identification Laboratory will also be expected to participate in external proficiency testing and, where an approved test provider is available, agree for test results to be reviewed by the appropriate Proficiency Review Committee (PRC), as outlined in the *ASCLD/LAB Proficiency Review Program* document available on our website. The laboratory will be expected to conduct an annual, internal audit and submit a Performance Declaration to ASCLD/LAB in accordance with program requirements.

During the first accreditation cycle, accredited laboratories will undergo an on-site surveillance visit approximately every twelve (12) months. During subsequent accreditation cycles the frequency of surveillance visits will be determined by ASCLD/LAB based on the performance of the laboratory during the previous accreditation cycle and the full reassessment. In addition, during the first accreditation cycle, accredited laboratories shall submit a Performance Declaration to ASCLD/LAB thirty (30) calendar days prior to the on-site surveillance visit. During subsequent accreditation cycles, the frequency of Performance Declaration submission will be determined in conjunction with the determination of surveillance visit frequency. Surveillance visits and activities for the laboratory will be coordinated with you at the appropriate time.

The laboratory will be invoiced for an annual accreditation fee near the end of each calendar year. The fee is based on the number of proficiency tested personnel in the program. The invoice will be calculated in accordance with the budget approved at the annual ASCLD/LAB Delegate Assembly meeting, and it will include an additional amount to cover the cost of the annual surveillance visit.

As the director of an accredited laboratory you are a voting member of the ASCLD/LAB Delegate Assembly. You are invited and encouraged to participate in the accreditation process and to exercise your vote on issues which are presented to the Delegate Assembly. Should you desire to appoint an alternate delegate, please notify ASCLD/LAB in writing.

On behalf of the Board, I extend my sincere congratulations to you and to all personnel of your laboratory. If you have any questions or if we might assist you in any way please feel free to get in touch with us.

My best wishes to you and your staff.

Sincerely,

Renee Romero

Renee Romero
Chair, ASCLD/LAB Board

cc: Mark Wadhams, Quality Manager
John K. Neuner, ASCLD/LAB Executive Director
Pamela L. Bordner, ASCLD/LAB Senior Accreditation Program Manager
Troy Hamlin, ASCLD/LAB Accreditation Program Manager-Testing
Jo Ann Given, ASCLD/LAB Staff Assessor

American Society of Crime Laboratory Directors
Laboratory Accreditation Board

declares to all Advocates of Truth, Justice and the Law that the management and technical operations of the

Armed Forces Medical Examiners System
Armed Forces DNA Identification Laboratory

115 Purple Heart Drive, Dover AFB
Dover, Delaware 19902-5051

have been found through assessment to meet the requirements of
ISO/IEC 17025:2005
"General Requirements for the Competence of Testing and Calibration Laboratories"
the ASCLD/LAB-International Supplemental Requirements for Testing Laboratories: 2011
and all other requirements of the

ASCLD/LAB-*International*

program, and is granted this

Certificate of Accreditation

in the field of

Forensic Science Testing

for the categories of testing listed on the corresponding Scope of Accreditation

Renee Romero, ASCLD/LAB Chair

John K. Neuner, Executive Director

Pamela L. Bordner, Senior Accreditation Program Manager

Troy Hamlin, Accreditation Program Manager-Testing

Certificate Number

ALI-353-T

granted this

12th day of August, 2014

which expires on the

11th day of August, 2018

American Society of Crime Laboratory Directors / Laboratory Accreditation Board

ASCLD/LAB-*International* Program

SCOPE of ACCREDITATION

Corresponds to Certificate Number **ALI-353-T**

Name and Address of Accredited Laboratory

Armed Forces Medical Examiners System
Armed Forces DNA Identification Laboratory
115 Purple Heart Drive, Dover AFB
Dover, Delaware 19902-5051

Laboratory Contact Information

Capt. Edward A. Reedy, Laboratory Director
Phone: 302-346-8910
Fax: 302-346-8870
E-Mail: edward.a.reedy.mil@mail.mil

The management and technical operations of this laboratory were assessed and found to conform with **ISO/IEC 17025:2005**, the **ASCLD/LAB-*International* Supplemental Requirements for Testing Laboratories (2011)** and all other requirements of the ASCLD/LAB-*International* program. The laboratory was found to be competent and was accredited in the following area (s):

Field of Accreditation

Forensic Science Testing

Discipline (s)

3.0 Biology

Categories of Testing:

3.1 DNA - Nuclear
3.2 DNA - Mitochondrial

Customers Served: The Armed Forces DNA Identification Laboratory is a federal government laboratory that provides worldwide scientific consultation, research and education services in the field of forensic DNA analysis to the Department of Defense and other agencies.

Accreditation Dates

Date Granted: August 12, 2014
Date Expires: August 11, 2018
Date Last Updated: No Updates

Troy Hamlin
Accreditation Program Manager-Testing
ASCLD/LAB

EXHIBIT 2

AMERICAN SOCIETY OF CRIME LABORATORY DIRECTORS
LABORATORY ACCREDITATION BOARD

September 7, 2016

Lt. Col. Alice J. Briones
Armed Forces Medical Examiner System
Armed Forces DNA Identification Laboratory
115 Purple Heart Drive, Dover AFB
Dover, DE 19902

Dear Director Briones:

ASCLD/LAB has approved the continuation of accreditation of the Armed Forces Medical Examiner System - Armed Forces DNA Identification Laboratory, based upon the results of the surveillance visit, conducted by ASCLD/LAB Lead Assessor Barbara Caraballo on June 7-10, 2016.

Continuation of accreditation does not change the expiration date of accreditation. Rather, it is an annual, formal acknowledgement that a laboratory continues to operate in satisfactory conformance with all applicable accreditation requirements.

Two (2) Nonconformities (NCs) were issued as a result of the surveillance visit and your laboratory has responded appropriately to resolve the NCs. ASCLD/LAB does not view the issuance of these NCs as a reason to impose a sanction against the accredited status of your laboratory. Our focus is on monitoring your response to ensure that appropriate corrective action is taken within the specified time limits.

Enclosed with this correspondence you will find a final copy of the Surveillance Visit Report. As indicated in the final report, all other requirements checked during the surveillance visit reflected your commitment to ongoing conformance with accreditation requirements.

Congratulations for maintaining an on-going, acceptable level of conformance with ASCLD/LAB-*International* accreditation requirements.

Best Regards,

Brad Putnam
Accreditation Program Manager
ASCLD/LAB

cc: Mark Wadhams, Laboratory Quality Manager
ASCLD/LAB Office

ANSI-ASQ National Accreditation Board

ANAB

June 23, 2017

Timothy P. McMahon, PhD
Armed Forces Medical Examiner System
Armed Forces DNA Identification Laboratory
115 Purple Heart Drive, Dover AFB
Dover, DE 19902

Dear Director McMahon:

ANAB has approved the continuation of accreditation of the Armed Forces Medical Examiner System - Armed Forces DNA Identification Laboratory, based upon the results of the off-site surveillance review, completed by Lead Assessor Michael Kellett on June 20, 2017.

Along with this correspondence, you will find a copy of the Surveillance Report. Continuation of accreditation does not change the expiration date of accreditation. Rather, this action is an annual, formal acknowledgement that a forensic service provider continues to operate in satisfactory conformance with accreditation requirements.

Congratulations on your continued conformance with the International and ANAB accreditation requirements as well as your own management system. Should you have any questions, please feel free to contact us at QualityMatters@ANAB.org.

Sincerely,

Brad Putnam
Accreditation Manager
ANSI-ASQ National Accreditation Board

cc: Mark Wadhams, Laboratory Quality Manager
ANAB Office

Forensic Science International: Genetics 31 (2017) 198–206

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Research paper

Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: Implications for the use of a positive control in a next-generation sequencing procedure

Charla Marshall[a,b,*], Kimberly Sturk-Andreaggi[a,b], Jennifer Daniels-Higginbotham[a,b], Robert Sean Oliver[a,b], Suzanne Barritt-Ross[a,b], Timothy P. McMahon[a]

[a] Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL), Department of Defense DNA Operations, 115 Purple Heart Dr., Dover AFB, DE 19902, United States

[b] ARP Sciences, LLC, 9210 Corporate Blvd., Rockville, MD 20850, United States[1]

ARTICLE INFO

Keywords:
Hybridization Capture
Mitochondrial Genome
Degraded DNA
Next-Generation Sequencing

ABSTRACT

Next-generation ancient DNA technologies have the potential to assist in the analysis of degraded DNA extracted from forensic specimens. Mitochondrial genome (mitogenome) sequencing, specifically, may be of benefit to samples that fail to yield forensically relevant genetic information using conventional PCR-based techniques. This report summarizes the Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory's (AFMES-AFDIL) performance evaluation of a Next-Generation Sequencing protocol for degraded and chemically treated past accounting samples. The procedure involves hybridization capture for targeted enrichment of mitochondrial DNA, massively parallel sequencing using Illumina chemistry, and an automated bioinformatic pipeline for forensic mtDNA profile generation. A total of 22 non-probative samples and associated controls were processed in the present study, spanning a range of DNA quantity and quality. Data were generated from over 100 DNA libraries by ten DNA analysts over the course of five months.

The results show that the mitogenome sequencing procedure is reliable and robust, sensitive to low template (one ng control DNA) as well as degraded DNA, and specific to the analysis of the human mitogenome. Haplotypes were overall concordant between NGS replicates and with previously generated Sanger control region data. Due to the inherent risk for contamination when working with low-template, degraded DNA, a contamination assessment was performed. The consumables were shown to be void of human DNA contaminants and suitable for forensic use. Reagent blanks and negative controls were analyzed to determine the background signal of the procedure. This background signal was then used to set analytical and reporting thresholds, which were designated at 4.0X (limit of detection) and 10.0X (limit of quantiation) average coverage across the mitogenome, respectively. Nearly all human samples exceeded the reporting threshold, although coverage was reduced in chemically treated samples resulting in a ~58% passing rate for these poor-quality samples. A concordance assessment demonstrated the reliability of the NGS data when compared to known Sanger profiles. One case sample was shown to be mixed with a co-processed sample and two reagent blanks indicated the presence of DNA above the analytical threshold. This contamination was attributed to sequencing crosstalk from simultaneously sequenced high-quality samples to include the positive control. Overall this study demonstrated that hybridization capture and Illumina sequencing provide a viable method for mitogenome sequencing of degraded and chemically treated skeletal DNA samples, yet may require alternative measures of quality control.

1. Introduction

Next-generation sequencing (NGS) technologies have significantly expanded the ability to recover genetic information from ancient and degraded DNA samples. In particular, hybridization capture has proven useful for targeted enrichment of genomic DNA (gDNA) [1–5], as well as smaller targets such as the human exome [6] and the mitochondrial genome (mitogenome) [1,7]. Capture is amenable to DNA fragments

* Corresponding author at: Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL), Department of Defense DNA Operations, 115 Purple Heart Dr., Dover AFB, DE, 19902, United States.
E-mail address: charla.k.marshall.ctr@mail.mil (C. Marshall).

[1] Contractor supporting the Armed Forces Medical Examiner System.

http://dx.doi.org/10.1016/j.fsigen.2017.09.001
Received 12 May 2017; Received in revised form 8 August 2017; Accepted 5 September 2017
Available online 08 September 2017

that are too degraded for NGS workflows involving PCR enrichment, which are more common to forensics [8–12]. The technological advancement offered by hybridization capture and NGS therefore create the potential to obtain genetic data from forensic specimens that were heretofore too degraded for DNA analysis. In particular, next-generation mitogenome sequencing holds promise for human identification efforts and missing persons cases such as those of the Defense Personnel Accounting Agency (DPAA). Working in conjunction with DPAA, the Armed Forces Medical Examiner System's Armed Forces DNA Identification Laboratory (AFMES-AFDIL) performs all of the DNA testing to assist in the identification of missing service members from past military conflicts. The samples typically submitted by DPAA to the AFMES-AFDIL for genetic testing are aged skeletal remains that were exposed to post-traumatic (e.g., fire) and environmental insults and sometimes chemical treatment with formaldehyde or powdered hardening compounds. As a result, endogeneous DNA is often limited in quantity, degraded and damaged, and plagued with co-extracted inhibitory molecules common to ancient DNA such as microbial DNA and humic acid. Though the AFMES-AFDIL currently has a > 90% mtDNA success rate using traditional Sanger-type-sequencing (STS) methods [13,14], an NGS method for mitochondrial genome sequencing was needed for the most challenging samples that yield partial or unreproducible results from standard techniques.

The present study summarizes a performance evaluation of a mitogenome capture and Illumina sequencing procedure for non-probative case-type samples. The procedure is a low-throughput, labor-intensive process with numerous tube transfers requiring approximately ten hours of hands-on time. The laboratory workflow involves several clean-lab steps, including an initial enzymatic damage mitigation step to minimize cytosine deamination typical of aged and degraded DNA [15], followed by purification and dual-indexed Illumina library preparation. In the post-amplification laboratory, a limited-cycle PCR completes the library preparation procedure, then indexed libraries are quantified to confirm successful amplification. Libraries are subsequently enriched for the mitogenome in an overnight in-solution hybridization incubation with biotinylated RNA baits. After DNA capture, purification, and amplification of the captured product, samples are purified and quantified before being pooled for paired-end sequencing on an Illumina MiSeq. The procedure takes roughly two weeks to complete from bone preparation to mitogenome profile for a sample set of three case samples and associated controls. As summarized below, 40 non-probative, case-type samples libraries and associated controls were processed for a total of more than 100 libraries sequenced by ten DNA analysts over the course of five months. These NGS data confirm the reliability, specificity, cleanliness, and sensitivity of this novel forensic DNA testing procedure.

2. Materials and methods

2.1. Samples

A total of 22 previously reported and therefore non-probative AFMES-AFDIL case samples were selected for this study, and were categorized into four types: chemically treated, degraded, high quality, and nonhuman (Table S1). The chemically treated samples represent a set of Korean War unknowns that were treated with formaldehyde before being buried in the National Memorial Cemetery of the Pacific (NMCP) in Honolulu, Hawaii. The degraded samples were chosen from World War II (WWII), Korean War, or Vietnam War cases. Two high-quality, non-probative bone samples were included for the purposes of the mixture study. All 22 non-probative case samples were previously tested with Sanger-type sequencing (STS) technology (following methods outlined in [13,14,16]), indicating a range of mtDNA quality. The high quality samples generated > 1000-bp amplicons, the degraded and nonhuman samples generated 150–300-bp amplicons, and the chemically treated samples yielded either very small amplicons or failed to amplify. This sample quality information was utilized to assess the sensitivity of the procedure in lieu of a traditional sensitivity series with control DNAs, which lack the environmental contaminants that necessitate the capture procedure.

2.2. Contamination monitoring and controls

DNA extraction and library preparation were carried out in a pre-PCR DNA laboratory dedicated for low copy and degraded DNA sample testing. Reagent blanks (RBs) were introduced during DNA extraction and were carried through the entire downstream procedure. A non-template library negative control (NC) and a library positive control (PC) were introduced during library preparation. The inclusion of a PC is unconventional for the ancient DNA community because the PC itself presents a potential source of contamination. However, the Federal Bureau of Investigation (FBI) Quality Assurance Standards (QAS) mandates that a PC be utilized to monitor the success of a DNA testing procedure [17]. In order to adhere with the QAS for accreditation purposes, this study employed a PC but utilized the following measures to minimize its potential to contaminate associated samples: 1) only 1.0 ng of fragmented gDNA was utilized in library prep (four-fold less DNA than the manufacturer's recommendation), 2) the PC was spiked into the library pool for sequencing to control the number of reads that it consumed (see below), and 3) the control DNA exhibited a rare mitogenome profile (K562; Promega Corporation, Madison, WI) to be readily identified in the sequence data (Table S2). The K562 control DNA was enzymatically fragmented with NEBNext Fragmentase (New England Biolabs Inc. (NEB), Ipswich, MA) in a separate pre-PCR laboratory, then diluted to 0.2 ng/μL prior to being taken to the degraded sample laboratory for library preparation. Overall, a total of 27 RBs, 18 NCs and 17 PCs were processing simultaneously with the non-probative samples.

2.3. DNA extraction and USER treatment

DNA was isolated from 0.2 g–1.0 g powdered bone sample. Bone powder was first demineralized in a buffer containing 0.5 M EDTA, 1% sarkosyl and 20 mg/ml proteinase K by incubating overnight with agitation at 56 °C [18]. DNA was purified using one of three protocols: organic [18], QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany) [19], or a variation of the [19] procedure that utilized the MinElute PCR Purification Kit (QIAGEN) instead of the QIAquick. As with the PC, the high quality DNA samples were enzymatically fragmented with NEBNext Fragmentase following the manufacturer's protocol in order to make the DNA suitable for library preparation. When additional bone sample was available, replicate DNA extracts were prepared for reproducibility testing (Table S3). Additionally, three mixtures were created from the remaining DNA extract volume after neat sample processing of the two high quality DNA extracts. Mixtures of 30 μL were created using a volume-based mixing scheme in ratios of 1:9, 1:1, and 9:1. To keep the input volumes consistent with neat samples, mixed samples were suspended in 70 μL volumes using Tris-EDTA (10 mM Tris-HCl, 0.1 mM EDTA, pH 7.5).

In the absence of a commercially available kit for DNA fragments averaging ~75 bp (such as those intended for the present procedure), a quantitative PCR (qPCR) was not performed. Instead, total extracted DNA (human and microbial) was quantitated to confirm the success of the DNA extraction procedure. An aliquot of each DNA extract was quantified using the Agilent 2100 Bioanalyzer using the High Sensitivity (HS) or 7500 assay (Santa Clara, CA) to confirm the success of DNA extraction. All DNA extracts were subjected to a 1-h Uracil-Specific Excision Reagent (USER) (NEB) treatment to mitigate the effects of cytosine deamination [20,21], followed by purification with the MinElute kit with elution in Tris-EDTA.

2.4. Library preparation

Each library preparation sample set included one to three case samples and one or two associated RBs, a library NC, and a library PC (fragmented K562). To increase sample size and to evaluate the reproducibility of the procedure, replicate libraries were prepared from the same DNA extract when excess volume was available. In total, 40 DNA libraries were prepared from the 31 DNA extracts (including the three intentional mixtures) by ten different scientists. The libraries in each sample set were taken through downstream processing together. Quantitation of samples was completed using the Qubit 2.0 Fluorometer (Invitrogen, Waltham, MA) to calculate DNA quantity for input into library preparation. DNA libraries were prepared using the NEBNext Ultra DNA Library Prep kit for Illumina (NEB) with the following modifications to the manufacturer's version 1.0 protocol: 1.3 x AMPure XP (Beckman Coulter, Brea, CA) purification of adapter-ligated samples, and elution with 42 µL of Tris-EDTA. Two PCR enrichment reactions of 20 µL adapter-ligated DNA were prepared using the NEBNext Multiplex Oligos for Illumina Dual Index Primers Set 1. Twelve cycles of PCR were carried out on a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA) following the NEBNext High-Fidelity 2X PCR Master Mix recommendations. The two PCR reactions were pooled and purified with AMPure XP using a 0.9 x bead:sample ratio, then eluted in 20 µL of Tris-EDTA. Successful amplification of DNA libraries was confirmed using the 2100 Bioanalyzer HS or 7500 assay by the presence of a DNA smear in the 170–1000 bp range.

2.5. Hybridization capture

An in-solution hybridization capture method using a custom MYbaits RNA probe array (MYcroarray, Ann Arbor, MI) was employed to enrich for the mitogenome [2,7]. This kit includes 20,000 copies of over 1800 unique, 50 bp–70 bp biotinylated RNA oligos that scaffold the revised Cambridge Reference Sequence (rCRS) [22,23] with approximately 5 x tiling (Fig. S1). Probe density was augmented in the mitochondrial control region (CR) to improve CR coverage for comparison with family reference specimen (FRS) CR data. The bait sequences consist primarily of rCRS sequences, although sequences with commonly observed variant motifs in the AFMES-AFDIL FRS database were included in the design (Table S4). The hybridization conditions followed the MYbaits V2 protocol with a 24-h incubation at 62 °C using the Veriti thermal cycler (Applied Biosystems). This hybridization temperature was chosen to enable hybridization between the baits and mtDNA sequences from diverse human haplotypes. This was a modification to the manufacturer-recommended 65 °C hybridization temperature for greater specificity and higher on-target percentage. The captured product was amplified in four PCR reactions per sample, each having 10 µL of captured DNA, using the Herculase II Fusion Enzyme with dNTPs Kit (Agilent) with HPLC-purified universal Illumina primers (P5: 5′-AAT GAT ACG GCG ACC ACC GA-3′; P7: 5′-CAA GCA GAA GAC GGC ATA CGA-3′) (Integrated DNA Technologies, Coralville, IA). Nineteen PCR cycles were completed on a 9700 thermal cycler. Amplified capture product was pooled by sample and purified with 1X AMPure XP, and eluted in 25 µL Tris-EDTA.

2.6. Library pooling and multiplexed sequencing

Each sample set was sequenced individually by pooling the two to three case samples, associated RB(s), and the NC by volume without the PC. High quality samples were sequenced separately from degraded and chemically treated samples. The molarity of the pool was determined using the 2100 Bioanalyzer. Each pool was diluted to 4 nM and prepared for sequencing following the manufacturer's recommendations. The associated PC and Phi-X Sequencing Control V3 (Illumina, San Diego, CA) were diluted and denatured separately, then spiked into the pool for loading. A final loading concentration of approximately 8 pM was targeted with 95% 20 pM sample pool, 2.5% 12.5 pM PhiX, and 2.5% 12.5 pM PC. Paired-end sequencing was completed using an Illumina MiSeq Reagent Kit v2 (300-cycle, 2 × 150) on the MiSeq Desktop Sequencer (Illumina). A total of 17 MiSeq runs were performed.

2.7. Read mapping and data analysis

MiSeq Reporter software (Illumina) generated demultiplexed FASTQ files from the raw data, and reads were imported into the CLC Genomics Workbench v7.5.1 (QIAGEN) for data analysis. The following steps were incorporated into an automated workflow to analyze sample sequence data and generate a mtDNA profile. Paired reads were mapped to the rCRS using stringent parameters (0.85 length fraction and 0.95 similarity fraction) to minimize the inclusion of off-target reads. As such, a sequence read of 75 bp, for example, must match the rCRS for a minimum stretch of 65 bp and within those 65 bp only three differences from the rCRS are permitted. These mapping parameters may result in reduced coverage of insertion/deletion (indel) clusters and SNP clusters depending on DNA fragment length and individual haplotype; but they are necessary to prevent bacterial DNA from mapping to the rCRS. Mapped duplicates were removed, and a local realignment was performed to assist in gap (indel) alignment. Coverage metrics were generated per individual read mapping, and only paired reads were analyzed in order to exclude broken pairs from instrument crosstalk in coverage calculations.

When paired coverage met or exceeded a 10X variant calling threshold, variants were reported if the variant frequency (VF) was ≥10% and variant count was ≥4 to overcome background noise from DNA damage and/or PCR and sequencing error. To conform the resulting variant profile to forensic guidelines (with 3′ alignment of indels and use of appropriate International Union of Pure and Applied Chemistry (IUPAC) codes for mixed positions), a custom mtDNA analysis tool developed by the AFMES-AFDIL for the CLC Genomics Workbench known as AQME (AFDIL-QIAGEN mtDNA Expert) was employed [24]. Haplogroups were determined using AQME, which performs an alignment of the consensus sequence against virtual haplogroup sequences in Phylotree Build 16 [25]. The highest scoring haplogroup based on shared polymorphisms was reported.

2.8. Concordance

Previously obtained STS sequence data from the mtDNA control region (CR) were used for concordance evaluation. The STS data originated from either the same skeletal element and/or an associated FRS. In some cases, identification of the remains was based solely on anthropological analysis [26,27] and/or circumstantial evidence; therefore these cases lacked direct STS data from the same skeletal element for comparison. Regardless, mtDNA comparison data from the CR were available for all non-probative case samples − either from data generated directly from the sample or from a FRS, or both.

3. Results and discussion

3.1. MiSeq run assessment

The sequencing runs produced an average cluster density of 740.41 K/mm^2, with an average of 11.93 million (M) reads passing filter (PF) (Table 1). The 17 runs produced an average of 87.46% of bases with a quality score greater than 30 (Q30, which corresponds to a 99.9% base call accuracy), and all but two runs (with 79.2% and 79.3% Q30) exceeded the manufacturer expectation of 80% Q30. One run produced a cluster density roughly twice the overall average at 1399 K/mm^2. This run contained high quality samples that may have affected accurate quantitation of the pool on the Bioanalyzer. At the other

Table 1
Summary of the run metrics from the 17 MiSeq runs.

Run Metric	Minimum	Maximum	Average	Standard Deviation
Cluster Density (K/mm^2)	136.00	1,399.00	740.41	319.35
Clusters Passing Filter (PF) (%)	76.51	95.43	89.08	5.18
Reads (Million (M))	2.77	25.27	13.70	5.60
Reads PF (M)	2.43	19.35	11.93	4.36
Q30 (%)	79.20	94.20	87.46	4.41
Aligned (%) (Phi-X)	1.48	21.96	5.14	4.79
% Reads PF Identified	65.52	96.30	86.51	9.08
% PC	2.18	38.49	9.02	8.82
% Sample Pool	27.03	93.05	77.49	15.69

extreme, one run produced much lower cluster density at 136 K/mm^2, and this run contained three chemically treated samples that produced low coverage data. The aligned reads (Phi-X) percentage was 4.09% on average, slightly higher than the 2.5% expected from the pool. Despite this variability, the Phi-X control was successfully sequenced and aligned to the reference genome in all runs, demonstrating its utility as a sequencing control. The PC comprised an even higher average of 9.02% of the sequenced reads despite the targeted proportion of 2.5%. And maximally the PC consumed over one-third of the sequenced reads (in the pool with the lowest cluster density) at 38.49% of all reads identified passing filter (PF). The percent of total reads apportioned to control DNAs (Phi-X and PC) was dependent on the overall quality of the pool; lower quality pools generated less reads per sample and consequently a higher percentage of reads were assigned to the PC and Phi-X control (Table S5). Despite the variation observed across the sequencing runs, each was successful and produced high quality DNA sequence data.

3.2. DNA quantity

Sample quality, as determined from STS success, was consistent with DNA quantitation results (Table 2). Higher quality DNA samples resulted in higher DNA concentrations after DNA extraction, resulting in increased library DNA yield. All human and nonhuman samples including the PC generated sufficient DNA libraries for hybridization capture, whereas the RBs and NCs contained only dimer products in Bioanalyzer traces after library purification (but were captured and sequenced regardless). RBs had more primer dimer than NCs due to their earlier initiation than NCs in the NGS procedure. As shown in Fig. 1, samples with increased DNA input generally produced more capture product. The chemically treated samples fell into one of two clusters – the first with the RBs (representing quantifiable primer dimer) and the second with other non-probative case samples. Therefore the capture DNA yield from the lowest-template chemically treated samples was of comparable DNA concentration to the primer dimer quantified from several RBs. The PCs exhibited consistently high

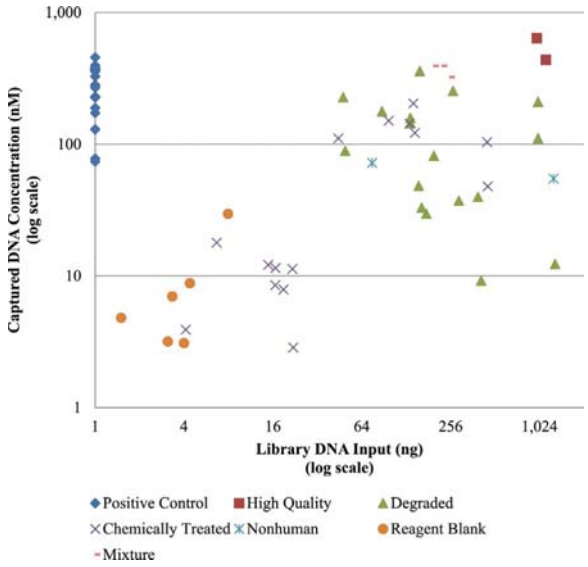


Fig. 1. Capture yield in relation to DNA input by sample type. Scatterplot excludes samples and controls with no quantifiable DNA (0 ng).

capture product (each > 77 nM) despite the reduced library input of one ng, likely due to the pristine nature of the human DNA present in the control sample. Overall, the samples with higher DNA quantity prior to library PCR also yielded more capture product for NGS.

3.3. DNA sequence data

The total number of reads per sample was variable (Table 3), yet this was expected due to the variability in run composition (e.g., two samples and two RBs vs. three samples and one RB) as well as variation in the number of reads PF per run. Each non-probative case sample garnered an average of eight million reads, whereas RBs and NCs produced substantially fewer than reads on average (250,000 and 50,000, respectively). RBs produced more reads than NCs, although a smaller portion of the RB reads mapped to the rCRS (5.1%) than NC reads (17.3%). When considering the non-probative case samples, the higher the sample quality, the higher percentage of reads that mapped to the rCRS. However, even the PC and high quality samples exhibited high proportions of unmapped reads (34.5% and 54.2%, respectively). It is possible that the stringent mapping parameters precluded some human mtDNA sequences from mapping to the rCRS, such as reads containing indels or PCR/sequence errors. Yet preliminary mapping of the high quality sample dataset to the human reference genome using the same mapping stringency (0.85 length fraction and 0.95 similarity fraction) indicated that the remaining reads were likely of nuclear DNA origin. Although some inefficiency in capture is expected based on findings from the ancient DNA literature (e.g., [2]), future optimization of the extraction and capture procedures aim to improve on-target read

Table 2
Quantitation results by sample type.

		Library dsDNA Input (ng)		Library DNA Yield (ng/μL)		Capture DNA Yield (nM)	
Sample Type	Library Count	Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Positive Control	17	1.0	0.0	5.0	3.6	275.3	116.7
High Quality	2	1055.4	104.7	58.0	7.6	538.5	142.3
Degraded	18	310.6	349.3	14.3	20.7	119.5	103.5
Chemically Treated	15	106.3	150.2	8.7	18.8	64.0	68.1
Reagent Blank	27	0.9	2.0	0.4	0.6	10.3	14.0
Negative Control	18	0.0	0.0	0.1	0.2	4.6	9.2
Nonhuman	2	674.5	847.5	5.3	1.0	63.5	12.3
Mixture	3	224.1	27.6	88.4	9.6	369.9	40.2

Table 3
Sequence data by sample type.

Sample Type	Library Count	Total Reads		Reads Mapped to rCRS (%)		# Unique Reads Mapped to rCRS		Mean Mapped Read Length (bp)		Average Coverage		% of Mitogenome Covered ≥ 10X	
		Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.	Avg.	St. Dev.
Positive Control	17	1,683,300	831,602	65.1	9.8	628,553	253,332	125.5	6.3	2828	1148	100	0.0
High Quality	2	6,253,915	149,311.3	45.8	7.3	1,995,842	552,722.7	119.1	3.0	8291	1878	100	0.0
Degraded	18	9,480,890	6,391,180	13.4	12.5	337,680	326,530	81.9	13.2	822	797	99.9	0.2
Chemically Treated	15	6,134,300	6,205,300	13.3	20.9	37,000	55,850	68.5	11.9	76	110	61.7	44.9
Reagent Blank	27	250,400	423,300	5.1	8.4	580	300	113.0	14.4	1.5	2.1	2.4	5.8
Negative Control	18	42,000	109,000	17.3	21.9	390	500	112.7	17.3	0.4	0.5	0.4	0.8
Nonhuman	2	9,782,069	3,152,937	0.0	0.0	96.9	117	100.1	2.7	1.9	1.0	1.8	0.7
Mixture	3	11,999,350	1,980,190	42.7	4.4	3,555,603	276,754.6	113.3	3.8	13,428	1608	100	0.0

percentage [28–30]. Compared to the PCs and high quality samples, the degraded and chemically treated samples generated much less on-target reads, averaging just 13%. Without capture, the chemically treated samples were shown to exhibit less than 0.01% human mtDNA, 15% bacterial DNA, and 85.5% DNA with no matches in GenBank [31]. Thus, the capture procedure effectively enriched the small amount of endogenous human mtDNA present in the chemically treated samples. Although outside the scope of this report, future investigation into the sources of the unmapped reads may shed light on the postmortem microenvironment of the skeletal remains. From the reads that successfully mapped to the rCRS, only 337,676 unique mitogenome reads remained on average from the degraded and 37,004 from the chemically treated samples after removal of mapped duplicates. Therefore despite an average library input of more than 100 ng gDNA for chemically treated and degraded samples, these samples are low in mtDNA template because human DNA composes only a very small fraction of the total DNA extracted.

Table 3 furthermore shows that mapped read length increased with increasing sample quality. The PCs and high quality samples received the same DNA fragmentation treatment; therefore the mean read lengths of these two sample types were similar (approximately 120 bp). Degraded and chemically treated samples expectedly produced shorter read lengths, averaging 81.9 bp and 68.5 bp, respectively. Interestingly, mean read length of the RBs and NCs was > 110 bp. Therefore it is unlikely that the mapped reads in the RBs and NCs originated from co-processed case samples due to the observed length discrepancy. Similar to read length, coverage also increased with increasing sample quality (Fig. 2). The PCs exceeded an average of 2800X average coverage, and all PCs produced over 1000X average coverage. These results from the 17 PCs demonstrate that one can expect consistent, robust results from one ng dsDNA input of pristine, fragmented DNA. The high-quality case samples, which garnered three times the reads of most PCs, averaged more than 8000X average coverage across 100% of the mitogenome. In contrast, the degraded samples resulted in an average of 820X average coverage, but coverage depth varied considerably by sample. Out of 18 tested, only one degraded sample had incomplete (< 100%) coverage of the mitogenome above the 10X variant calling threshold, which brought the degraded sample average down to 99.89%. This degraded sample (17A) produced an average coverage depth of 7.4X with 24.43% of bases covered above the 10X variant calling threshold. In retrospect, it was identified that more than five µg of DNA was used in library prep of 17A. Thus, the incomplete coverage can be attributed to an error in processing, as the DNA extract was not appropriately diluted to the maximum input of 1 µg dsDNA prior to library preparation. When this outlier is excluded, all 17 of the degraded case samples resulted in 100% mitogenome coverage with an average coverage depth of 914X. The chemically treated samples exhibited the least mitogenome coverage of the human sample types, and coverage also varied considerably by sample. Average depth of coverage was 76.4X, and the portion of mitogenome coverage averaged 61.7%. Coverage was minimal in the nonhuman samples, consistent with data generated from RBs and NCs. Thus, nonhuman mtDNA is not captured and sequenced as efficiently as human mtDNA using the present protocol.

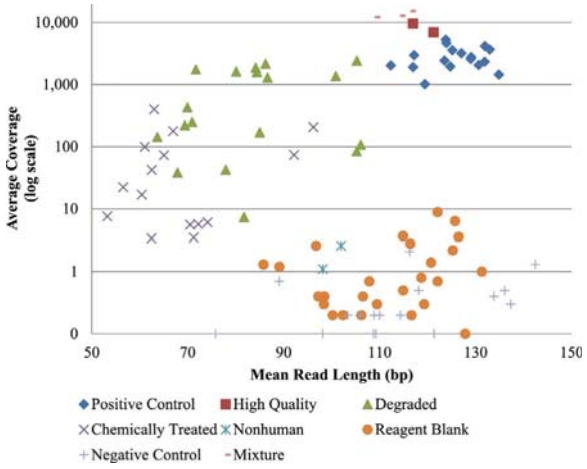


Fig. 2. Reads that mapped to the rCRS by sample type.

3.4. Background signal assessment

The limit of detection (LOD) and limit of quantitation (LOQ) were calculated from the mitogenome coverage of the control blanks (RBs and NCs). The LOD and LOQ represent the 99.73% and > 99.99% confidence limits of average coverage, respectively, and were utilized to establish the analytical (LOD) and reporting (LOQ) thresholds. By constructing the analytical and reporting thresholds from the control blank data, the background signal of the entire procedure was captured − from DNA extraction through sequencing. Of the 45 control blanks that were sequenced, two reagent blanks were determined to be contaminated (discussed below) and excluded from the background assessment. A total of 43 remaining control blanks (25 RBs and 18 NCs) were used to determine the background signal of the procedure. The LOD (average + 3 standard deviations (SDs)) was 3.599X average coverage, which was rounded to 4.0X for simplicity. Mapped sequence data exceeding 4.0X average coverage were therefore above background noise (analytical threshold). The LOQ (average + 10 SD) was 10.217X average coverage, rounded to 10.0X. This 10.0X average coverage LOQ was determined to be the threshold for reporting sample data generated from a single library (reporting threshold). Sample data between 4.0X and 9.99X average coverage would require additional sequencing of the captured library to meet the 10.0X minimum

coverage threshold for reporting. For resequenced libraries from the same capture product, mapped duplicates should be removed simultaneously to avoid the inclusion of separately sequenced PCR duplicates in variant calling.

When the analytical (LOD) and reporting (LOQ) thresholds were applied to the non-probative sample data, all PCs and high quality samples (including the mixtures) passed the reporting threshold. Seventeen out of eighteen (94%) degraded samples passed the reporting threshold, yet one sample fell between the analytical and reporting thresholds (17 A) with 7.4X average coverage. As discussed above, this degraded sample had excess DNA in library preparation that likely inhibited library preparation. The chemically treated samples yielded more variable results: only 60% (9 of 15) passed the reporting threshold, 27% (4 of 15) passed the analytical threshold, and 13% (2 of 15) failed to exceed 4.0X average coverage. Although the results are less robust for chemically treated samples than the nearly 100% passing rate for degraded DNA samples, the NGS results outperformed the 6% Sanger sequencing success rate from the same samples (data not shown). Of further note is the finding that the contaminated RBs that were excluded from the LOD and LOQ calculations were shown to be above the analytical threshold yet below reporting threshold. Consequently a 4.0X maximum coverage threshold for control blanks is consistent with the presense of contamination determined by the analyst. Finally, the nonhuman samples failed to meet the analytical threshold with no reportable data.

3.5. Mitogenome profile assessment

The variant profiles generated from the non-probative samples and PCs are summarized in Table 4, and haplotypes are shown in Table S6. The haplotypes of samples above the analytical threshold were consistent among 11 of the 12 cases within NGS replicates. One sample (4A) appeared to be contaminated with DNA from an unknown source, although the associated RB was clean. The major profile of 4A was consistent with the NGS data from the sample replicate. The 4A extract was generated from 0.2 g tooth powder four years prior for STS processing, and the extract was handled several times for PCR amplification attempts before NGS was initiated. Therefore the low-level (~10%) contaminant affecting sample 4A may have been introduced prior to the start of the present study. Importantly, the reduced average VF (91.85%) and seven nucleotide positions with mixed bases observed in the 4A profile were not observed in the replicate 4B profile generated from a freshly prepared DNA extract. Therefore due to the sensitivity of NGS, starting from freshly prepared DNA extracts to minimize repeated tube handling may be worthwile to prevent unnecessary reprocessing. Despite the one instance of a low-level contaminant, the NGS data, which were replicated by the same and different scientists, were determined to be repeatable and reproducible.

The NGS profiles were then compared with STS CR data in order to authenticate the mitogenome data obtained. When consistent with the known CR profile, the data were considered to be single-source. When multiple nucleotide positions with two or more bases (mixed positions) were observed or inconsistencies were identified between NGS and STS, the profiles were determined to be mixed (Table 5). All PCs and the two high quality samples exhibited single-source profiles that were consistent with STS (Table 5). Only minor fluctuation in low-level variants (5%–10%) was observed in the PCs (Table S7). The three intentional mixtures (i.e., volume:volume mixtures) exhibited reduced average VF and/or excess mixed positions compared to the high-quality samples and PCs. One of the three intentional mixtures was a 1:9 mixture that appeared to be umixed because it showed zero minor variants above the 10% variant calling threshold. The mixture was identifiable, however, because of the reduced average VF (95.76% for the 1:9 mixture compared to an average of 99.13% in the neat components). The remaining two intentional mixtures each produced 31 mixed positions. Hence, the two metrics of average VF and number of mixed positions across the mitogenome may be useful for mixture detection. By applying a 99.73% confidence interval (average − 3 SDs) to the average VF of single-source degraded and chemically treated samples (above the reporting threshold), a VF reduction below 94.0% may indicate a mixture. Given that average VF is haplotype- and sample-type- dependent, and that it encompasses the signal of background noise, this metric serves only as a means to flag potential mixtures for further scrutiny. And secondly, an excessive number of nucleotide positions with two or more bases may indicate a mixed profile, as three point heteroplasmies (PHPs) is the maximum observed in a single-source profile based on high quality STS haplotypes [32].

The majority of degraded (16 of 18 (89%)) and chemically treated (11 of 13 (84%)) samples above the analytical threshold produced profiles that were single-source and concordant with Sanger CR data. Variants were largely consistent with the expected profile. Only three unreplicated instances of mixed positions were observed among the single-source degraded and chemically treated case samples, all of which were consistent with cytosine deamination (4464 G(89.8%)/A(10.2%) in sample 6A1, 16519 C(88.6%)/T(11.4%) in sample 9B, and 3498 C(90.0%)/T(10.0%) in sample 15). Nucleotide position 16519 was previously noted by Rathbun et al. to be a damage hotspot in their CR analysis [9], further supporting the hypothesis that DNA damage is evident in these data. Therefore the initial USER treatment to remove damaged cytosine bases may not be 100% effective. Two degraded sample libraries generated from the same DNA extract (12A1a and 12A1b) were inconsistent with Sanger results and determined to be mixed with a co-processed, degraded sample (16A). Due to the presence of contamination in both 12A1 libraries (a and b), it is possible that only one cross-contamination event occurred before the libraries were prepared. The contaminating case sample was of much higher quality with > 30% of reads mapping to the rCRS and > 1700X average coverage in both replicates (16A1a and 16A1b).

The remaining instances of contamination also originated from co-processed high quality samples. Two of the three such instances were attributed to the PC, which affected one chemically treated DNA extract (2A2) and one RB (RB-6). A second RB (RB-7) produced a profile consistent with an associated high quality sample (18). Average coverage for each of these three contaminated profiles was above the analytical but below the reporting threshold, with 2X-9X average coverage originating from the contaminant. The contamination was evident in the RBs (at a rate of 7.4%), and was a detriment to data interpretation of the low-quality samples for which the procedure was developed. Of the 31 degraded and chemically treated sample libraries that were sequenced, three (9.7%) were contaminated with sequence data from a high quality sample or PC. This rate of contamination was not expected given the measures that were taken to minimize contamination both in the laboratory and during the sequencing runs. As explained in the methods above, the clean-lab steps prior to indexing were performed by trained DNA analysts who routinely process degraded DNA samples from sample preparation through sequencing. Therefore cross-contamination between co-processed samples was expected to be minimal. Moreover, the treatment of the PC was done in a way to minimize the potential for cross-contamination and crosstalk. This included artificial degradation and dilution of the K562 control DNA, barcoding the PC to remove PC sequences from the sample data bioinformatically, using a relatively low 8 pM loading concentration to prevent overclustering on the flow cell, and spiking the PC into the pool at a limited quantity to minimize its impact on the sequencing run.

In order to evaluate the possibility that the contamination occurred from index cluster crosstalk during the sequencing runs, seven sample pools were resequenced without the PC (using 97.5% sample pool and 2.5% Phi-X control). Additionally, the high quality sample (18) that was believed to be the source of the RB-7 contamination was not included in its pool. As shown in Table 6, the PC was entirely removed from the sequence data in the runs without the PC and was no longer a source of contamination. And, the RB that was contaminated with the co-

Table 4
Mitogenome profile summaries.

Sample ID	Sample Type	Average Coverage	% of mtG Covered ≥ 10X	Haplogroup	# Variants Called	Average Variant Forward/Reverse Balance	Average Variant Quality	# mixed positions	Average VF
1	Chemically Treated	22.3	89.28	H4a1a1	13	0.39	37.50	0	97.27
2A1	Chemically Treated	3.4	0.93	Undefined	0			0	
2A2	Chemically Treated	5.6	8.28	Undefined	4	0.39	36.76	3	70.24
3A1a	Chemically Treated	400.5	100	H4a1a4b	21	0.34	34.51	0	96.65
3A1b	Chemically Treated	17	80.9	H4a1a4b	14	0.37	36.69	0	97.78
3B	Chemically Treated	42.5	99.98	H4a1a4b	21	0.43	37.01	0	100.00
4A	Chemically Treated	206.4	100	H1 + 16189	16	0.42	35.98	7	91.85
4B	Chemically Treated	178.4	100	H	14	0.44	36.96	0	98.87
5	Chemically Treated	7.6	22.63	H44a	3	0.40	35.89	1	96.67
6A1a	Chemically Treated	5.7	9.8	H4a	4	0.21	37.72	0	100.00
6A1b	Chemically Treated	73.9	100	H4a1a1a	22	0.36	34.47	2	93.53
7A1	Chemically Treated	99.6	99.96	U2e3a	46	0.40	36.43	1	95.68
7A2	Chemically Treated	72.9	99.92	U2e3a	42	0.45	36.80	1	96.01
8	Chemically Treated	3.5	2.63	Undefined	1	0.31	35.69	0	100.00
9A1	Degraded	143.9	100	I4a	36	0.44	36.64	0	98.55
9A2	Degraded	221.3	100	I4a	37	0.46	36.59	1	93.95
10A1	Degraded	84.5	99.38	V1a1	18	0.41	36.90	0	98.94
10A2	Degraded	2,417.6	100	V1a1	17	0.41	37.15	0	98.95
11A1	Degraded	1,589.7	100	H2a2a1	4	0.37	36.54	2	88.18
11A2	Degraded	2158	100	H2a2a1	4	0.37	36.48	2	87.11
12A1a	Degraded	42.6	99.76	T2a1b1a1	46	0.43	36.37	27	85.25
12A1b	Degraded	38.3	99.12	T2a1b1a1	40	0.44	37.36	12	93.62
13	Degraded	1,366.9	100	T2b	38	0.40	35.61	1	97.25
14A1a	Degraded	1,296.2	100	B4b1a2a	32	0.43	36.04	0	97.70
14A1b	Degraded	429.1	100	B4b1a2a	32	0.43	36.89	0	98.53
15	Degraded	249	100	B4e	41	0.34	36.40	1	95.28
16A1a	Degraded	1,745.3	100	N9b1c	28	0.48	37.06	0	97.36
16A1b	Degraded	1,868.6	100	N9b1c	29	0.46	35.17	0	95.69
16B1a	Degraded	1,615.3	100	N9b1c	27	0.48	36.24	0	97.31
16B1b	Degraded	170.2	100	N9b1c	28	0.43	36.90	0	97.73
17A1	Degraded	7.4	24.43	T	10	0.34	37.09	0	98.48
17A2	Degraded	108	100	T1a	34	0.36	37.09	1	95.86
18	High Quality	6946	100	V10a	18	0.36	37.49	0	98.72
19	High Quality	9,580.5	100	U1b	37	0.38	37.60	0	99.48
20	Nonhuman	2.6	1.64	Undefined	23	0.44	36.79	10	87.80
21	Nonhuman	1.1	1.16	Undefined	6	0.40	37.48	0	95.35
22	Chemically Treated	6.1	10.92	Undefined	6	0.41	36.40	0	98.48
23A	Mixture 1A:1B	15,204	100	Undefined	44	0.38	36.02	31	77.76
23B	Mixture 1A:9B	12,156	100	U1b	38	0.41	35.97	0	95.76
23C	Mixture 9A:1B	12,792.8	100	Undefined	44	0.38	35.99	31	78.95
RB-6	Reagent Blank	6.4	6.1	T2	9	0.31	36.80	1	95.37
RB-7	Reagent Blank	8.94	29.24	V10	9	0.38	38.05	1	95.34
PC	Average all	2,877.9	100	T2a1b1a1a	43	0.36	37.1	3	99.11

processed sample was no longer above LOD when sequenced without the PC, indicating that the contamination observed in the first run was due to sequencing crosstalk. One of the three samples (16A1b) believed to be truly contaminated with a co-processed sample was also re-sequenced, and the profile was still mixed when the PC was removed. This finding of the persistent contaminating sequence supports the plausibility of 16A being mixed at the library level and not a result of sequencing crosstalk. Thus, the contamination indication in RBs and NCs was not indicative of contamination in a sample (and vice versa).

Overall, the rate of contamination decreased when the PC was removed from the sequencing runs, as it contributed approximately 1–2X coverage to the samples and control blanks. (This estimate was based upon the data from the pools resequenced without the PC, which were compared back to the original data). These results therefore indicate that sequencing crosstalk may have caused the mis-assignment of sequence data to an incorrect index. This can happen via “spatial crosstalk” among neighboring DNA clusters, which is a known issue with Illumina sequencing technology [33,34]. The crosstalk observed in the present study was caused by high-quality samples (including PCs) in which a high proportion of reads mapped to the rCRS. Yet it may be just

Table 5
NGS-Sanger data concordance of sample profiles above the analytical threshold.

Sample Type	Library Count	Concordant with STS Data	Mixed with Positive Control	Mixed with Associated Sample	Mixed with Unknown Source
Positive Control	17	100%			
High Quality	2	100%			
Degraded	18	89%		11%*	
Chemically Treated	13	84%	8%		8%
Reagent Blank	2		50%	50%	

*These two libraries were generated from the same DNA extract; therefore it is possible that only one contamination event occurred.

Table 6
Crosstalk observed in samples sequenced with and without the PC.

		Sequenced With PC		Sequenced Without PC	
Sample Type	Library Count	# Mixed Profiles Above LOD	Observations of PC in rCRS read mapping	# Mixed Profiles Above LOD	Observations of PC in rCRS read mapping
High Quality	1	0	0	0	0
Degraded	6	1	1	1	0
Chemically Treated	8	1	5	0	0
Reagent Blank	12	2	6	0	0
Negative Control	7	0	1	0	0

as likely that chemically treated and degraded DNA samples produced sequence data that were mis-assigned to a neighboring cluster on the flow cell. Due to the fact that the majority of reads (averaging ~87%) generated from degraded and chemically treated samples did not map to the rCRS, the crosstalk from one of these samples would not be recognized in the mitogenome data because the majority of reads were off-target. Consequently, the source of the observed crosstalk was more likely to be high-quality samples than samples with low endogenous mtDNA content. Therefore it is paramount that samples of disparate quality – in particular the endogenous mtDNA content – are sequenced separately.

The removal of the PC was of little consequence to the success of the sequencing runs and analytical thresholds. In fact, when the side-by-side data were compared from the runs with and without the PC, the sequence analysis metrics were very similar (Table S8). The number of reads passing filter, total reads identified per sample, and average coverage per sample were not statistically significantly different between runs ($p > 0.3$). Moreover, when only the RBs and NCs were considered (used to calculate LOD and LOQ), average coverage was not statistically significantly different between the sequencing runs with and without the PC ($p = 0.145$).

Until Illumina sequencing technology is adapted to mitigate spatial crosstalk [34], best scientific practice for this particular procedure is to forego the use of a traditional positive control. Instead, other means of determining whether a reaction was successful can be utilized as QC measures. These include quantitation of DNA after each step in the NGS workflow, checking for adapter dimer in control blanks to confirm that library preparation was performed properly, replication of NGS data through two independent DNA extractions and library preparation events to verify the sequence result, and the use of Phi-X as a sequencing control to monitor the sequencing reaction. And given that all of the DNA extract is consumed during library preparation, a failed library will always result in starting over from DNA extraction – regardless of the PC result. The need for the traditional positive control, a control DNA that is initiated at the amplification step, can be satisfied by the adoption of these alternative, yet robust, QC procedures.

4. Conclusion

The present study evaluated a NGS mitogenome capture protocol for sequencing mtDNA from non-probative case-type samples typically processed by the AFMES-AFDIL. The results showed NGS methods to be exceedingly sensitive, capable of generating entire mitogenome data from samples that failed to yield reliable sequences with standard PCR-based techniques. All degraded and high-quality samples generated complete mitogenome profiles with the exception of one sample that exceeded the maximum library DNA input. The workflow produced reliable, reproducible mitogenome profiles that were concordant with Sanger data (when available) as well as consistent between NGS replicates. The data herein demonstrate the cleanliness of the reagents involved in the laboratory workflow. Yet the results showed that low-level crosstalk can be expected during the sequencing run, especially from high quality samples and positive controls, which may lead to the appearance of contamination. The comparison of sample data from sequencing runs performed with and without the positive control suggest that the inclusion of a positive control in this method interfered with data analysis of the low-quality samples that necessitated the capture procedure. Thus, it is recommended for this procedure that degraded DNA samples are not sequenced in conjunction with high-quality samples or controls. Although the FBI's Quality Assurance Standards requires the use of a positive control, it was shown here to be detrimental to the interpretation of sample and reagent blank data. Instead of a traditional positive control, alternative quality control procedures may be utilized to ensure the authenticity of the obtained data. This will reduce unnecessary costly reprocessing of seemingly contaminated samples that do not contain exogenous mtDNA. These findings underscore the fact that quality assurance measures must be specific to the DNA testing procedure at hand, thus forcing a paradigm shift in forensic DNA analysis.

Disclaimer

The assertions herein are those of the authors and do not necessarily represent the official position of the U.S. Department of Defense or its entities including the Armed Forces Medical Examiner System.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

None.

Acknowledgments

The authors would like to thank two anonymous reviewers for their helpful feedback on this report. The authors would furthermore like to thank the following individuals for their assistance in the project: Mark J. Wadhams, Taylor M. Dickerson III, Julie Demarest, Michelle A. Peck, Joseph D. Ring, Erin M. Gorden, Spence A. Fast, Odile P.M. Loreille, Jacqueline Tyler Thomas, Bruche E. Trotter, Christina M. Neal, Holly M. Krantz, Kerriann K. Meyers, Daniel J. Kahline, Jennifer A.G. O'Rourke, Megan V. Poffenberger, Lt Col Alice Briones (USAF), Lt Col Laura Garner (USAF), Timmathy Cambridge, James J. Canik, James Ross, Jon Norris, Michael Meyers, Michael A. Fasano, Lorenzo Andreaggi, Shairose Lalani, and AFMES leadership and support staff.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2017.09.001.

References

[1] Y. Cui, J. Lindo, C.E. Hughes, J.W. Johnson, A.G. Hernandez, B.M. Kemp, et al., Ancient DNA analysis of mid-Holocene individuals from the Northwest Coast of North America reveals different evolutionary paths for mitogenomes, PLoS One 8 (2013) e66948.

[2] J.M. Enk, A.M. Devault, M. Kuch, Y.E. Murgha, J.M. Rouillard, H.N. Poinar, Ancient whole genome enrichment using baits built from modern DNA, Mol. Biol. Evol. 31 (2014) 1292–1294.

[3] M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing, Cold Spring Harb. Protoc. 2010 (2010) (pdb.prot5448).

[4] M.C. Avila-Arcos, E. Cappellini, J.A. Romero-Navarro, N. Wales, J.V. Moreno-Mayar, M. Rasmussen, et al., Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA, Sci. Rep. 1 (2011) 74.

[5] M.L. Carpenter, J.D. Buenrostro, C. Valdiosera, H. Schroeder, M.E. Allentoft, M. Sikora, et al., Pulling out the 1%: whole-Genome capture for the targeted enrichment of ancient DNA sequencing libraries, Am. J. Hum. Genet. 93 (2013) 1–13.

[6] J. Lindo, E. Huerta-Sánchez, S. Nakagome, M. Rasmussen, B. Petzelt, J. Mitchell, et al., A time transect of exomes from a Native American population before and after European contact, Nat. Commun. 7 (2016) 13175.

[7] J.E. Templeton, P.M. Brotherton, B. Llamas, J. Soubrier, W. Haak, A. Cooper, et al., DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification, Invest. Genet. 4 (2013) 1–13.

[8] M.A. Peck, M.D. Brandhagen, C. Marshall, T.M. Diegoli, J.A. Irwin, K. Sturk-Andreaggi, Concordance and reproducibility of a next generation mtGenome sequencing method for high-quality samples using the Illumina MiSeq, Forensic Sci. Int. Genet. 24 (2016) 103–111.

[9] M.M. Rathbun, J.A. McElhoe, W. Parson, M.M. Holland, Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data, Forensic Sci. Int. Genet. 26 (2017) 1–11.

[10] W. Parson, C. Strobl, G. Huber, B. Zimmermann, S.M. Gomes, L. Souto, et al., Evaluation of next generation mtGenome sequencing using the ion torrent personal genome machine (PGM), Forensic Sci. Int. Genet. 7 (2013) 543–549.

[11] S. Desmyter, M. Bodner, G. Huber, S. Dognaux, C. Berger, F. Noël, et al., Hairy matters: mtDNA quantity and sequence variation along and among human head hairs, Forensic Sci. Int. Genet. 25 (2016) 1–9.

[12] W. Parson, G. Huber, L. Moreno, M. Madel, M.D. Brandhagen, S. Nagl, et al., Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples, Forensic Sci. Int. Genet. 15 (2015) 8–15.

[13] S.M. Edson, J.P. Ross, M.D. Coble, T.J. Parsons, S.M. Barritt, Naming the dead confronting the realities of rapid identification of degraded skeletal remains, Forensic Sci. Rev. 16 (2004) 63–90.

[14] M.N. Gabriel, E.F. Huffine, J.H. Ryan, M.M. Holland, T.J. Parsons, Improved MtDNA sequence analysis of forensic remains using a mini-primer set amplification strategy, J. Forensic Sci. 46 (2001) 247–253.

[15] T. Lindahl, Instability and decay of the primary structure of DNA, Nature 362 (1993) 709–715.

[16] B. Balitzki-Korte, K. Anslinger, C. Bartsch, B. Rolf, Species identification by means of pyrosequencing the mitochondrial 12S rRNA gene, Int. J. Legal Med. 119 (2005) 291–294.

[17] The FBI Director's Forensic Quality Assurance Standards for DNA Testing Laboratories, (2011).

[18] O.M. Loreille, T.M. Diegoli, J.A. Irwin, M.D. Coble, T.J. Parsons, High efficiency DNA extraction from bone by total demineralization, Forensic Sci. Int. Genet. 1 (2007) 191–195.

[19] O.M. Loreille, R.L. Parr, K.A. McGregor, C.M. Fitzpatrick, C. Lyon, D.Y. Yang, et al., Integrated DNA and fingerprint analyses in the identification of 60-year-old mummified human remains discovered in an Alaskan glacier, J. Forensic Sci. 55 (2010) 813–818.

[20] A.W. Briggs, U. Stenzel, M. Meyer, J. Krause, M. Kircher, S. Paabo, Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA, Nucleic Acids Res. 38 (2010) e87.

[21] D. Reich, R.E. Green, M. Kircher, J. Krause, N. Patterson, E.Y. Durand, et al., Genetic history of an archaic hominin group from Denisova Cave in Siberia, Nature 468 (2010) 1053–1060.

[22] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, et al., Sequence and organization of the human mitochondrial genome, Nature 290 (1981) 457–465.

[23] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (1999) 147.

[24] K. Sturk-Andreaggi, M.A. Peck, C. Boysen, P. Dekker, T.P. McMahon, C. Marshall, AQME: A Forensic Mitochondrial DNA Analysis Tool for Next Generation Sequencing Data, (2017).

[25] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (2009) E386–394.

[26] C.N. Stephan, A.P. Winburn, A.F. Christensen, A.J. Tyrrell, Skeletal identification by radiographic comparison: blind tests of a morphoscopic method using antemortem chest radiographs, J. Forensic Sci. 56 (2011) 320–332.

[27] C.N. Stephan, B. Amidan, H. Trease, P. Guyomarc'h, T. Pulsipher, J.E. Byrd, Morphometric comparison of clavicle outlines from 3D bone scans and 2D chest radiographs: a shortlisting tool to assist radiographic identification of human skeletons, J. Forensic Sci. 59 (2014) 306–313.

[28] D.I. Cruz-Dávalos, B. Llamas, C. Gaunitz, A. Fages, C. Gamba, J. Soubrier, et al., Experimental conditions improving in-solution target enrichment for ancient DNA, Mol. Ecol. Res. 17 (3) (2016) 508–522.

[29] M.C. Ávila-Arcos, M. Sandoval-Velasco, H. Schroeder, M.L. Carpenter, A. Malaspinas, N. Wales, et al., Comparative performance of two whole-genome capture methodologies on ancient DNA Illumina libraries, Methods Ecol. Evol. 6 (2015) 725–734.

[30] P. Korlević, T. Gerber, M. Gansauge, M. Hajdinjak, S. Nagel, A. Ayinuer-Petri, et al., Reducing microbial and human contamination in DNA extractions from ancient bones and teeth, Biotechniques 59 (2015) 87–93.

[31] O. Loreille, H. Koshinsky, V.Y. Fofanov, J.A. Irwin, Application of next generation sequencing technologies to the identification of highly degraded unknown soldiers' remains, Forensic Sci. Int. Genet. Suppl. Ser. 3 (2011) e540–e541.

[32] R.S. Just, M.K. Scheible, S.A. Fast, K. Sturk-Andreaggi, A.W. Röck, J.M. Bush, et al., Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three US populations, Forensic Sci. Int. Genet. 14 (2015) 141–155.

[33] J.A. Sleep, A.W. Schreiber, U. Baumann, Sequencing error correction without a reference genome, BMC Bioinf. 14 (2013) 367.

[34] B. Wang, L. Wan, A. Wang, L.M. Li, An adaptive decorrelation method removes Illumina DNA base-calling errors caused by crosstalk between adjacent clusters, Sci. Rep. 7 (2017) 41348.

Forensic Science International: Genetics 31 (2017) 189–197

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Short communication

AQME: A forensic mitochondrial DNA analysis tool for next-generation sequencing data

Kimberly Sturk-Andreaggi[a,b,*], Michelle A. Peck[a,b], Cecilie Boysen[c,1], Patrick Dekker[c,2], Timothy P. McMahon[a], Charla K. Marshall[a,b]

[a] Armed Forces DNA Identification Laboratory, A Division of the Armed Forces Medical Examiner System, 115 Purple Heart Drive, Dover AFB, DE 19902, United States
[b] ARP Sciences, LLC, Contractor Supporting the Armed Forces Medical Examiner System, 9210 Corporate Boulevard, Suite 120, Rockville, MD 20850, United States
[c] QIAGEN Bioinformatics, Silkeborgvej 2, 8000 Aarhus C, Denmark

ARTICLE INFO

Keywords:
Mitochondrial DNA
Next-generation sequencing (NGS)
Forensics
Data analysis
Haplogroup

ABSTRACT

The feasibility of generating mitochondrial DNA (mtDNA) data has expanded considerably with the advent of next-generation sequencing (NGS), specifically in the generation of entire mtDNA genome (mitogenome) sequences. However, the analysis of these data has emerged as the greatest challenge to implementation in forensics. To address this need, a custom toolkit for use in the CLC Genomics Workbench (QIAGEN, Hilden, Germany) was developed through a collaborative effort between the Armed Forces Medical Examiner System – Armed Forces DNA Identification Laboratory (AFMES-AFDIL) and QIAGEN Bioinformatics. The AFDIL-QIAGEN mtDNA Expert, or AQME, generates an editable mtDNA profile that employs forensic conventions and includes the interpretation range required for mtDNA data reporting. AQME also integrates an mtDNA haplogroup estimate into the analysis workflow, which provides the analyst with phylogenetic nomenclature guidance and a profile quality check without the use of an external tool. Supplemental AQME outputs such as nucleotide-per-position metrics, configurable export files, and an audit trail are produced to assist the analyst during review. AQME is applied to standard CLC outputs and thus can be incorporated into any mtDNA bioinformatics pipeline within CLC regardless of sample type, library preparation or NGS platform.

An evaluation of AQME was performed to demonstrate its functionality and reliability for the analysis of mitogenome NGS data. The study analyzed Illumina mitogenome data from 21 samples (including associated controls) of varying quality and sample preparations with the AQME toolkit. A total of 211 tool edits were automatically applied to 130 of the 698 total variants reported in an effort to adhere to forensic nomenclature. Although additional manual edits were required for three samples, supplemental tools such as mtDNA haplogroup estimation assisted in identifying and guiding these necessary modifications to the AQME-generated profile. Along with profile generation, AQME reported accurate haplogroups for 18 of the 19 samples analyzed. The single errant haplogroup assignment, although phylogenetically close, identified a bug that only affects partial mitogenome data. Future adjustments to AQME's haplogrouping tool will address this bug as well as enhance the overall scoring strategy to better refine and automate haplogroup assignments. As NGS enables broader use of the mtDNA locus in forensics, the availability of AQME and other forensic-focused mtDNA analysis tools will ease the transition and further support mitogenome analysis within routine casework. Toward this end, the AFMES-AFDIL has utilized the AQME toolbox in conjunction with the CLC Genomics Workbench to successfully validate and implement two NGS mitogenome methods.

1. Introduction

The use of next-generation sequencing (NGS) for mitochondrial DNA (mtDNA) analysis has been thoroughly evaluated by the forensics community as a high-throughput, sensitive method compared to traditional Sanger sequencing [1–5]. Although similar to Sanger, the analysis of the comparatively large amount of data produced with NGS technologies has emerged as the greatest challenge to implementation, necessitating the development of forensic-specific bioinformatics tools. Until recently, NGS software development was driven by other fields

* Corresponding author at: Armed Forces DNA Identification Laboratory, 115 Purple Heart Dr Dover AFB, DE, 19902, United States.
E-mail address: kimberly.s.andreaggi.ctr@mail.mil (K. Sturk-Andreaggi).
[1] Present address: PACT Pharma, 3928 Point Eden Way, Hayward, CA 94545, United States.
[2] Present address: KeyGene N.V., Agro Business Park 90, 6708 PW Wageningen, The Netherlands.

http://dx.doi.org/10.1016/j.fsigen.2017.09.010
Received 25 May 2017; Received in revised form 21 August 2017; Accepted 16 September 2017
Available online 19 September 2017
1872-4973/

such as evolutionary genetics and biomedical research, which focus on genome assembly and variant discovery versus haplotype comparison that is the thrust of forensic mtDNA analysis.

Consequently, forensic mtDNA analysts must convert a variant list generated by the traditional NGS software into a profile for forensic use [6–8]. The forensic profile requires 3′ placement of indels, the separation of grouped adjacent variants (i.e. multiple nucleotide variants, MNVs), the assignment of IUPAC characters for heteroplasmic positions (i.e. point heteroplasmies, PHPs), and accordance with forensic nomenclature including calling of the major molecule (MM) in regions of length heteroplasmy (LHP). Without specialized forensic tools for mtDNA analysis, these profile manipulations have to be performed manually and are therefore susceptible to errors in addition to being time consuming.

In recent years, forensic scientists have begun to develop bioinformatics tools for mtDNA profile generation. For example, mitoSAVE [9] is an Excel-based solution designed as an ad hoc nomenclature modifier to be used after a bioinformatics workflow has produced a VCF file. MitoSAVE not only generates a haplotype consistent with current forensic nomenclature guidelines, but it furthermore applies user-defined thresholds for profile reporting (e.g., coverage, heteroplasmy). More recently, GeneMarker HTS (SoftGenetics, Inc., State College, PA) was introduced for mtDNA analysis within a forensic context [10]. This program offers a user-friendly interface to analyze FASTQ data and generate mtDNA profiles adhering to forensic nomenclature. Such forensically developed tools have significantly helped to make the analysis of mtDNA NGS data generated from high-quality samples straightforward.

However, the remaining challenge in performing mtDNA analysis of NGS data is the ability to determine the breadth of the mtDNA reference genome that meets the coverage threshold for variant detection (i.e. the interpretation range). This necessity for the reporting of an interpreted range is particularly relevant to low-quality samples that often contain coverage drop-out. Furthermore, NGS chemistries that result in mtDNA genome (mitogenome) coverage imbalance regardless of sample quality will require interpreted range designation (e.g., [11,12]). With Sanger data, analysts visually confirm the presence of the requisite number (and direction) of sequences in order to generate the corresponding interpretation range for the mtDNA profile. While seemingly straightforward, sequence pileup and read filtering parameters employed for variant detection make this task extremely difficult for NGS data and certainly impractical when analyzing the 16,569-bp mitogenome and/or hundreds of samples per worklist. As a result of the lacking interpreted range from analyzed NGS data, a polymorphism may appear absent from the mtDNA profile when in fact the position is below the coverage threshold and therefore not subject to variant detection [1]. Consequently an incorrect profile may be reported if the interpretation range is assumed as the entire target region instead of accurately reflecting only those positions sufficiently covered.

A second need for NGS mtDNA data analysis is a software tool that will produce a haplogroup estimate during profile generation. The haplogroup information assists the analyst in adhering to phylogenetic nomenclature according to international guidelines [6,8]. In the absence of integrated haplogroup assignment, it is necessary to generate the mtDNA profile and/or consensus sequence and then enter that output into external haplogrouping tools (e.g., [13–18]). The ability to determine the mtDNA haplogroup without user intervention during NGS data analysis provides the analyst with accurate, real-time phylogenetic nomenclature guidance. The other major advantage of estimating the haplogroup during the initial profile review is that it can act as a critical quality assurance check of the validity of a haplotype by flagging artificial recombination [19], phantom mutations [20], and other potential errors in the data. Furthermore, the haplogroup can be used to predict maternal ancestry that may provide valuable information to an investigation.

Though forensic mtDNA analysis tools are now available, none provide all of the desired functionality within a single software solution. To this end, the Armed Forces Medical Examiner System — Armed Forces DNA Identification Laboratory (AFMES-AFDIL) with QIAGEN Bioinformatics created a custom tool for forensic analysis of mtDNA data within the CLC Genomics Workbench (QIAGEN, Hilden, Germany). The AFDIL-QIAGEN mtDNA Expert, or AQME, is a collection of tools designed to be employed at the end of any mtDNA analysis workflow in CLC to generate a profile according to user-defined parameters and forensic conventions. AQME has been designed for the analysis of mtDNA data from all sample types regardless of the sample quality, targeted range (e.g., mtDNA control region (CR), mitogenome), library preparation method or sequencing platform. To demonstrate the functionality of the AQME tools, mitogenome NGS data from both high- and low-quality samples were analyzed. The outputs were evaluated to assess the reliability of AQME in profile generation and haplogroup assignment as well as the gains in efficiency observed for mitogenome NGS data analysis.

2. Materials and methods

2.1. Forensic mtDNA analysis tool

AQME requires two CLC input files: a read mapping (sequence reads aligned to the mtDNA reference genome) and the associated variant track (list of detected mtDNA variants). This allows AQME to be used on any mtDNA data independent of sample type, library preparation, and NGS platform. The AQME toolbox provides the CLC user with three primary tools for forensic mtDNA analysis: Realign Variants, Create Mitochondrial Variant Table and Mitochondrial Haplogrouper (Fig. 1). Supplemental tools also included in the AQME toolbox allow for easy import of files utilized by the three main tools, creation of a consensus sequence from the mtDNA profile, and analysis of read counts in specified regions (e.g., regions of LHP). The AQME tools are employed following defined steps for variant calling such as sequence trimming, mapping to the targeted portion of the revised Cambridge Reference Sequence (rCRS) [21–23], and other read handling steps specific to the data (Fig. S1). Because the CLC mapping tool aligns indels 5′ in homopolymeric tracts [24], the conventional CLC variant track lists indels incorrectly according to forensic guidelines [6,8]. To address this issue, the Realign Variants tool virtually shifts indels 3′ within the homopolymeric region (Fig. S2). Variants are shifted only in the variant track but not in the mapping; however both the original and shifted position are documented in the Realigned Variant Track (and thus included in the mtDNA Table) to maintain an auditable edit trail.

The Realigned Variant Track and Read Mapping outputs are input into the Create Mitochondrial Variant Table tool. Multiple outputs allow the user access to detailed information about the mtDNA data (Fig. 1). The main output, the mtDNA Table (Figs. 2 and S3), includes both original and final variant calls with edited positions flagged by color as well as tool comments identifying modifications performed automatically. These tool-edited positions include PHPs that are automatically assigned the appropriate IUPAC character, variants shifted by the Realign Variants tool, and variants previously categorized as MNVs that were separated into single nucleotide polymorphisms. Additionally, the ability to filter entire regions from the profile or certain types of variants (e.g., substitution or indel) based on their frequency has been incorporated into the Create Mitochondrial Variant Table tool. These modifications are also displayed and flagged in the mtDNA Table. The mtDNA Table displays twelve metrics for each variant position in order to allow the analyst to comprehensively review the profile (Table S1).

Another output of the Create Mitochondrial Variant Table tool, the Coverage Table, provides six metrics for each nucleotide at every position in the mtDNA target region (i.e. rCRS) based upon the variant detection parameters (Table S1). These metrics provide forensic DNA analysts, who are accustomed to visually inspecting each base in

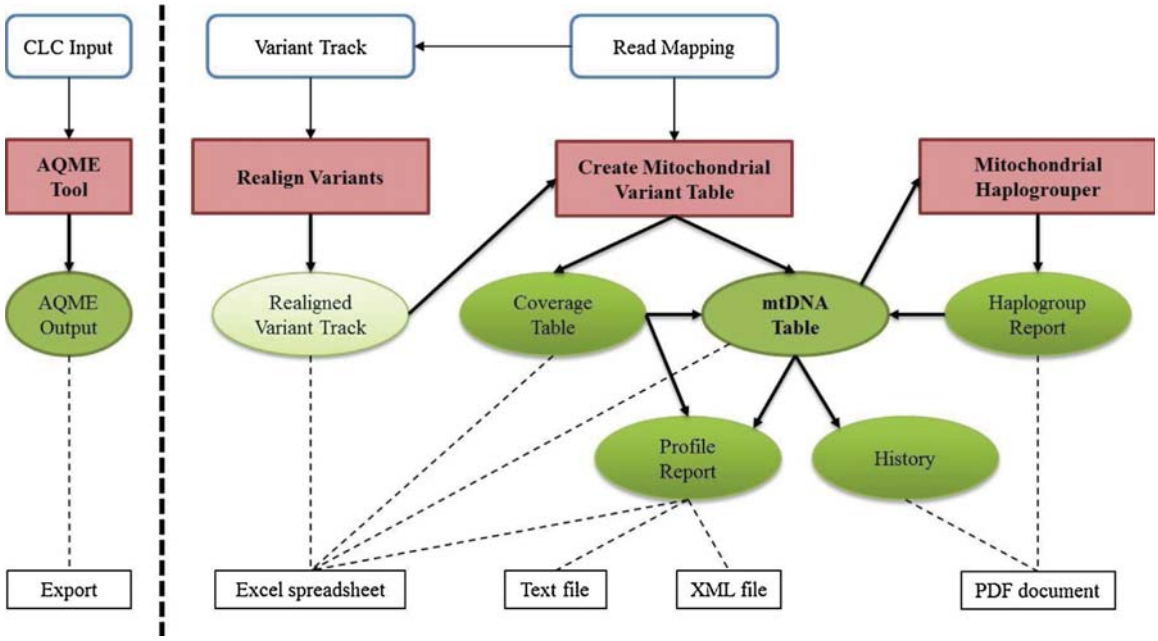


Fig. 1. Flowchart of the AQME tools, inputs, outputs and exports.

Sanger-generated data, with the same level of scrutiny not previously possible due to the sheer volume of NGS data. The quantitative data now accessible to the analyst offers a means to validate the variants reported as well as all non-variant positions. This information can be used to assess the background noise of NGS data, which is useful for determining variant detection thresholds. AQME utilizes the information in the Coverage Table to determine the interpretation range, and allows the analyst to exclude specific regions from the range as necessary. The analyst can edit, delete, or add variants based upon their review of the data provided in these two outputs and the read mapping, which can be displayed alongside the mtDNA Table (Fig. S4). Any edit requires entry of the user's initials and a comment, and edit details are recorded in the History (Fig. S5). The History report also includes detailed information about each step in the automated analysis workflow that generated the mtDNA Table, and can then be exported as a PDF file for inclusion of the profile audit trail in the case file.

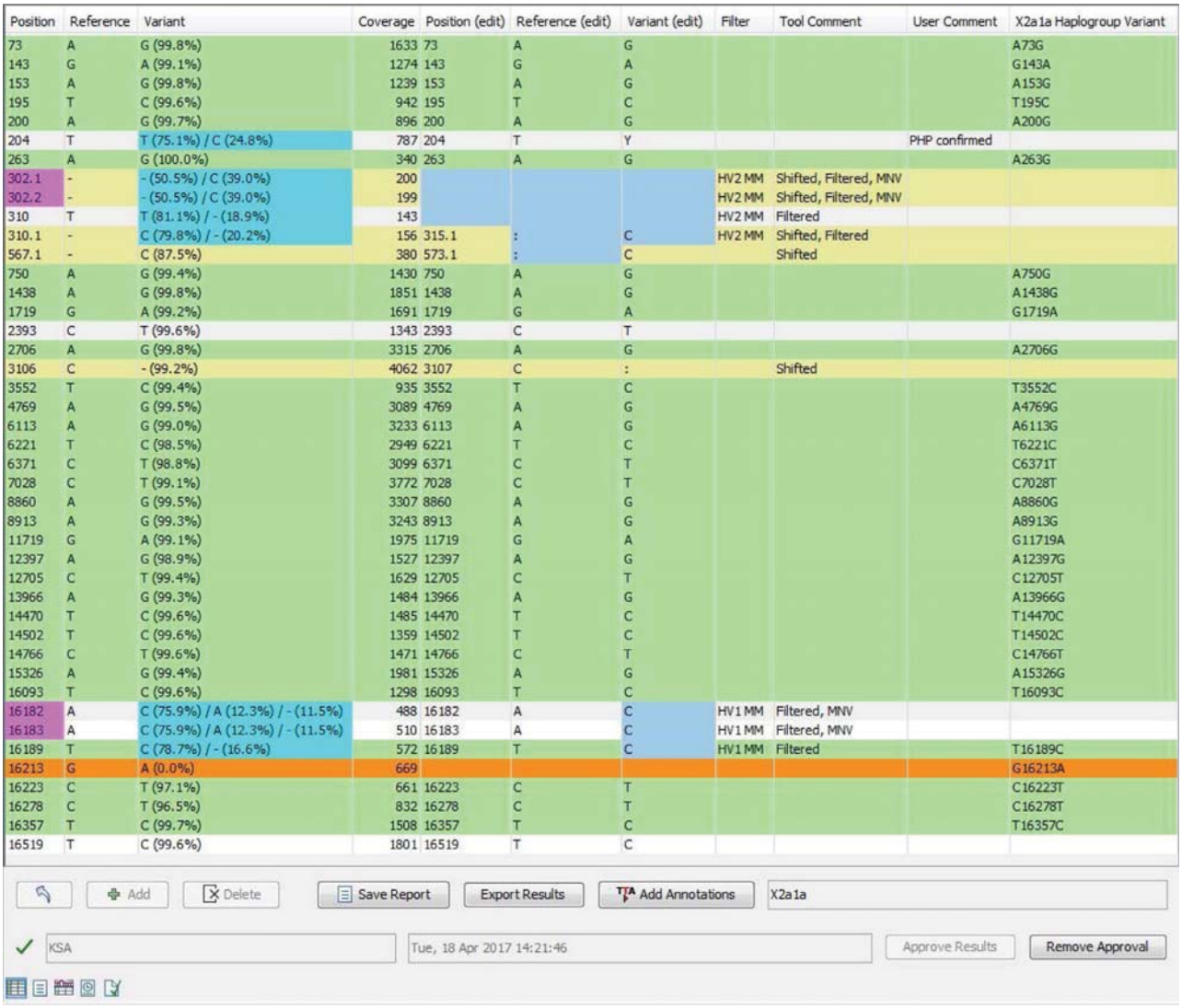
The Mitochondrial Haplogrouper tool provides an integrated haplogroup estimation of the generated consensus sequence. To utilize the Mitochondrial Haplogrouper, a variant list and virtual mtDNA sequences are imported into CLC using a supplemental AQME tool to create an annotated haplogroup database for comparison. This database can be based upon the haplogroups in the Phylotree mtDNA tree [25] and updated as new mtDNA phylogenetic trees are released, or even utilized for other purposes with alternative haplotype sources (e.g., staff elimination database). The Haplogrouper tool performs the assignment by first generating a consensus sequence from the interpretation range of the profile in the mtDNA Table, and then aligning that sequence to all of the reference haplogroup sequences in the database. The tool returns the haplogroup with the greatest alignment score. When multiple haplogroups share the top alignment score, all are listed in the report. The number of reported (i.e., top-scoring) haplogroups is typically one for full mitogenome data and multiple if only a portion of the mitogenome is covered. From those select sequences, the haplogroup-defining SNPs are compared to the sample variants and the haplogroups are ranked based on user-defined scoring. For this study the following scoring strategy was used: +10 for a shared SNP, −8 for a missing SNP, −5 for a missing unstable/uncertain SNP, and −3 for a private SNP. The current version of AQME considers phylogenetically neutral variants (e.g., T16519C, 315.1C, 523–524 AC indels) as private mutations even though they are excluded from phylogenetic reconstruction by Phylotree [25]. This two-step process assigns the haplogroup first by string search and then variant profile comparison; therefore nomenclature can be informed by the phylogeny without impacting accurate mtDNA haplogroup estimation. Once the haplogroup assignment is complete, a Haplogroup Report is generated (Fig. S6). The mtDNA Table is automatically updated with the haplogroup-defining variants if the report contains only one haplogroup (Figs. 2 and S3); conversely, the analyst must select the haplogroup to add to the mtDNA Table if multiple lineages are identified by the Mitochondrial Haplogrouper tool.

An analyst then approves the mtDNA Table once the generated profile is deemed complete, allowing for the Profile Report to be exported. This report includes the sample name, profile approval information, the software and plug-in versions, the coverage threshold used, variant profile, interpretation range, and haplogroup (if determined). The haplogroup represented in the mtDNA Table will be reflected in the Profile Report; however, "Undefined" will be reported if the analyst did not select one haplogroup when multiple were identified by AQME or "N/A" if no haplogroup assignment was performed. The Profile Report can be exported as a text file (Fig. S7) and/or XML file, which is in the Combined DNA Index System (CODIS) mtDNA format for electronic transfer into a compatible database. Moreover, all components of the mtDNA Table can be exported through the workflow as individual files or separate tabs in a single Excel workbook.

The AQME toolkit can be used on a stand-alone CLC Genomics Workbench or on a high-capacity server to increase efficiency. Furthermore, the incorporation of the AQME tools into an automated workflow enables an administrator to lock all parameters to ensure that only the validated pipeline is utilized to analyze forensic data.

2.2. Data collection and analysis

CLC read mappings and variant tracks from 12 high-quality reference-type samples and five low-quality non-probative case samples were analyzed with the AQME tools to generate forensic mtDNA profiles and haplogroups. Associated controls were also included in the analyses. Sanger sequencing data were available for the entire mitogenome of the reference-type samples [26]. MtDNA CR data were previously generated with Sanger sequencing from the case samples and/or related family reference sample. NGS data were generated on



Position	Reference	Variant	Coverage	Position (edit)	Reference (edit)	Variant (edit)	Filter	Tool Comment	User Comment	X2a1a Haplogroup Variant
73	A	G (99.8%)	1633	73	A	G				A73G
143	G	A (99.1%)	1274	143	G	A				G143A
153	A	G (99.8%)	1239	153	A	G				A153G
195	T	C (99.6%)	942	195	T	C				T195C
200	A	G (99.7%)	896	200	A	G				A200G
204	T	T (75.1%) / C (24.8%)	787	204	T	Y			PHP confirmed	
263	A	G (100.0%)	340	263	A	G				A263G
302.1	-	- (50.5%) / C (39.0%)	200				HV2 MM	Shifted, Filtered, MNV		
302.2	-	- (50.5%) / C (39.0%)	199				HV2 MM	Shifted, Filtered, MNV		
310	T	T (81.1%) / - (18.9%)	143				HV2 MM	Filtered		
310.1	-	C (79.8%) / - (20.2%)	156	315.1	:	C	HV2 MM	Shifted, Filtered		
567.1	-	C (87.5%)	380	573.1	:	C		Shifted		
750	A	G (99.4%)	1430	750	A	G				A750G
1438	A	G (99.8%)	1851	1438	A	G				A1438G
1719	G	A (99.2%)	1691	1719	G	A				G1719A
2393	C	T (99.6%)	1343	2393	C	T				
2706	A	G (99.8%)	3315	2706	A	G				A2706G
3106	C	- (99.2%)	4062	3107	C	:		Shifted		
3552	T	C (99.4%)	935	3552	T	C				T3552C
4769	A	G (99.5%)	3089	4769	A	G				A4769G
6113	A	G (99.0%)	3233	6113	A	G				A6113G
6221	T	C (98.5%)	2949	6221	T	C				T6221C
6371	C	T (98.8%)	3099	6371	C	T				C6371T
7028	C	T (99.1%)	3772	7028	C	T				C7028T
8860	A	G (99.5%)	3307	8860	A	G				A8860G
8913	A	G (99.3%)	3243	8913	A	G				A8913G
11719	G	A (99.1%)	1975	11719	G	A				G11719A
12397	A	G (98.9%)	1527	12397	A	G				A12397G
12705	C	T (99.4%)	1629	12705	C	T				C12705T
13966	A	G (99.3%)	1484	13966	A	G				A13966G
14470	T	C (99.6%)	1485	14470	T	C				T14470C
14502	T	C (99.6%)	1359	14502	T	C				T14502C
14766	C	T (99.6%)	1471	14766	C	T				C14766T
15326	A	G (99.4%)	1981	15326	A	G				A15326G
16093	T	C (99.6%)	1298	16093	T	C				T16093C
16182	A	C (75.9%) / A (12.3%) / - (11.5%)	488	16182	A	C	HV1 MM	Filtered, MNV		
16183	A	C (75.9%) / A (12.3%) / - (11.5%)	510	16183	A	C	HV1 MM	Filtered, MNV		
16189	T	C (78.7%) / - (16.6%)	572	16189	T	C	HV1 MM	Filtered		T16189C
16213	G	A (0.0%)	669							G16213A
16223	C	T (97.1%)	661	16223	C	T				C16223T
16278	C	T (96.5%)	832	16278	C	T				C16278T
16357	T	C (99.7%)	1508	16357	T	C				T16357C
16519	T	C (99.6%)	1801	16519	T	C				

Fig. 2. The mtDNA Table for sample mtGAfrWI0001 with 11 of the 33 optional columns displayed. Tool modifications are flagged with highlighting and/or comments such as shifted positions (yellow row, tool comment), heteroplasmic sites (cyan variant, IUPAC edited variant), MNVs (magenta position, tool comment), filtered LHP regions (filter note, tool comment), and changes to positions or variants (blue edited position, reference and/or variant). One user comment is shown at np 204 confirming the presence of a PHP; no other user edits were necessary. The variants diagnostic for haplogroup X2a1a are also shown in the mtDNA Table, and the assigned haplogroup is indicated at the bottom of the profile as well as the top of the haplogroup variant column. Rows highlighted in green indicate mutations observed in the sample profile and assigned haplogroup with unhighlighted/yellow rows denoting private or ignored mutations. The orange row is an example of a haplogroup-defining variant not present in the sample and the details for that variant are displayed. The profile was approved by analyst "KSA" on 18 April 2017. See Fig. S3 for a detailed view of this table.

the MiSeq System (Illumina, San Diego, CA) using procedures described in Peck et al. for the high-quality samples [1] and Marshall et al. for the low-quality samples [27].

The analysis workflows utilized to generate the read mappings and variant tracks in the CLC Workbench were based upon the specific needs of each library preparation and sample type. Sample information, processing methods, and data analysis parameters are summarized in Table 1 for reference. Following read mapping and variant detection in each workflow, the Realign Variants tool was used to virtually shift variants 3′ according to forensic convention. Then, the Create Mitochondrial Variant Table tool converted the Realigned Variant Track based on the variant detection parameters utilized. The mtDNA tool also applied a filter in regions of known LHP (nps 302–316, 452–464, 513–525, 568–574, 956–966, 5891–5900, 8270–8290, 12418–12426,

Table 1
Information about the two sample sets including sample type, processing method and analysis details. Each sample set also included one positive and one negative control, which are not represented in the count.

Sample Set	Count	Substrate	DNA Quality	Mitogenome Enrichment	Library Preparation	Analysis Workflow	Variant Detection
Reference-Type	12	Serum	High	Long-Range PCR	Nextera XT	Sequences trimmed; Reads mapped to rCRS (default); Local realignment	100X minimum coverage; 10 minimum variant count, 10% variant frequency
Non-Probative	5	Degraded Bone	Low	Hybridization Capture	NEBNext Ultra	Reads mapped to rCRS (stringent); Mapped duplicate reads removed; Local realignment	10X minimum coverage; 4 minimum variant count, 10% variant frequency

16180–16194) in order to automatically report the major molecule (variants with ≥50% frequency). Each workflow incorporated the Mitochondrial Haplogrouper to estimate the mtDNA haplogroup based on the phylogenetic haplotypes presented in Phylotree.org mtDNA tree Build 16 [25]. AQME-generated haplogroup assignments were compared to haplogroup estimations generated by a new tool developed as part of the EDNAP Mitochondrial DNA Population Database (EMPOP [28]; W. Parson, personal communication) based upon Build 17 of Phylotree [29]. Haplogroups were manually confirmed using Phylotree.

3. Results and discussion

3.1. Analysis time

Each workflow was completed in roughly two and a half minutes when a FASTQ file of 60–500 MB was analyzed using the Genomics Server extension of the CLC tools on a high-capacity server with three compute nodes (each node includes 16 × 2.6 GHz processors, 64 GB RAM, 2 TB hard drive). The same analysis took about seven minutes using the CLC Genomics Workbench on a personal computer (2.53 GHz processor, 12 GB RAM, 465 GB hard drive). The generation of the read mapping and variant detection using the standard CLC tools required less than two minutes regardless of the system used, with the remaining time required for the AQME-specific steps. Though haplogroup assignment took only 10 s when performed on a high-capacity server, the process was significantly longer (approximately five minutes) on the Workbench due to the number of haplogroup reference sequences used for comparison (> 4800). Therefore the haplogrouping process was the most time-consuming step, but only when performed on the Genomics Workbench. If an experienced user performed all of the AQME-specific steps manually on the standard CLC outputs, analysis would take substantially more time (≥15 min/sample). In fact, manual generation of 96 mtDNA forensic profiles and haplogroups would likely take over 24 h, introducing greater opportunity for error. In comparison, batch analysis of 96 samples (~200,000 paired reads of ~225 bp per FASTQ) on the high-capacity server took just over an hour (68 min), equating to less than a minute per sample.

3.2. Interpretation range

The depth and breadth of coverage was reported in both the interpretation range as well as the variants detected for all samples (Tables S2-S3). Only one reference-type sample (mtGAfrOH0001) had coverage less than the 100X minimum coverage threshold for variant calling. The low coverage was observed across portions of the hypervariable region 2 (HV2), resulting in drop-out of the expected 315.1C insertion in this high-quality profile. However, the interpretation range accurately denoted that the position was not covered, and thus excluded from variant detection. Therefore even reference-type specimens may exhibit incomplete mitogenome coverage that necessitates interpretation range information. The three low-quality samples without full coverage resulted in varying breadths of mitogenome coverage ranging from 3.2% (Sample 5) to 99.6% (Sample 3). In these cases, the interpretation range assisted in the comparison of profiles by clarifying the presence or absence of variants (Table S4). The negative controls had no positions above the specified coverage threshold, and consequently no mtDNA variants were detected. For all samples and controls, the interpretation range was shown to be accurate and consistent with the variants that were called, enabling a reliable mtDNA profile to be reported and used for comparison.

3.3. Profile edits

A total of 698 variants were reported in the 21 samples and controls included in the study (Table 2). Without the use of AQME, 130 (18.6%) variant positions would have required manual adjustment of the CLC-reported variants by the analyst. In total, AQME successfully performed 211 edits with 99 (46.9%) attributed to the indel realignment in homopolymeric (or repeat) regions such as the three polycytosine stretches (C-stretches) located in the CR. All other edits (i.e. MNV separation, LHP MM filtering and IUPAC assignment for PHPs) were completed by the Create Mitochondrial Variant Table tool. Each variant profile modification performed by the AQME toolkit was identified with color highlighting and/or a comment in the mtDNA Table, visually showing the user the position and type of edit. During validation studies

Table 2
Sample profile details including the edits required to generate the forensic haplotype. Particular variants required more than one type of tool edit (e.g., AC deletions needed to be split into individual variants and then shifted from nps 514–515 to 523–524) and were therefore included in multiple "Tool Edits" columns. The "Edited Positions" columns represent the number of variant positions that required editing by the tool and/or user. The edit counts and final variant number included the np 3107 (deletion) placeholder.

Set	Sample	Bases Covered	Tool Edits				Edited Positions		Final Variants
			Shift Indels	MNV	MM	PHP	Tool	User	
Reference-Type									
	mtGAfrCO0001	16,569	14	14	1	0	17	0	102
	mtGAfrOH0001	16,511	1	0	0	0	1	0	54
	mtGAfrVA0010	16,569	6	11	3	0	16	10	58
	mtGAfrWI0001	16,569	5	2	7	1	10	0	39
	mtGCaucMN0001	16,569	1	0	0	0	1	0	2
	mtGCaucNY0002	16,569	3	0	0	2	5	0	40
	mtGCaucOH0003	16,569	5	2	2	0	5	0	43
	mtGCaucWI0008	16,569	2	0	0	0	2	0	37
	mtGHispCA0015	16,569	10	4	2	0	10	0	51
	mtGHispCA0031	16,569	12	11	4	1	16	1	41
	mtGHispPR0005	16,569	5	2	0	0	5	0	43
	mtGHispTX0008	16,569	12	9	10	0	12	0	41
	2800 M (Positive)	16,569	2	0	1	1	3	0	13
	Negative 1	0	0	0	0	0	0	0	0
Non-Probative									
	Sample 1	16,569	7	5	1	0	7	0	37
	Sample 2	16,569	6	4	2	2	8	0	23
	Sample 3	16,509	3	0	0	0	3	0	19
	Sample 4	2805	0	0	0	0	0	0	6
	Sample 5	533	0	0	0	0	0	0	1
	K562 (Positive)	16,569	5	6	0	2	9	5	48
	Negative 2	0	0	0	0	0	0	0	0
Total		**284,893**	**99**	**70**	**33**	**9**	**130**	**16**	**698**

it will be the onus of the laboratory to determine the tool flags that require additional scrutiny.

Three AQME-generated profiles required manual edits (16 variant positions) in order to adhere to forensic nomenclature (Table 2). Sample mtGHispCA0031 needed manual adjustment of the variant at np 16183 as part of the HV1 C-stretch. LHP caused by the T16189C in the HV1 region often results in A-C transversions at nps 16180–16183, which further complicate the alignment and variant detection in that region. However, review of the read mapping, Coverage Table, and supplemental read count output of this sample allowed the analyst to make an informed decision to modify the position to the expected A16183C (Fig. S8). The other two samples requiring manual edits both contained indel clusters that were not shifted properly by the Realign Variants tool due to the presence of a SNP in the same region. Although the haplogroup-diagnostic 9-bp deletion was correctly placed at nps 8281–8289 in two other samples (mtGAfrCO0001 and mtGHispCA0031), the presence of a variant at np 8270 in sample mtGAfrVA0010 prevented the group of deletions from shifting. Notably, the incorrect alignment was flagged by the correctly assigned haplogroup since the consensus sequence and not a nomenclature-dependent variant list is utilized by the Haplogrouper tool (discussed further below). Guided by the haplotype phylogeny, the analyst shifted the deletions and added the SNP to the variant profile in order to reflect the consensus sequence with the recommended nomenclature. Similarly, the analyst edited the HV2 C-stretch for K562, which contained a T310C and cytosine deletions that failed to shift properly. Better alignment and/or incorporation of supplemental tools could help to improve variant detection in these regions, recognize problematic motifs, and reduce the incidences of these types of manual edits. However, in each case AQME flagged these positions for further inspection by the analyst and the edits to the profile were tracked in the History (e.g., Fig. S5).

3.4. Haplogroup assignment

The AQME haplogroup assignments were identical to those determined by EMPOP for 15 of the 17 full mitogenomes (Table 3). The first full mitogenome discrepancy was a minor discordance observed between the assignments for K562 that resulted from the use of two different builds of Phylotree. In this case, the mtDNA haplogroup T2a1b1a1a2 did not exist in Build 16 (AQME) and was added in Build 17 (EMPOP). When the T2a1b1a1a2 annotated sequence was added to the AQME haplogroup database, the precise haplogroup was assigned. The second difference in haplogroup assignment involved reference-type sample mtGHispTX0008. Its full mitogenome haplotype was assigned to a more refined haplogroup by EMPOP (Y2a1a) than AQME (Y2a1). The profile contained one Y2a1a SNP (C2856T) but was missing the other Y2a1a haplogroup-defining SNP (G13135A). The inconsistency between the AQME and EMPOP haplogroup assignments can be explained by differences in the tools' respective haplogrouping methods, particularly how missing and private mutations impact the assignment. The EMPOP tool utilizes variant fluctuation rates within a particular clade to derive the most likely haplogroup, even when diagnostic polymorphisms are missing from the profile [13]. By contrast, AQME uses a simplified scoring strategy that was designed for the present study to err on the side of a conservative haplogroup assignment, favoring private over missing SNPs regardless of their individual mutation rates. As a result, EMPOP haplogroup calls may be more precise than AQME − though both are accurate − especially when a profile lacks haplogroup-defining polymorphisms as was the case for mtGHispTX0008. With the exception of this one sample, the AQME-assigned haplogroups were shown to be reliable and concordant with EMPOP when analyzing full mitogenome profiles (and the same Phylotree mtDNA tree version was utilized).

For full mitogenome data, the real-time haplogroup information provided by AQME enabled profiles to be checked for inconsistencies with phylogenetic nomenclature during review. For example, the haplogroup assignment of sample mtGAfrVA0010 was critical in identifying nomenclature inconsistent with its phylogeny. The consensus sequence generated from mtGAfrVA0010 was assigned to mtDNA haplogroup M1a1d, but the annotations added to the mtDNA Table showed that several phylogenetically diagnostic variants were missing from the profile (Fig. S9a). As discussed above, the C8270T and 9-bp

Table 3
AQME and EMPOP haplogroup assignments for samples including controls with mitogenome coverage (n = 19). Samples are split into two groups based upon mitogenome coverage: full/nearly full (> 16,500 bp) and partial. AQME assignments identical to those determined by EMPOP are indicated with a "-" in the table. The adjusted SNP score ignores phylogenetically neutral variants that are excluded from the Phylotree.org mtDNA tree. The maximum SNP score is the highest possible score based on the haplogroup-diagnostic SNPs present in the queried range, and if different, the maximum score for the entire mitogenome is listed in parentheses.

Mitogenome Coverage	Sample	EMPOP	AQME				
		Reported Haplogroup	Reported Haplogroup	Haplogroups Identified	Reported SNP Score	Corrected SNP Score	Maximum SNP Score
Full/Nearly Full							
	mtGAfrCO0001	L0a2a1b	–	1	911	929	950
	mtGAfrOH0001	L2a1b1	–	1	138	147	150
	mtGAfrVA0010	M1a1d	–	1	415	442[a]	490
	mtGAfrWI0001	X2a1a	–	1	278	293	320
	mtGCaucMN0001	H2a2a1b	–	1	7	10	10
	mtGCaucNY0002	U5b2b3	–	1	270	279	300
	mtGCaucOH0003	K1a4a1a2	–	1	308	326	350
	mtGCaucWI0008	T1a1b	–	1	331	340	340
	mtGHispCA0015	C1b7a	–	1	380	395	410
	mtGHispCA0031	B2m	–	1	332	344	350
	mtGHispPR0005	L3b1a + 152	–	1	313	328	340
	mtGHispTX0008	Y2a1a	Y2a1	1	241	250	280
	Sample 1	I4a	–	1	292	310	310
	Sample 2	H4a1a1a	–	1	126	144	150
	Sample 3	V1a1	–	1	324	339	360
	2800M	H1c + 152	–	1	78	87	90
	K562	T2a1b1a1a2	T2a1b1a1a	1	501	504	510
Partial							
	Sample 4	K	U8b	98	50	47[a]	50 (190)
	Sample 5	L3	–	3750	10	10	10 (180)

[a] The corrected score also includes adjustments to the value to correct for errors in scoring caused by bugs in the current version of AQME.

deletion at nps 8281–8289 expected in the profile were represented as 8270–8278 deletions. Although shifting the deletions to nps 8281–8289 does not adhere to maximum parsimony because of the additional polymorphism (C8270T), it is consistent with the M1a1d motif and therefore the recommended nomenclature [8]. The analyst was able to edit the profile and adjust the variants reported to reflect the appropriate call during review (Fig. S9b). The nomenclature discrepancy was also apparent in the Haplogroup Report because the 9-bp deletion was misclassified as both found and private (Fig. S9c). The duplication of the misaligned deletions in the report, caused by the way the haplogroup and sample variants are compared, resulted in a reduced score (415). If the bug affecting the classification of this indel cluster was corrected (i.e. only as found since the mtGAfrVA0010 and M1a1d sequences are identical in that region), the score would increase to 442 (or 457 with phylogenetically neutral variants ignored). Although the fix would ensure appropriate scoring that may be useful for validating an assignment, the haplogroup derived by AQME Haplogrouper was accurate and provided an invaluable resource in recognizing incorrect profile nomenclature contrary to the phylogenetic alignment.

The AQME Haplogrouper was tested with two incomplete mitogenomes and the results were less robust than EMPOP. This is due in part to the tool differences that predominantly affect haplogroup assignment of partial profiles. In particular, the EMPOP tool automatically reports the most conservative node of haplogroups with the same likelihood, which is beneficial when multiple haplogroups of equal likelihood/score are identified. On the other hand, AQME reports all of the top-scoring haplogroups and the user must identify the basal haplogroup for a conservative call. For example, the basal haplogroup assignment for Sample 5 (L3) was automatically output by the EMPOP tool; yet AQME reported over 3500 haplogroups with equal alignment and SNP scores (Table 3), therefore requiring the analyst to pinpoint the ancestral node. Aside from AQME's inability to automatically deduce the most basal haplogroup from a long list of top-scoring options, which is likely to occur with partial mitogenome data, AQME furthermore errored by producing a list of haplogroups for Sample 4 that included two wrong lineages. Although the basal haplogroup of this sample (K) was correctly reported by EMPOP, the most basal haplogroup of the 98 reported by AQME was U8b (which was determined by the analyst). Although U8b is phylogenetically close to K and includes five of the six variants called in the sample profile, this haplogroup is incorrect for Sample 4. Upon investigation of this inaccuracy, it was discovered that the AQME Haplogroup Report only included these five SNPs instead of the six total profile variants (Fig. S10). This incorrect Haplogroup Report for Sample 4 revealed a bug in the Haplogrouper in which some portions of the fragmented interpretation range did not align to the haplogroup sequences; therefore the variant in that region was not considered towards the alignment and SNP scores. In the case of Sample 4, a haplogroup K diagnostic SNP (T14798C) was in the variant profile but, due to the bug in the tool, was not used in the haplogrouping alignment against the Phylotree haplogroups. This caused the inaccurate assignment to U8b (instead of K). Furthermore, the SNP score was incorrectly reported as 50 due to the exclusion of T14798C when it should have been 47 (5 shared SNPs, 1 private SNP). Due to these issues inherent to partial profile haplogrouping in the current AQME version, the results may be inaccurate until the bugs in the tool are fixed. Although the interpretation range (i.e. region sequenced) and the specific haplogroup motif limit the precision of haplogrouping for partial profiles regardless of the tool utilized (AQME, EMPOP, etc.), AQME at this time requires substantial analyst discretion to ensure accurate haplogroup assignments are generated from incomplete mitogenome data.

Modifications to the Mitochondrial Haplogrouper in future versions of AQME would improve upon its functionality to enable more precise and accurate haplogroup assignment in the CLC Workbench without user intervention. One such enhancement would automatically adjust the SNP score to be reported without phylogenetically neutral variants (such as T16519C, for example) that artificially reduce the score. Additionally, the result would include normalization based on the maximum possible score for both the queried region and full mitogenome. These modifications to the SNP scoring would provide the user with metrics to better evaluate the validity of a haplogroup assignment. For example, mtGCaucMN0001 was assigned to haplogroup H2a2a1b with a corrected score of only 10. However, the maximum score possible for H2a2a1b is 10 because this haplogroup contains a single diagnostic variant. Therefore the comparison of the queried profile score to the maximum possible haplogroup score would help to support the assignment in spite of a relatively low score; or, conversely, that a high score is not necessarily reliable if the maximum score is substantially greater. Though the Haplogrouper may be useful in identifying errant (or at least questionable) haplogroup assignments for full mitogenome data, partial profiles pose a more challenging task due to reduced coverage of the mitogenome and haplogroup-diagnostic variants. A majority of the time this results in more than one haplogroup with the greatest alignment score, and therefore requires the user to determine the basal node from a potentially long list of divergent options. AQME improvement plans include the ability to automatically identify the most conservative haplogroup from those of equal alignment and SNP scores. This would streamline the analysis and allow for useful phylogenetic information to be gleaned from even partial mitogenome profiles. The addition of specific AQME features will further enable the haplogroup assignment reported by AQME to serve as a means to quickly QC each haplotype [6,30,31]. Although not observed in this study, a large number of private mutations or missed haplogroup-defining mutations in a mitogenome profile may indicate the occurrence of an error during processing (e.g., flagging potential artificial recombination [19] based on visual inspection of the profile as demonstrated in Fig. S11). Future enhancements would allow for these types of issues to be automatically flagged by the tool. The use of the haplogroup assignment as a QC measure will also require the development of specific guidelines to identify questionable profiles for further analyst inspection. Overall, the real-time application of mtDNA haplogroup assignment during analyst review allows for an increase in efficiency, assists in phylogenetic nomenclature guidance, and can be an invaluable QC assessment of the profile.

3.5. Profile review

Sample profiles were approved following completed review by the analyst. Changes to the mtDNA Table were not possible in approved profiles, and removing the approval to make edits was subsequently tracked in the History including the user and an explanation note. The three files exported for each sample included an Excel workbook, CODIS XML and History PDF. The Excel file contained the mtDNA Table, Profile Report and Coverage Table along with seven supplemental reports (Supplemental File 1); therefore more than 200 reports were generated and available for ancillary review of the 21 samples analyzed. Meta-analysis of the data was then performed to determine various metrics (e.g., average variant frequency, average variant quality) utilized for sample interpretation (i.e. profile acceptability). The Profile Report was also exported as a text file allowing for easy review of the haplogroup, range and profile (Fig. S7). All CODIS-formatted XML files were successfully imported into the AFMES-AFDIL's laboratory information management system, Laboratory Information System Applications (LISA; Future Technologies, Inc.). The range, variants and haplogroup imported into the LISA database were consistent with those approved by the analyst within the CLC Workbench. Additionally, the NGS mitogenome haplotypes were compared to the previously generated Sanger data, and the profiles were 100% concordant in overlapping regions with the exception of minor variation of low-level and length variants (Table S4). Overall, the AQME toolkit assisted in the generation of reliable mtDNA profiles consistent with forensic guidelines.

4. Conclusion

AQME is a toolkit developed for the forensic analysis of NGS mitogenome data within the CLC Genomics Workbench. The AQME tools function as part of a user-defined bioinformatics pipeline (CLC workflow) to automate mtDNA analysis from sequence data to forensic profile. This software solution is agnostic to sample type, library preparation method, and sequencing platform – and can furthermore accommodate Sanger data by utilizing the CLC Workbench for analysis. Thus, AQME can facilitate the forensic community's transition to NGS for mtDNA analysis by providing a toolkit that is backward compatible with CE methods. Moreover, AQME offers integrated mtDNA haplogroup assignment to inform phylogenetic nomenclature necessary for profile comparisons. This automatic haplogrouping feature may expand the use of the mtDNA haplogroup in forensics as a QC measure and maternal ancestry predictor. Although enhancements may improve upon certain functionalities, the CLC Genomics Workbench with the AQME toolkit allows for robust analysis pipelines that generate reliable mtDNA profiles for forensic use. At this time, AQME may be shared by the AFMES-AFDIL with other laboratories using CLC for the purposes of validation, testing, or collaboration on a specified project. Tool sharing must be approved by QIAGEN and there is a cost to the receiving laboratory for tool maintenance and support. In the future, expanded use/interest may drive down AQME tool maintenance costs and/or promote potential commercial availability of AQME (or similar forensic-specific mtDNA analysis tool) by QIAGEN Bioinformatics.

The broader implementation of NGS and specialized analysis tools like AQME will allow mtDNA sequencing to become more practical for routine forensic use. At the AFMES-AFDIL, specifically, the AQME toolbox was instrumental to the validation of two NGS mitogenome methods [27,32] in which it was shown to produce accurate forensic profiles for high-quality, degraded, and chemically treated DNA samples. In just one year since implementation, the AQME toolbox has been used to analyze thousands of mtDNA haplotypes for both casework and research applications (e.g., [12]). With efficient methods for data analysis now available, mitogenome data from forensically relevant population samples can be generated for public use. The availability of accurately reported population reference haplotypes will maximize the utility of the mitogenome in forensic applications.

Declaration of authorship

KSA, MAP, TPM and CM contributed to experimental design, direction of the project, and writing of the manuscript. KSA and MAP were responsible for the overall software design and testing iterations, while CB and PD were responsible for the software development.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

CB and PD were employees of QIAGEN Bioinformatics at the time of the tool development, and the software was the result of a U.S. federal government procured contract. No other conflicts of interest exist for these or any of the authors.

Acknowledgements

The authors are extremely grateful to Toni Diegoli (Defense Forensic Science Center) for invaluable tool design input and manuscript review, Timmathy Cambridge (AFMES) for technical support, and Walther Parson (EMPOP) for mtDNA haplogroup assignments. The authors would also like to thank Erin Gorden, Jennifer Higginbotham and Joseph Ring (ARP Sciences, LLC; AFMES-AFDIL) for laboratory and analysis assistance; James Canik (ARP Sciences, LLC), Nicole Wilkins, Lt Col Garner, Lt Col Briones, COL Ladd Tremaine, and COL Louis Finelli (AFMES) for administrative and logistical support; and the two anonymous reviewers for constructive input on both the manuscript and AQME. The opinions or assertions presented are the private views of the authors and should not be construed as official or as reflecting the views of the Department of Defense, its branches, the U.S. Army Medical Research and Materiel Command, the Defense Health Agency, or the Armed Forces Medical Examiner System. Mention of commercial products does not constitute a recommendation or endorsement by the authors and/or their associated organization/institute.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2017.09.010.

References

[1] M.A. Peck, M.D. Brandhagen, C. Marshall, T.M. Diegoli, J.A. Irwin, K. Sturk-Andreaggi, Concordance and reproducibility of a next generation mtGenome sequencing method for high-quality samples using the Illumina MiSeq, Forensic Sci. Int. Genet. 24 (2016) 103–111.
[2] W. Parson, C. Strobl, G. Huber, B. Zimmermann, S.M. Gomes, L. Souto, et al., Evaluation of next generation mtGenome sequencing using the ion torrent personal genome machine (PGM), Forensic Sci. Int. Genet. 7 (2013) 543–549.
[3] J.A. McElhoe, M.M. Holland, K.D. Makova, M.S. Su, I.M. Paul, C.H. Baker, et al., Development and assessment of an optimized next-generation DNA sequencing approach for the mtgenome using the Illumina MiSeq, Forensic Sci. Int. Genet. 13 (2014) 20–29.
[4] W. Parson, G. Huber, L. Moreno, M.B. Madel, M.D. Brandhagen, S. Nagl, et al., Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples, Forensic Sci. Int. Genet. 15 (2015) 8–15.
[5] J.L. King, B.L. LaRue, N.M. Novroski, M. Stoljarova, S.B. Seo, X. Zeng, et al., High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq, Forensic Sci. Int. Genet. 12 (2014) 128–135.
[6] W. Parson, L. Gusmao, D.R. Hares, J.A. Irwin, W.R. Mayr, N. Morling, et al., DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing, Forensic Sci. Int. Genet. 13 (2014) 134–142.
[7] H.J. Bandelt, W. Parson, Consistent treatment of length variants in the human mtDNA control region: a reappraisal, Int. J. Legal Med. 122 (2008) 11–21.
[8] Scientific Working Group on DNA Analysis Methods, Interpretation Guidelines for Mitochondrial DNA Analysis by Forensic DNA Testing Laboratories, (2013).
[9] J.L. King, A. Sajantila, B. Budowle, mitoSAVE: mitochondrial sequence analysis of variants in Excel, Forensic Sci. Int. Genet. 12 (2014) 122–125.
[10] M.M. Holland, E.D. Pack, J.A. McElhoe, Evaluation of GeneMarker(R) HTS for improved alignment of mtDNA MPS data haplotype determination, and heteroplasmy assessment, Forensic Sci. Int. Genet. 28 (2017) 90–98.
[11] J.D. Churchill, J.L. King, R. Chakraborty, B. Budowle, Effects of the Ion PGM Hi-Q sequencing chemistry on sequence data quality, Int. J. Legal Med. (2016).
[12] J.D. Ring, K. Sturk-Andreaggi, M.A. Peck, C.K. Marshall, A performance evaluation of nextera XT and KAPA HyperPlus for rapid illumina library preparation of long-range mitogenome amplicons, Forensic Sci. Int. Genet. 29 (2017) 174–180.
[13] A.W. Rock, A. Dur, M. van Oven, W. Parson, Concept for estimating mitochondrial DNA haplogroups using a maximum likelihood approach (EMMA), Forensic Sci. Int. Genet. 7 (2013) 601–609.
[14] A. Kloss-Brandstatter, D. Pacher, S. Schonherr, H. Weissensteiner, R. Binna, G. Specht, et al., HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups, Hum. Mutat. 32 (2011) 25–32.
[15] R. Clima, R. Preste, C. Calabrese, M.A. Diroma, M. Santorsola, G. Scioscia, et al., HmtDB 2016: data update, a better performing query system and human mitochondrial DNA haplogroup predictor, Nucleic Acids Res. 45 (2017) D698–D706.
[16] H. Weissensteiner, D. Pacher, A. Kloss-Brandstatter, L. Forer, G. Specht, H.J. Bandelt, et al., HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing, Nucleic Acids Res. 44 (2016) W58–63.
[17] D. Navarro-Gomez, J. Leipzig, L. Shen, M. Lott, A.P. Stassen, D.C. Wallace, et al., Phy-Mer: a novel alignment-free and reference-independent mitochondrial haplogroup classifier, Bioinformatics 31 (2015) 1310–1312.
[18] M.T. Lott, J.N. Leipzig, O. Derbeneva, H.M. Xie, D. Chalkia, M. Sarmady, et al., mtDNA variation and analysis using mitomap and mitomaster, Curr. Protoc. Bioinf. 44 (1.23) (2013) 1–26.
[19] H.J. Bandelt, T. Kivisild, Quality assessment of DNA sequence data: autopsy of a mis-sequenced mtDNA population sample, Ann. Hum. Genet. 70 (2006) 314–326.
[20] A. Brandstatter, T. Sanger, S. Lutz-Bonengel, W. Parson, E. Beraud-Colomb, B. Wen, et al., Phantom mutation hotspots in human mitochondrial DNA, Electrophoresis 26 (2005) 3414–3429.
[21] T.D. Anderson, J.P. Ross, R.K. Roby, D.A. Lee, M.M. Holland, A validation study for

the extraction and analysis of DNA from human nail material and its application to forensic casework, J. Forensic Sci. 44 (1999) 1053–1056.
[22] R.M. Andrews, I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, N. Howell, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA, Nat. Genet. 23 (1999) 147.
[23] S. Anderson, A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, et al., Sequence and organization of the human mitochondrial genome, Nature 290 (1981) 457–465.
[24] QIAGEN, CLC Genomics Workbench User Manual, Version 7.5, 1 ed., QIAGEN, 2014.
[25] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (2009) E386–94.
[26] R.S. Just, M.K. Scheible, S.A. Fast, K. Sturk-Andreaggi, A.W. Rock, J.M. Bush, et al., Full mtGenome reference data: development and characterization of 588 forensic-quality haplotypes representing three U.S. populations, Forensic Sci. Int. Genet. 14 (2015) 141–155.
[27] C.K. Marshall, K. Sturk-Andreaggi, J. Daniels-Higginbotham, R.S. Oliver, S. Barritt-Ross, T.P. McMahon, Performance evaluation of a mitogenome capture and Illumina sequencing protocol using non-probative, case-type skeletal samples: implications for the use of a positive control in a next-generation sequencing procedure, Forensic Sci. Int. Genet. (2017), http://dx.doi.org/10.1016/j.fsigen.2017.09.001 (in press).
[28] W. Parson, A. Dur, EMPOP–a forensic mtDNA database, Forensic Sci. Int. Genet. 1 (2007) 88–92.
[29] M. van Oven, PhyloTree build 17: growing the human mitochondrial DNA tree, Forensic Sci. Int.: Genet. Suppl. Ser. 5 (2015) e392–e394.
[30] H.J. Bandelt, M. van Oven, A. Salas, Haplogrouping mitochondrial DNA sequences in legal medicine/forensic genetics, Int. J. Legal Med. 126 (2012) 901–916.
[31] A. Salas, H.J. Bandelt, V. Macaulay, M.B. Richards, Phylogeographic investigations: the role of trees in forensic genetics, Forensic Sci.Int. 168 (2007) 1–13.
[32] M.A. Peck, K. Sturk-Andreaggi, J.T. Thomas, R.S. Oliver, S.M. Barritt-Ross, C.K. Marshall, Developmental validation of a Nextera XT mitogenome Illumina MiSeq sequencing method for high quality samples, Forensic Sci. Int. Genet. (2017) (in review).

EXHIBIT 5

FAMILY REFERENCE COLLECTION FORM

Armed Forces DNA Identification Laboratory AFDIL Case #:_____

DONOR INFORMATION

FIRST NAME	MIDDLE NAME	LAST NAME

HOME TELEPHONE	DATE OF BIRTH (Month/Day/Year)	GENDER (Check Box) ☐ Male ☐ Female

HOME STREET ADDRESS	DCIPS-FAMILY MEMBER NUMBER

CITY	STATE	ZIP CODE	COUNTRY OF BIRTH (If not the United States)

Check Box For Your Ethnic Group (See Below For Group Classifications)
☐ Caucasian ☐ African American ☐ Hispanic ☐ American Indian
☐ Asian Or Pacific Islander ☐ Other (Specify): _____

FOR AFDIL USE ONLY (Check All Applicable)
☐ MtDNA Reference ☐ NucDNA Reference
☐ Y DNA Reference ☐ Direct Reference
☐ Ineligible Reference ☐ Exclusion Reference

ETHNIC GROUP

Caucasian: A person having origins in any of the peoples of Europe, North America, or the Middle East (*not of Hispanic origin*).

African American: A person having origins in any of the black racial groups of Africa (*not of Hispanic origin*).

Hispanic: A person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish cultural origin, regardless of race.

American Indian: A person having origins in any of the original peoples of North America, and who maintains a cultural identification through tribal affiliation or community recognition.

Pacific Islander or Asian: A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian Subcontinent or the Pacific Islands.

MISSING INDIVIDUAL INFORMATION

FIRST NAME	MIDDLE NAME	LAST NAME	Date of Birth (Month/Day/Year)

CONFLICT (Check Applicable Box)
☐ WW I ☐ WW II ☐ Korean War ☐ Cold War ☐ Vietnam War ☐ Other (Specify):_____

BRANCH OF SERVICE (Check Applicable Box)
☐ US Army ☐ USAAF (WWII) ☐ US Navy ☐ USMC ☐ US Air Force ☐ Other (Specify):_____

SSN Or SERVICE NUMBER	RANK	DCIPS CASE NUMBER	JPAC ISN NUMBER	JPAC INCIDENT NUMBER

REFNO # (SEA Only)	FIELD SEARCH CASE NUMBER	MACR (WWII USAAF Only)	BUNO (USN Only)

FAMILY RELATIONSHIP INFORMATION

(See Page 3 and circle your relationship to the missing individual)

Please list your relationship to the missing individual:_____

Are you adopted? ☐ YES ☐ NO

Are you a step-sibling to the missing individual (no shared biological parent)? ☐ YES ☐ NO

Are you a half-sibling to the missing service member (shared biological parent)? ☐ YES* ☐ NO

* If yes, do you share the same: ☐ Mother ☐ Father

Highlighted Information MUST Be Completed By Donor For Sample To Be Accepted By AFDIL

POTENTIAL LIVING OR DECEASED BIOLOGICAL DONORS FOR DNA ANALYSIS

FATHER/MOTHER OF MISSING INDIVIDUAL

NAME	RELATIONSHIP	ADDRESS	PHONE

SPOUSE/BIOLOGICAL CHILDREN OF MISSING INDIVIDUAL

NAME	RELATIONSHIP	ADDRESS	PHONE

BROTHERS AND SISTERS OF MISSING INDIVIDUAL

NAME	RELATIONSHIP	ADDRESS	PHONE

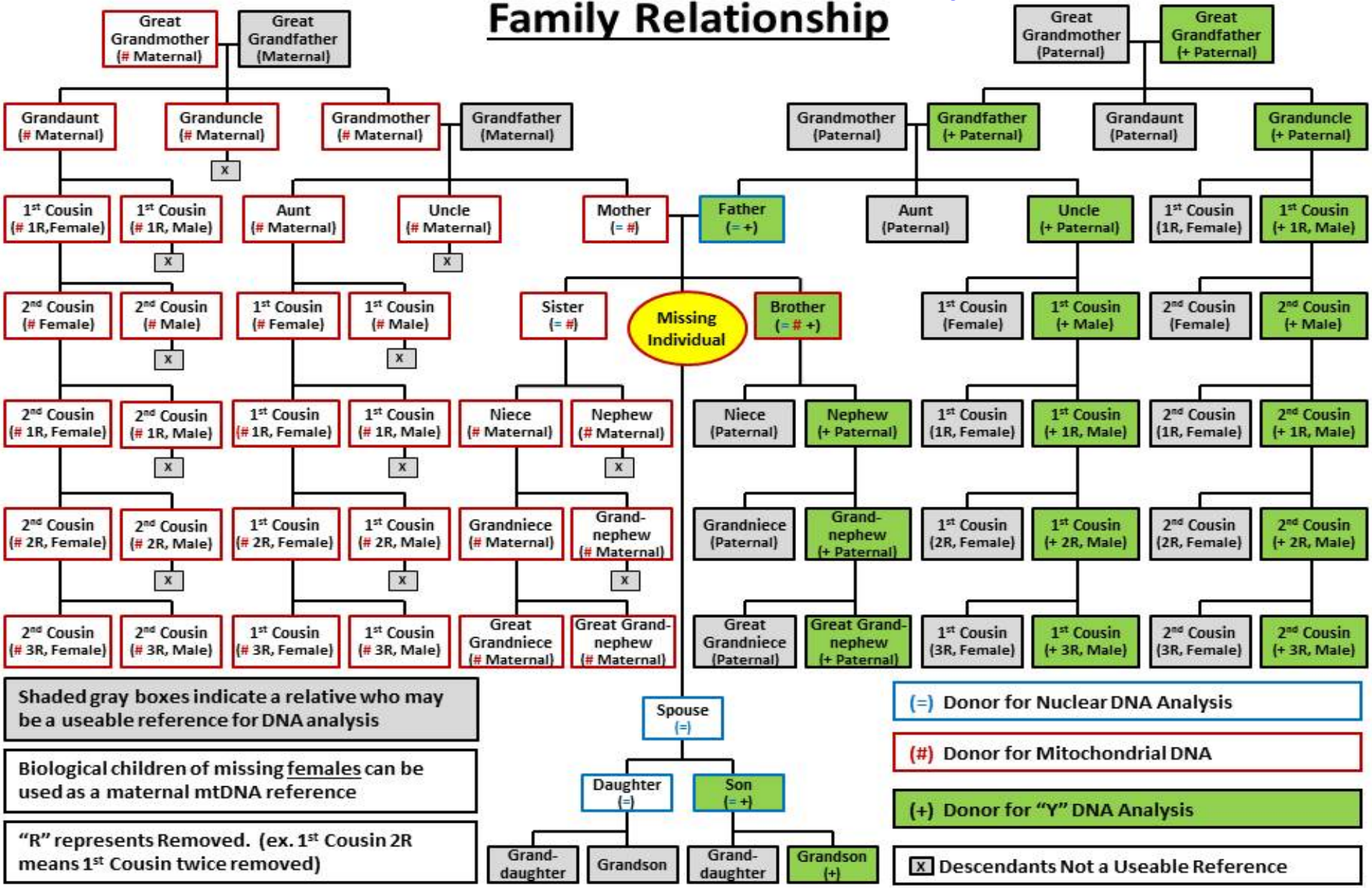
UNCLES/AUNTS OF MISSING INDIVIDUAL

NAME	RELATIONSHIP	ADDRESS	PHONE

NEPHEWS/NIECES/COUSINS OF MISSING INDIVIDUAL

NAME	RELATIONSHIP	ADDRESS	PHONE

Family Relationship



If your relationship is not represented in the chart above, please describe, in detail, your relationship to the missing individual: _____

DONOR CONSENT FORM

AFDIL Case #: _____

PRIVACY ACT STATEMENT

In accordance with the Privacy Act of 1974 (Public Law 93-579), this notice informs you of the purpose of the form and how it will be used. Please read it carefully.

AUTHORITY: 10 U.S.C. §1471; Public Law 104-191; Deputy Secretary of Defense Memorandum, "Establishment of a Repository of Specimen Samples, December 16, 1991; and DoDI 5154.30.

PRINICIPAL PURPOSES: To establish a DNA reference specimen repository and database of information from kindred family members of unaccounted for/unidentified service members or other individuals needing to be identified. DNA will be extracted from a biological specimen or personal effect and used in identifying human remains.

ROUTINE USE: Use and disclosure of your records outside of DoD may also occur in accordance with the DoD Blanket Routine Uses published at http://dpclo.defense.gov/privacy/SORNs/blanket_routine_uses.html and as permitted by the Privacy Act of 1974, as amended (5 U.S.C. 552a(b)). Any protected health information (PHI) in your records may be used and disclosed generally as permitted by the HIPAA Privacy Rule (45 CFR Parts 160 and 164), as implemented within DoD.

DISCLOSURE: Voluntary. Failure to provide a reference sample or requested information may render DNA identification impossible.

STATEMENT OF CONSENT

The above answers are correct to the best of my knowledge and belief, and I understand that my answers are important in determining my kindred family relationship to an unaccounted for service member or other unaccounted for individual. I have also read the Privacy Act statement above. Realizing that nuclear or mitochondrial deoxyribonucleic acid (DNA) may be extracted from my biological specimen or personal effect and used in the identification of a kindred family member, I agree to donate a biological specimen or personal effect, to have my DNA control region analyzed and if necessary the whole mitochondrial DNA (mtDNA) genome, and to have my name and other relevant typing information placed in a confidential registry or database for identification and statistical analysis. I am voluntarily donating a biological specimen including, but not limited to, blood, buccal swab, or personal effect, as required and consent to the Department of Defense using the information and specimens for the identification of any unaccounted for family member.

DISCLOSURE: Mitochondrial DNA (mtDNA) sequencing data results will be reported as differences compared to the revised Cambridge Reference Sequence (rCRS). Certain differences may have medical implications. I understand that the Armed Forces DNA Identification Laboratory (AFDIL) is not a medical genetic testing laboratory and is not engaged in the practice of medicine. If I have medical concerns about my mtDNA sequencing data, I understand that I should consult my doctor.

Use Of Your Sample By AFDIL For Training, Research, or Validation

Use of your anonymized/confidential DNA information will allow AFDIL scientists to improve laboratory testing protocols, test the validity of new techniques, and generally advance the overall DNA identification process. Your refusal to consent for the use of your anonymized sample and/or data will in no way affect the use of your sample for identification of your family member.

Please Check Appropriate Box And Initial:

☐ _____ YES, I consent to the use of my sample/DNA information for training, research and/or validation purposes.
Initials

☐ _____ NO, I do not want my sample/DNA information utilized for training, research or validation purposes.
Initials

DISCLOSURE: Failure to provide this information will be taken as consent by the donor to use the donor's anonymized DNA information for training, research and/or validation purposes to assist in the identification of unaccounted for family members.

DNA Report Request

Please Check The Appropriate Box and Initial:

☐ _____ YES, I authorize my DNA report(s) to be sent to me at the address shown on page 1.
Initials

☐ _____ NO, I do not want my DNA report sent to me.
Initials

DISCLOSURE: Failure to provide this information will be taken as consent by the donor to have their DNA report sent to them.

SIGNATURE OF DONOR	PRINT DONOR NAME	DATE

SIGNATURE OF COLLECTOR	PRINT COLLECTOR NAME	DATE

Check FRS Collection Source: ☐ Service Collection ☐ DPAA Collection ☐ Family Update Collection ☐ Other (Please Specify):_____

Highlighted Information MUST Be Completed By Donor For Sample To Be Accepted By AFDIL

EXHIBIT
6

REPORT OF THE

DEFENSE SCIENCE BOARD

TASK FORCE

ON

THE USE OF DNA TECHNOLOGY FOR IDENTIFICATION OF ANCIENT REMAINS

JULY 1995

OFFICE OF THE UNDER SECRETARY OF DEFENSE
FOR ACQUISITION & TECHNOLOGY
WASHINGTON, D.C. 20301-3140

This report is a product of the Defense Science Board (DSB). The DSB is a Federal Advisory Committee established to provide independent advice to the Secretary of Defense. Statements, opinions, conclusions and recommendations in this report do not necessarily represent the official position of the Department of Defense.

This document is UNCLASSIFIED.

Security review completed 13 February 1995 by OASD (Public Affairs) directorate for Freedom of Information and Security Review Case number 95-J- 0529.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp. Date: Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS N/A
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Distribution Statement A. Approved for Public Release: Distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) N/A	5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A
6a. NAME OF PERFORMING ORGANIZATION Defense Science Board, Ofc of the Under Secy of Def (A&T) 6b. OFFICE SYMBOL (If applicable) DSB/OUSD(A&T)	7a. NAME OF MONITORING ORGANIZATION N/A
6c. ADDRESS (City, State, and ZIP Code) The Pentagon, Room 3D865 Washington, DC 20301-3140	7b. ADDRESS (City, State, and ZIP Code) N/A
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Science Board, OUSD(A&T) 8b. OFFICE SYMBOL (If applicable) DSB/OUSD(A&T)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N/A
8c. ADDRESS (City, State, and ZIP Code) The Pentagon, Room 3D865 Washington, DC 20301-3140	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO.: N/A PROJECT NO.: N/A TASK NO.: N/A WORK UNIT ACCESSION NO.: N/A
11. TITLE (Include Security Classification) Report of the Defense Science Board Task Force on The Use of DNA Technology for Identification of Ancient Remains, Unclassified.	
12. PERSONAL AUTHOR(S) N/A	
13a. TYPE OF REPORT Final 13b. TIME COVERED FROM N/A TO N/A	14. DATE OF REPORT (Year, Month, Day) 1995 July 15. PAGE COUNT 155
16. SUPPLEMENTARY NOTATION N/A	
17. COSATI CODES FIELD / GROUP / SUB-GROUP	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
19. ABSTRACT (Continue on reverse if necessary and identify by block number)	

DTIC QUALITY INSPECTED 5

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL Diane L.H. Evans	22b. TELEPHONE (Include Area Code) (703) 695-4157/8 22c. OFFICE SYMBOL DSB/OUSD(A&T)

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

OFFICE OF THE SECRETARY OF DEFENSE
WASHINGTON, D.C. 20301-3140

DEFENSE SCIENCE
BOARD

16 JUL 1995

MEMORANDUM FOR UNDER SECRETARY OF DEFENSE (ACQUISITION & TECHNOLOGY)

SUBJECT: Report of the Defense Science Board (DSB) Task Force on the Use of DNA Technology for Identification of Ancient Remains

I am pleased to forward the final report of the DSB study on the Use of DNA Technology for Identification of Ancient Remains. The report focuses on a new and exciting area of life science technology that can have a dramatic impact on the Department's ability to resolve current and future issues concerning the fullest possible accounting of prisoners of war and missing in action (POW/MIAs).

The report's recommendations could serve to support broad policy adjustments for the Department of Defense on issues concerning identification of ancient war remains. I concur with the observations and recommendations of the Task Force, and recommend that you forward the report to the Secretary of Defense.

Craig I. Fields
Chairman

OFFICE OF THE SECRETARY OF DEFENSE
WASHINGTON, D.C. 20301-3140

DEFENSE SCIENCE
BOARD

MEMORANDUM FOR CHAIRMAN, DEFENSE SCIENCE BOARD

SUBJECT: Report of the Defense Science Board (DSB) Task Force on the Use of DNA Technology for Identification of Ancient Remains

Attached is the report of the DSB study on the use of DNA Technology for Identification of Ancient Remains. This DSB Task Force was formed to study the issues and provide findings concerning the use of DNA comparison techniques for ancient remains identification. The primary purpose was to determine the feasibility of utilizing DNA techniques to identify unassociated ancient remains from past conflicts.

The Task Force heard presentations from a wide range of scientific and medical experts from within and outside the Department of Defense. We also reviewed written information from published and unpublished sources that was pertinent to our terms of reference.

The DSB Task Force has found that mitochondrial DNA (mtDNA) sequencing currently offers the best means of identifying those skeletal remains that cannot be identified through traditional means. The Task Force finds that current DNA identification efforts are supported by sufficient scientific evidence to proceed, in particular with application of mtDNA sequencing to ancient remains from the Korean conflict.

The Task Force also finds that operations at the Armed Forces DNA Identification Laboratory (AFDIL) are conducted appropriately such that families can rely on the results generated.

Joshua Lederberg
Task Force Chairman

Report
of the

DEFENSE SCIENCE BOARD TASK FORCE
ON
THE USE OF DNA TECHNOLOGY
FOR IDENTIFICATION OF ANCIENT REMAINS

January 20, 1995

Joshua Lederberg
Rockefeller University
Task Force Chairman

Jan Bashinski
California Department of Justice

Bruce Budowle
Federal Bureau of Investigation

Peter Gill
British Forensic Science Service

Mark Stoneking
Pennsylvania State Univeristy

Douglas Wallace
Emory University

Bruce Weir
North Carolina State University

George Whitesides
Harvard University

CDR Mark Jensen, USN
Executive Secretary

LTC John Dertzbaugh, USA
DSB Secretariat Representative

REPORT SUMMARY

The Defense Science Board (DSB) was requested by the Defense POW/MIA Office (DPMO) to address key issues arising from efforts to identify skeletal remains using new DNA testing technologies (Annex A and B). The Armed Forces DNA Identification Laboratory (AFDIL) of the Armed Forces Institute of Pathology (AFIP), a Department of Defense organization, is currently performing this testing under a Memorandum of Agreement for the U.S. Army Central Identification Laboratory, Hawaii (CILHI) of the U.S. Army Casualty and Memorial Affairs Operations Center (CMAOC) with funds from the U.S. Army Deputy Chief of Staff for Personnel (DCSPER).

A priority of the United States government has been the recovery and identification of the remains of American servicemembers from Southeast Asia, the Cold War era, and Korea. Currently, there are over 2,200 servicemembers from Southeast Asia, 132 servicemembers from the Cold War period, and over 8,100 servicemembers from the Korean conflict, whose remains have not been recovered and/or identified.

Current remains testing by the AFDIL involves mitochondrial DNA (mtDNA) sequencing, a new technology used only by a few laboratories in the world for forensic identification purposes. The Assistant Secretary of Defense (Health Affairs) [ASD(HA)] with input from civilian organizations developed a Quality Assurance Program for mtDNA testing of ancient remains. An important component of this program is the formation of an oversight committee composed of civilian technical consultants.

MtDNA testing is currently performed on repatriated remains with a name association, primarily from Southeast Asia. It is estimated that 500 cases from Southeast Asia will require mtDNA testing over approximately 5 years.

It is estimated that 3,000 remains could be repatriated from North Korea. This is in addition to the 200 remains repatriated by North Korea already at CILHI and the 865 unidentified American remains from Korea interred in the National Memorial Cemetery of the Pacific, Honolulu, Hawaii. The vast majority would require

mtDNA testing due to lack of adequate dental and medical records. These remains would not generally have name associations.' Accordingly, a database of family reference mtDNA sequences would be constructed and mtDNA sequences of these remains would be compared to it.

The DSB Task Force finds that mtDNA sequencing currently offers the best means of identifying those skeletal remains that cannot be identified through traditional means. The Task Force finds that operations at the AFDIL are conducted appropriately, such that identification of military remains using mtDNA technology, is defendable and that families can rely on the results generated in these cases. The Task Force finds that mtDNA sequencing in conjunction with other nonDNA evidence could provide identifications on the unassociated Korean remains. The cost of this program over the next 12 years would be approximately $2 million annually over existing funding levels.

Principal Conclusions:

The Task Force finds that current DNA identification efforts are supported by sufficient scientific evidence to proceed, in particular with application of mtDNA sequencing to identify ancient remains from the Korean conflict.

The Task Force supports the Assistant Secretary of Defense (Health Affairs) in the creation of a scientific advisory board composed of civilian technical consultants.

CONTENTS

Annexes

Figures

Bibliography

Report
of the
DEFENSE SCIENCE BOARD TASK FORCE
ON THE USE OF DNA TECHNOLOGY
FOR IDENTIFICATION OF ANCIENT REMAINS

INTRODUCTION

Throughout United States history, the military services have to the best of their ability attempted to recover and identify it's deceased military personnel. In 1981, President Reagan placed the issue of accounting for American servicemembers from Southeast Asia as a matter of highest national priority. This position has been reaffirmed by all Presidents since. The Department of Defense has been tasked to investigate and account, to the greatest extent possible, for the "unaccounted for" Americans and repatriate, identify, and return the remains to their families. Today, there are over 2,200 servicemembers from Southeast Asia, 132 servicemembers from the Cold War era, and over 8,100 servicemembers from Korea, whose remains have not been recovered and/or identified (POW/MIA Fact Book, Department of Defense, October 1992).

The United States Army Central Identification Laboratory, Hawaii (CILHI) of the U.S. Army Casualty and Memorial Affairs Operation Center (CMAOC) is responsible for the recovery, identification, and processing of human remains from previous conflicts. CILHI uses traditional forensic odontological and anthropological methods, as well as other state of the art methods to identify human remains, including the Computer Assisted Post Mortem Identification (CAPMI) dental system and computerized craniofacial superimposition. CILHI uses state of the art photographic, microscopic, and radiographic equipment to accomplish their mission. The availability of records, the passage of time, and the environment to which remains have been exposed are obstacles to traditional identification efforts. Emerging technologies offer new opportunities.

In 1991, the Army contacted the Armed Forces DNA Identification Laboratory (AFDIL), a division of the Office of the Armed Forces Medical Examiner (OAFME) at the Armed Forces Institute of Pathology (AFIP) to apply mitochondrial DNA (mtDNA) technology to

the identification of human remains. The mtDNA analysis and other corroborating evidence has since successfully identified the remains of Americans recovered in Southeast Asia. This technology now offers the prospect of identifying remains in the absence of name association, dental, medical, fingerprint, or circumstantial evidence.

The AFDIL has been in the world-wide vanguard of activities to identify remains through the use of DNA technology, and thus carries on the fine tradition of military biomedical research. From the first command-directed immunization program, inoculation for smallpox in President Washington's Army, up to and including the present time, many military and civilian medical scientists continue to make seminal contributions to military and general medicine. Among the many contributions, are Beaumont's studies of digestion in 1824; the founding of the first American School of Preventive Medicine and Public Health in 1893; and Walter Reed's proof that mosquitoes transmit yellow fever in 1900; anti-malarial drugs such as mefloquine halotantrine; vaccines such as VEE, typhoid, hemorrhagic fever, adenovirus and meningococcus; plasma and albumin blood products, and CPDA-1, AS-1 and AS-3 blood preservatives. The fields of burn therapy and emergency medicine have their roots in the military and are largely patterned after developments in military medicine. Through the Advanced Research Projects Agency, the military funded efforts that led to the development of the CAT scan and MRI. (Dora Strother, Army Science Board, 50 Years of Accomplishments in Army Research and Development, Social & Scientific Systems, Inc. Bethesda, MD)

The benefit of this military initiative and leadership can now be seen in the specialized area of DNA analysis of human remains. The use of mtDNA identification on remains by AFDIL continues that legacy of pioneering advances in medicine that contribute to society at large. This national resource can be used to assist in the identification of remains not only of servicemen and women who died in battle, but also to other disaster related deaths such as aircraft mishaps, earthquakes, explosions, and fires.

Despite difficulties in extracting mtDNA from ancient remains, the AFDIL has been successful in positively correlating mtDNA extracted from skeletal remains to their maternal relatives in

many CILHI cases. In addition, the AFDIL has performed the mtDNA sequence analysis on the skeletal remains and awaits family blood reference specimens for comparison in numerous other CILHI cases.

In one case, "X-6", conflicting results were received from the mtDNA analysis of remains analyzed by the AFDIL and a laboratory of the University of California-Berkeley. The source of the discrepancy has not been determined. This discrepancy case caused the Department of Defense to re-examine the use of mtDNA and to take several actions.

First, uncertainty concerning the efficacy of mtDNA technology created by the discrepancy in case "X-6" caused the U.S. Army on February 3, 1994, to suspend using mtDNA to determine the identity of war remains without corroborating evidence. The Department of Defense supported that position.

Second, the Army recommended that the Department of Defense Science Board (DSB) establish a task force to examine the issues associated with using mtDNA to identify remains. In May 1994, the Defense POW/MIA Office (DPMO) accepted responsibility to be the Department of Defense sponsor for this Task Force. On June 20, 1994, the Under Secretary of Defense requested the Chairman of the Defense Science Board establish a Task Force on the use of DNA technology for identification of ancient remains. The issues (Annex A) for the members of the DSB Task Force (Annex B) were incorporated into a set of Terms of Reference (TOR).

Third, the military reviewed the standards used to perform mtDNA testing and noted that a set of formally recognized and widely accepted technical and quality assurance standards did not exist specifically for mtDNA testing of ancient skeletal remains. It has become well-understood that minuscule levels of contaminants can lead to erroneous results and that extraordinary measures of quality control are needed. The ASD(HA), with input from the forensic and genetics communities, developed a quality assurance program for the Department of Defense that would have credibility and acceptability within the scientific and legal communities, to the families, and to the general public (Annex C).

Fourth, portions of the contested remains were sent to the British Forensic Science Service (FSS, also known as the British Home Office) for analysis. Results confirmed AFDIL's findings

(Annex D). Furthermore, when subsequently and unknowingly challenged by CILHI with skeletal remains of the same case, the AFDIL twice more obtained the same mtDNA sequence result.

On January 6, 1995, the Army concluded that they had confidence in the AFDIL testing results. The Department of Defense then gave approval to respond to all Army requests for analysis and to consider that Army requests have had a waiver, at Army level, to the Army imposed moratorium (Annex E).

Meanwhile, the issue of repatriation and identification of remains from the Korean conflict has come to the forefront. On August 24, 1993, the Korean People's Army (KPA) signed an agreement with the United Nations Command (UNC) for cooperation in the recovery, repatriation, and identification of UNC remains located north of the Demilitarized Zone (DMZ). Since 1990, the North Koreans have repatriated 208 coffins containing purported American remains. Because of the condition of these remains and the paucity of relevant personal, medical, and dental records for servicemembers serving in the Korean action, mtDNA analysis offers the best prospect to identify these remains.

The TOR represent the issues that needed to be addressed before the military proceeded with mtDNA testing of skeletal remains, particularly "unassociated" Korean remains. Paramount is assurance that the technology is cost effective and that families can depend on the methods of identification used.

I. FEASIBILITY

DSB TOR: To determine the feasibility of using DNA techniques for identification of ancient remains as evidenced, in part, by success in identification efforts thus far. *[Is the conceptual basis for mtDNA identification of ancient skeletal remains workable? Is the discriminatory potential of mtDNA as currently obtained, and with or without other identification data, sufficient for individuation of skeletal remains from Southeast Asia and Korea?]*

Nuclear DNA typing has the capacity to be used for identification because DNA is different among all individuals, except identical twins. The potential exists for DNA tests to provide identifications which cannot be made in any other way. Any portion of skeletal remains could potentially be useful for DNA

identification. Since reference specimens for DNA comparison can be obtained from family members, it can be useful in situations even though premortem specimens are not available. In contrast, premortem records must be available for conventional identification methods using medical, dental or anthropological comparisons.

A. Forensic DNA Identification

Molecular techniques have revolutionized the biological sciences. Procedures for rapid DNA sequencing were developed in the 1970s, the Southern blot technique for DNA fragment sizing was developed a few years later, and the polymerase chain reaction (PCR) for DNA fragment amplification in 1985. These techniques have become well established and now are at the heart of innumerable research efforts in the biologic sciences. Molecular biologic techniques have long since moved from the research laboratory to the clinical service laboratory. The revolution created by this new technology has spread to the forensic sciences where DNA typing is taking its place alongside fingerprinting in terms of its impact on the criminal justice system. The basic molecular biologic principles are at this point well established and documented.

The Office of Technology Assessment released its report on the forensic uses of DNA typing in 1990. They concluded that "no scientific doubt remains that technologies already available can accurately detect genetic differences between humans." Similarly, the National Research Council (NRC) of the National Academy of Sciences issued a report in 1992, confirming the capability of DNA testing as a new and important technology to identify the origin of biologic trace evidence. Traditional serologic testing is based on genetic differences that are best characterized at the DNA level.

Courts of law have generally embraced the new DNA technology. The passage of the DNA Identification Act as a part of the 1994 Crime Bill by Congress, to spur creation of a national network of state DNA databases of convicted sex offenders and other felons, is a recognition of the value and validity of this DNA identification technology.

Application of DNA typing to the identification of human remains is obvious. Identification of tissue origin is being performed by many crime labs around the world. The AFDIL has been a leading laboratory devoted to the identification of human remains, and assisting other Federal Government agencies.

However, most DNA identification efforts have thus far focused on the typing of nuclear DNA of relatively recent vintage. Identification of ancient skeletal remains through mtDNA sequencing presents new issues which have not been a significant part of the larger discussions of the application of molecular biology to forensic identification.

B. MtDNA Sequence Identifications

Dr. Mary-Claire King first employed mtDNA to identify the Argentina "disappeared". In these cases, Dr. King would match the mtDNA sequence of children, whose parents were killed for political reasons, to that of purported maternal grandmothers. In 1991, Dr. Mark Stoneking described the use of mtDNA sequencing for the identification of a skull found in the Mojave desert, approximately four years after a 3 year-old girl was reported missing. In 1994, Dr. Peter Gill described the use of mtDNA to identify the Russian Romanov Tsar Nicholas II and his family.

In 1991, the AFDIL first successfully employed mtDNA sequencing to identify the skeletal remains of a servicemember killed in the Southeast Asian conflict; a report of this case was published in the *Journal of Forensic Sciences* in 1993. Subsequently, AFDIL has performed testing which has led to other identifications.

C. Mitochondria

Human mitochondria are thought to have evolved through incorporation of an intracellular symbiont (parasite) into early life forms (Lynn Margulis, Symbiosis and Cell Evolution, 1981, Freeman Publishing). This intracellular symbiont, similar to a primitive bacterium, had its own DNA. The symbiont flourished within the cytoplasm in harmony with the host cell. Mitochondrial symbionts would pass into the daughter cells of every dividing cell. In time, cells came to depend on the efficient energy utilization mechanisms of this symbiont. This theory of endosymbiotic origin explains many of the peculiar features of this intracytoplasmic organelle.

Mitochondria are the primary means of oxidative respiration of the cell. They are critical to the utilization of oxygen from the air to generate usable energy in the form of phosphorylated compounds for the cell. Hence, mitochondria are considered to be the "powerhouses of the cell".

D. Mitochondrial DNA

Human mtDNA is a circular DNA "particle", 16,569 base pairs in length. The complete sequence from a composite of individuals was published in the journal *Nature* in 1981 by Anderson, et. al. An MboI restriction site within the major noncoding region was arbitrarily designated as the origin, and the base pairs are numbered sequentially proceeding clockwise (Figure 1). This published sequence is by convention used as a reference sequence in studies of human mtDNA variation, with polymorphisms usually indicated as differences from this reference sequence.

The mtDNA genome contains 37 genes, including 13 protein-coding genes, 2 ribosomal RNA (rRNA) genes, and 22 transfer RNA (tRNA) genes (Figure 2). The protein-coding genes include two ATP synthetase subunits, seven NADH dehydrogenase subunits, three cytochrome oxidase subunits, and cytochrome b. These proteins are all involved in electron transport and cellular respiration, the primary function of the mitochondria. All of the remaining several hundred proteins necessary for mitochondrial function (including those required for replication, transcription, and translation of mtDNA) are encoded in the nucleus, and hence must be imported from the cytoplasm.

The two complementary single DNA strands that comprise the double-stranded human mtDNA genome have an asymmetric distribution of guanine and thymine residues, and can be separated as heavy (H) and light (L) strands via ultracentrifugation. Most of the mtDNA genes are transcribed from the H strand, with only one protein-coding gene and eight tRNA genes transcribed from the L strand.

One of the most striking features of the human mtDNA genome is the extreme paucity of noncoding sequence. Only 7% does not encode proteins, rRNA, or tRNA, and the coding regions do not contain intervening sequences. By contrast, it is estimated that at least 90% of the nuclear DNA genome is noncoding, and large

and frequent intervening sequences are the rule for nuclear genes. Intergenic regions in mtDNA are usually less than 10bp in length, and for some of the genes polyadenylation of the mRNA transcript is required to form the termination codon. Approximately 90% of the noncoding mtDNA consists of the control region (displacement loop or D-loop), an 1,100 base pairs region that includes the H-strand origin of replication and origins of transcription for both strands. The L-strand origin of replication is located in a noncoding segment of 31 base pairs, located about 5,700 base pairs from the control region.

Cells typically have several hundred to several thousand copies of DNA, each with one to ten mtDNA molecules, whereas most nuclear genes exist in just a single paired complement per cell. Thus, for ancient remains where there may be very small amounts of surviving DNA that is highly degraded, the probability of obtaining a DNA type is greater for mtDNA. For some types of remains, such as telogen (shed) hairs or keratinized skin, nuclear DNA appears to be absent while mtDNA is still present.

Human mtDNA is strictly maternally inherited. Accordingly, the DNA in mitochondria is not present as pairs of genes, one maternal and one paternal, as is the nuclear DNA of chromosomes. Thus, despite the high number of copies, only a single sequence is found and recombinational events do not occur.

Spermatozoa have about 50 to 100 mitochondria in the midpiece, which provides the energy for the spermatozoa to swim. While the midpiece does penetrate the egg upon fertilization, it is not clear what subsequently happens to the paternal mitochondria. It may be that the paternal mitochondria are preferentially sequestered and destroyed. Prior to fertilization the maternal mitochondria increase to about 100,000 to 200,000 in the oocyte (egg). Maternal inheritance may simply reflect this greatly enhanced abundance of maternal mtDNA relative to paternal mtDNA in the egg. In addition, a bottleneck theory has been proposed in which only a few copies of the oocyte mtDNA are actually replicated, excluding the paternal copies from proliferating. Some studies have reported a low level of paternal mtDNA inheritance in Drosophila flies and in mice on the order of 0.01 to 0.001% per generation. However, these studies utilized interspecies crosses and could therefore reflect the peculiar nature of the hybrids. Regardless of the mechanism involved, no exception to maternal inheritance has ever been reported in

humans.

In the absence of new mutations, maternally-related individuals should have identical mtDNA types. For the purposes of individual identification, any maternal relative can therefore serve as a reference. The maternal line includes biological mothers, siblings, maternal aunts and uncles, children of sisters, and children in the case of an unidentified deceased female servicemember. The family relation can be distant. For example, fifth generation maternal relatives were used successfully to identify the remains of the last Tsar of Russia and the Royal family.

Critical for identification purposes, mtDNA is highly polymorphic, differing between most individuals.

E. Polymorphisms

The highly streamlined nature of the vertebrate mtDNA genome initially led to the expectation that it would be highly conserved evolutionarily. In the late 1970's it was postulated that mtDNA evolves, on average, 5-10 times more rapidly than nuclear DNA. Further studies in the laboratories of Dr. Douglas Wallace and Dr. Allan Wilson showed that there were high levels of mtDNA polymorphism within humans. These studies were based on analyses of restriction fragment length polymorphisms (RFLPs) across the entire human mtDNA genome. RFLP analyses have continued to be a valuable source of information concerning human mtDNA variation, evolution, and disease. However, for individual identification purposes, attention has focused instead on DNA sequence analysis of the control region which is known as the displacement loop or D-loop (Figures 3 and 4).

The control region is the major noncoding segment of human mtDNA, and is the most polymorphic segment as well. It is a region of greater than one thousand base pairs (16,024-576), which is instrumental in the regulation and initiation of synthesis of the gene products and replication of the mtDNA. Elsewhere, the mtDNA sequence is highly conserved. Consequently, the level of polymorphism in the coding region is about one third that of the

control region.

Sequence analysis of the entire control region demonstrates that variability is not distributed at random, but rather is concentrated in two hypervariable (HV) segments of about 400 base pairs each, with the first segment (HV1) having about twice as much variability as the second segment (HV2) (Figures 5 and 6). Virtually all subsequent studies of control region sequence variation have therefore analyzed either HV1 alone or both HV1 and HV2.

These studies have revealed much higher levels of mtDNA sequence variation than nuclear DNA sequence variation. Dr. Peter Gill of the Forensic Science Service, sequenced HV1 and HV2 from 100 British Caucasians and found the level of mtDNA nucleotide diversity to be 1.1%, while the level of nuclear DNA nucleotide diversity has been estimated to be at most about 0.11%; thus, the amount of mtDNA nucleotide diversity in the control region is about ten times the amount of nuclear DNA nucleotide diversity.

It is not clear why mtDNA evolves so rapidly. The rate of evolution is a function of the rate at which new mutations arise, and the rate at which mutations become fixed; there is reason to suspect that both of these factors are elevated for mtDNA. Highly mutagenic by-products of respiration, such as free radicals, are known to be present in mitochondria, and while some repair of mtDNA damage apparently does occur, it does not seem to be as efficient as DNA repair in the nucleus. The rate at which mutations become fixed in a population is inversely related to the effective size of the population, and because of the maternal inheritance of mtDNA, it has a smaller effective size than nuclear DNA. Still, further work is required to understand how mtDNA mutations arise and spread before a general theory relating mtDNA polymorphism, mutation rates, and evolutionary rates can be developed.

F. Ancient DNA

In order for DNA to be useful to the identification of individuals, the DNA must remain sufficiently intact and capable of extraction.

A new field of scientific endeavor has emerged which seeks to extract and amplify DNA from very old biological materials, such

as woolly mammoths, moas, sabre-toothed tigers, and mummies. The goals of these endeavors are diverse, but focused on evolutionary studies or population migrations and origins. This discipline is known as "molecular archaeology", "molecular anthropology" or "molecular paleontology", and those in this field study "ancient DNA".

Books have been written on the subject of ancient DNA and the community has organized itself. An informal newsletter is distributed and two international conferences on ancient DNA have been held (Nottingham, England, 1991; Washington, D.C. 1993). *Jurassic Park* was written by Michael Crichton from the machinations of the Extinct DNA Study Group at the University of California-Berkeley, a founding group of this community.

There is a background of important scientific work from which to draw in performing DNA analyses of ancient skeletal remains. The common ground between all scientific efforts in the field of ancient DNA and military skeletal remains identification of the Vietnam and Korean conflicts is the extraction of DNA information from samples in which the DNA is so extensively broken into small fragments and extensively damaged. The techniques and concerns of molecular anthropologists in the analysis of ancient DNA are applicable to military skeletal remains identification efforts.

G. Polymerase Chain Reaction

DNA from skeletal remains is degraded or broken down into very small fragments. This is particularly true of ancient DNA, where the average fragment size may be less than one hundred bases. Furthermore, most of the original DNA has been destroyed or washed away so that very low quantities of DNA are present and recovering DNA sequence information from old skeletal remains poses a significant technical challenge.

The polymerase chain reaction (PCR) is a method to amplify a target region of DNA. PCR is an enzymatic reaction resulting in the exponential production of copies of a given DNA segment, where the copies produced can be themselves copied. A million-fold increase in the number of copies of target DNA is often accomplished by PCR, permitting further analysis. Accordingly, PCR-based testing is exquisitely sensitive. This PCR amplification technique has revolutionized DNA testing and all molecular biology. PCR is of such significance that Dr. Kary

Mullis was awarded the 1993 Nobel Prize in Chemistry for its discovery. DNA testing of minimal and degraded DNA of ancient skeletal remains is made possible by the PCR method of amplification.

H. Military Skeletal Remains Identification

Substantial medical and/or dental records and nonDNA evidence exist to identify the majority of remains recovered from Southeast Asia using traditional methods; mtDNA testing would not be required in most cases where intact remains can be recovered. It is estimated by CMAOC that approximately 500 cases would require DNA tests, however, the number is unavoidably speculative. The AFDIL is currently resourced to perform mtDNA testing at a rate of ten cases per month. These skeletal remains involve cases with name associations and living family members are known in nearly all cases. The AFDIL will normally obtain the bone mtDNA sequences from the remains and then compare them to family reference blood specimens. Many families have already donated blood specimens in hopes that they may be useful for identification of their family member. The identification efforts using mtDNA have been highly successful and are anticipated to continue for the next five years.

The United Nations Command Military Armistice Commission (UNCMAC), representing the 16 nations that supported the South Korean government in the Korean conflict, has continued to press for the repatriation of the remains of UNC servicemembers since the termination of hostilities in 1953. During Operations Little Switch and Big Switch in 1953, a total of 3,748 U.S. POWs (out of a UNC total of 13,457) were repatriated. In 1954, the remains of 1,868 U.S. servicemen (out of a UNC total of 4,023) were repatriated in Operation Glory. Of these remains, 866 are declared unknown and 865 were buried in the National Memorial Cemetery of the Pacific (Punchbowl), Honolulu, Hawaii.

The total of over 8,100 U.S. servicemembers (out of a UNC total of over 10,200) includes those remains that have not been recovered, and were buried in known UNC cemeteries in North Korea, lost or buried at sea, and others who were unaccounted for with the body not recovered. Also included are 389 personnel (out of a UNC total of 2,233) about whom the Korean People's Army (KPA) and the Chinese People's Volunteers (CPV) should have knowledge. Information gathered from intelligence sources and

POW debriefings suggested that these were individuals who were possibly captured and died under KPA/CPV control.

In May 1990, for the first time since 1954, the KPA repatriated remains. Over 200 remains have been turned over in recent years. However, the change of power in North Korea has brought uncertainty into the future of remains repatriation.

The poor condition of remains recovered from Korea, the lack of records necessary to make an identification, and the inability to make joint recoveries have largely impeded CILHI's ability to identify these remains using traditional methods. In 1973, about 80% of the necessary medical and personnel records were destroyed in a fire at the National Personnel Records Center in St. Louis, Missouri. Existing records of that era provide a dearth of medical and dental information. Recent repatriations include many cases in which the midfacial portion of the skull is missing and thus the most useful anthropologic features for identification are absent. Accordingly, mtDNA testing is anticipated to be used in most Korean cases.

The Casualty Data section of CILHI has established preliminary figures for the number of recoveries through joint United States/KPA recovery operations and investigations in North Korea. There is sufficient information currently existing to establish a known location, where the recovery of remains is possible for 2,400 of 6,000 remains located in North Korea. These include 1,612 reported burials in former POW camps with known locations, 181 reported interments in known temporary cemeteries north of the demilitarized zone (DMZ), and 633 known aircraft loss sites north of the DMZ. In an additional 568 incidents, sufficient information is known which may lead to remains recovery. These include 535 reported POW camp burials with no known location, 1 reported interment in a cemetery with no location, and 32 aircraft losses without a fixed or general location. Most cases represent burials by American personnel during advances before reoccupation by the KPA. Accordingly, CILHI estimates the upper limit of the reasonable prospect of the number or individuals that may be recovered through joint investigations and recoveries is about 3,000.

I. NonDNA Identification Evidence

For Korean cases, nonDNA evidence, such as location of recovery

and forensic anthropologic and dental examinations, can provide identifying information, but due to the paucity of antemortem medical and dental records, these will serve primarily to limit the servicemembers considered for a potential mtDNA match. CILHI scientists estimate that they will be able to narrow the possible name associations to an average of 25 from the overall pool of the approximately 8,100 cases based on age, race, and dental characteristics in the vast majority of cases and in some cases to within five individuals.

J. Current Methodology

The process of mtDNA testing ancient remains is technically difficult and demanding, but AFDIL now has considerable experience. The first steps involve the preparation of the sample. A portion of bone (about two grams per extraction) is cleaned to prevent cross contamination and to remove any mineralization that inhibits mtDNA testing. The sample is pulverized; from the bone powder, DNA is extracted. A quick analysis of the quality and purity of the DNA recovered is performed at that stage. The region of the mtDNA to be analyzed is then amplified using PCR. Multiple copies of specific segments of the mtDNA that are to be analyzed are generated until sufficient material for the sequencing reaction is produced. Once this process is completed automated sequencing instrumentation can then read the exact sequence of the mtDNA molecule. Finally, the mtDNA sequence from the family reference specimen is compared to that of the skeletal samples. The amplification and sequencing scheme employs two overlapping sets of primer sets which allow for further confirmation of sequencing results (Figure 7).

The Task Force finds that identification of so-called ancient skeletal remains by a program of mtDNA testing is possible, particularly in association with other information. A few specimens may remain unresolved. Although contenders may emerge; at this time, mtDNA sequencing technology is the most appropriate technology.

II. FACTORS

DSB TOR: To evaluate factors that might influence the effectiveness of using mtDNA techniques. *[What factors might effect the utility of mtDNA testing for identification purposes, e.g. age of the skeletal remains, manner of interment, environmental temperature, acidity of soil, limitations of sample availability, commingling, etc? Do environmental factors affect the mtDNA sequence results obtained? How does somatic mutation and heteroplasmy affect mtDNA typing efforts? To what extent may the lack of family reference samples impede mtDNA typing efforts? What is the likely average number of bases obtained from Southeast Asian and Korean remains? If this data are not available, what size sampling would be sufficient for establishing reasonable estimates? How will these limitations of mtDNA sequence information available from Southeast Asian and Korean cases affect the ability to identify unassociated remains?]*

Many factors may affect the ability of the military to perform accurate and successful mtDNA identifications on the skeletal remains.

A. Skeletal Remains

Recovered remains from Southeast Asia vary drastically in their quantity and quality. Some skeletal remains are virtually complete, where others have completely disintegrated, dissolving in the acidic soil. The integrity of the remains is affected by the manner of interment, length of interment, environmental temperature, and the acidity of the soil. In some cases, a hot environment can preserve by permitting thorough desiccation; in some cases, total submersion can permit preservation. Anthropologists note that even within the same burial site remains may demonstrate great differences in their degree of preservation. Thus, small changes in the environment can cause substantially different rates of degeneration.

Bone samples from Southeast Asia have demonstrated that they harbor only small amounts of mtDNA, and that it is severely fragmented. The variance in the quality of mtDNA within the bone samples is significant and unpredictable. The quality of the bone and the ease of obtaining genetic information from the bone

varies between bones from the same individual. Skeletal remains of the last Russian Tsar, Nicholas II, tested 73 years after his death, yielded relatively abundant quantities of large mtDNA fragments despite their age. The determination of whether or not a sample of bone will yield mtDNA information is simply a matter of trial. A second sampling of bone may yield mtDNA when the first sample did not. Teeth may yield mtDNA when bone will not.

There are indications that extraction of mtDNA from skeletal remains from North Korea may be significantly easier and more productive than from skeletal remains from Southeast Asia. The climate of Korea is far cooler and drier than Southeast Asia. Visual inspection of the quality of remains from Korea reveal that they are in reasonable condition relative to remains recovered from Southeast Asia. The AFDIL experience with Korean remains suggests they may not be any more problematic than those from Southeast Asia.

Recent repatriation of Korean remains disclose an 80% rate of admixture of different remains or "commingling". This is often easily recognized when, for instance, two right femurs are submitted; other times commingling is more difficult to diagnose. Joint recovery operations, which could include the use of more experienced U.S recovery teams could substantially decrease the rate of commingling. Commingled remains may require multiple mtDNA samplings.

B. Family Reference Specimens

A lack of family reference samples may prohibit effective mtDNA identification. If necessary, family references could at least theoretically be obtained from exhumed remains of deceased family members. Occasionally, pre-mortem reference specimens such as locks of hair, paraffin-embedded material from prior biopsies, and neonatal bloodstains can be obtained to compare with the deceased individual. The ability of the military to find appropriate kindred is a factor in the success of mtDNA identifications. Unfortunately, the immediate kindred of servicemembers from the Vietnam and particularly the Korean conflicts are an aging population and hence the availability of family references is rapidly diminishing. There are indications

of interest by many families, suggesting they will cooperate by donating blood specimens (section V., Family Reference Database, p. 29).

C. Discriminatory Potential

The full discriminatory potential of the D-loop sequence is not yet known, since the region has been sequenced from relatively few people. Samples of sequences from large Caucasian populations have shown no sequence occurring with a frequency of greater than 3%. Sequences occurring with higher frequency would have almost certainly been detected. More frequent sequences have been found within some native African tribal populations. Coincidental matches between unrelated members are therefore likely to arise only a few percent of the time.

The amount of information obtained from family members will quickly allow greater precision to be attached to statements of discriminatory power, as will the sequences from remains. The latter sequences may, however, be of less than the complete hypervariable regions. Once 500 sequences are available from a population, a particular sequence can either be assigned its observed frequency or said (with over 99% confidence) to have a frequency of less than 1% if it is not observed in the sample. However, sufficient data are available to conclude that a high discriminatory power is achievable by current methods and can be used for reasonable estimates of mtDNA sequence population frequencies.

Although, the DNA information from hypervariable regions I and II will permit great discrimination, the sequence information actually obtained from ancient samples may be limited, reducing the discriminatory power. The full mtDNA sequence may not be obtained due to amplification failure by one or more sets of primers, fading of terminal sequence signal, or from internal sequence ambiguities. The AFDIL is very conservative in determination of sequence or "base calling" and hence may not call an ambiguous, though informative base. Accordingly, the lack of full sequence information may result in limitation of the full potential discriminatory power in a given case.

Another potential limiting factor is the occurrence of multiple servicemembers of the same lineage. For example, brothers and maternal first cousins will have the identical mtDNA sequence.

The rate of maternal kindred relations among servicemembers is not known. The rate for servicemembers serving in Korea is likely somewhat higher than current frequencies, due to the larger families during that time period. Regardless, families themselves should typically have some idea of the existence of other kindred servicemembers who might potentially confound an identification.

D. Mutations

The mtDNA sequence is identical throughout the body unless a somatic mutation arises. Investigations by Monnat and Reay demonstrated that the mtDNA control region does not differ among the various tissues of the body. An investigation of 83 retinal cell clones, resulting in 32,000 bases of DNA sequence (from both the D-loop and coding regions), revealed only one mutation (in a tRNA gene).

More extensive age-related changes have however been found in other post-mitotic cells (Annex F). Bone and teeth have not yet been extensively studied. While somatic and germline mutations should be vigilantly looked for, they should not vitiate mtDNA identification in the majority of cases.

The polymorphic nature of mtDNA, especially within the control region, is evidence of mutation, and surveys support a widely held belief that mutations occur at a higher rate in the mitochondrion than in the nucleus. Whatever the mechanism for this elevated rate, it makes the mitochondrion useful for human identification. The disadvantage with high mutation rates is that sequences within the same maternal lineage, especially when separated by several generations, may exhibit occasional differences.

Although no comprehensive studies have been performed, evolutionary studies have estimated that the average fixed mutation rate for the mtDNA control region is one nucleotide difference per every 300 generations, or one difference every 6,000 years. Consequently, one would not expect to observe many examples of nucleotide differences between maternal relatives.

The AFDIL has observed at least two examples of fixed sequence differences between mother and child in approximately 30 maternal relative comparisons evaluated by AFDIL during routine casework.

In one case, there were two nucleotide differences between a mother and her daughter, but not in her two sons. The maternal relationship was verified using 12 DNA markers, including 8 RFLP loci.

The AFDIL surveyed 46 mother-child comparisons in a controlled study of family specimens from the Centre d'Etude du Polymorphisme Humaine (CEPH), showing no sequence differences. The AFDIL is completing an extensive study across three hundred generations which will help to further define the fixed mutation rate. Preliminary data suggest that the observed rate will be closer to one nucleotide difference every 50 generations. Dr. Mark Stoneking has found similar results in a recent study of the resident population of Tristan da Cunha, a small island off the Atlantic coast of South Africa.

This higher than expected rate of mutation may be explained by so-called "hot spots" which have a higher mutation rate than other regions even within HV1 and HV2.

E. Heteroplasmy

The condition in which only a single mtDNA sequence (albeit in many copies) is present in a cell is termed homoplasmy. The possibility exists that a subpopulation of mitochondria could harbor a different mtDNA sequence due to mutation or paternal contribution; this is termed heteroplasmy. Heteroplasmy is most often seen in disease states, in which a mtDNA sequence is defective (e.g. deletions, duplications) and hence at a competitive disadvantage. Stable heteroplasmy in humans although rare has been described.

Measuring the fixed mutation rate does not take into account the rate of heteroplasmy during the process of fixing a particular mutation. It is not known whether most mutations are manifested through many generations via heteroplasmy or if they are fixed during a single generation. Examples of each have been observed (i.e., the Russian Tzar was reported to be heteroplasmic, despite the lack of heteroplasmy in other family members). The consensus of the Task Force Committee is the level of heteroplasmy within the control region of humans is generally low.

F. Environmental Damage

Environmental changes do not alter the mtDNA sequence. Long term exposure has resulted in depurination in which Guanines result in Adenines after PCR, but this has been observed only in material thousands of years old. This does not appear to be a problem with the identification of skeletal remains from the Vietnam and Korean conflicts. If environmental damage could be assessed in a rapid and cost effective manner, then that information could be used as indication of the quality of the data obtained or the need for repair of the DNA before amplification.

G. Contamination

The mtDNA present in ancient skeletal remains is minimal. The mtDNA that had existed has largely been enzymatically or chemically broken down into small fragments and much has been leached away. The AFDIL reports that only a few dozen target mtDNA fragments are obtained by their extraction techniques on skeletal remains from Vietnam. Therefore, the methods of mtDNA typing must be very sensitive. The enhanced sensitivity required for this type of testing (PCR-based) also means that it is capable of detecting trace contaminants.

MtDNA contaminants can come from a variety of sources. Contamination through proximity to other remains and through specimen handling in the field is not thought to be a significant source, especially when a laboratory removes the outer surfaces of bone samples prior to mtDNA extraction. Shed skin or exhaled droplets from individuals performing the extraction is another potential source of contaminating mtDNA. The mtDNA sequence of laboratory personnel must be determined so that their sequence types will be recognized as a possible contaminant. The most serious contamination concern is cross-contamination by PCR product within a laboratory. Millions of copies of the mtDNA target are generated during the PCR process. If a minute quantity of this product is allowed to contaminate another mtDNA extract, the contaminating product may mask the true mtDNA type of the extract. When all appropriate precautions are taken, random sporadic contaminants, not attributable to any known source, may be encountered.

Controlling contamination is vital to the success of a PCR-based mtDNA testing program, especially when targeting old skeletal remains. A range of precautionary measures tailored to the specific laboratory are necessary and when anticipating high

volumes of mtDNA testing, further measures must be considered to control the higher potential contaminant burden. Some examples of possible useful precautions are discussed below.

Control of contamination must begin with the physical separation of DNA extraction and amplification set-up activities (pre-PCR) from PCR product analysis (post-PCR). In addition, all steps in the extraction and amplification set-up procedures should be performed in a hood. Areas for extracting low levels of mtDNA from skeletal remains should be separated from high level extraction areas. Anterooms or vestibules may be placed at the entrance of each laboratory where mtDNA testing is being performed to help prevent the transfer of PCR product from one laboratory to another. When possible, mixing of the air circulation between pre-PCR and post-PCR areas should be minimized.

Dedicated equipment should be used for amplification setup and in PCR product areas. Positive displacement or plugged tip pipettors should be used for aliquoting samples and PCR reagents. Use of laboratory coats and gloves are essential. Laboratory coats should be disposable or dedicated to the areas where mtDNA PCR product is being handled. Care should be taken when wearing disposable gloves not to touch any surface which may contain a contaminant such as the surface of the skin, eye glasses, clothing or even a cleaned bench-top. As a common practice, before handling evidence or items which come in contact with evidence, always change gloves or wipe gloves with bleach, allowing the gloves to air dry.

General cleaning practices are important for controlling contamination. The universal cleaning agent for PCR contamination is 10% commercial bleach (7 mM sodium hypochlorite). The bench-tops, hoods, and any surface which comes in contact with the evidence or DNA extract should be washed with bleach frequently. In addition, the floors of each laboratory should be periodically washed with bleach.

To illustrate the impact of proper laboratory design and practices, a representative number of cases were evaluated from the AFDIL laboratory. One set of data represented the conditions under which cases were processed in 1991. The second set represented cases performed in 1994. The number of times PCR product was observed in an extraction reagent blank control or a

PCR reagent blank control was counted for each case. For seven cases processed in 1991, 15% of the extraction reagent blanks and 7% of the PCR reagent blanks showed PCR product following amplification. For five cases processed in 1994, only 8% of the extraction reagent blanks and 3% of the PCR reagent blanks showed PCR product following amplification. These numbers indicate how sensitive mtDNA testing is to contamination, but also illustrate that significant progress can be made to limit the occurrence of contamination with improvements to laboratory design and practices.

Appropriate controls and measures to detect contamination are imperative. Both extraction reagent blank and PCR reagent negative controls should be run in every case to detect the occurrence of contamination. Documentation of contamination will allow for review of the major sources and will assist in developing a comprehensive quality control program for controlling PCR product contamination.

Where a contaminant is detected, testing should be repeated, if possible, unless the sequence from the skeletal remains is otherwise determined to be reliable. There may be instances when contamination prevents the reporting of results.

In general, other than mtDNA testing of skeletal remains, it is uncommon to find amplification product in a reagent blank during any PCR-based testing. Accordingly, the AFDIL rarely finds a positive amplification product in the extraction reagent blank for a whole blood maternal reference.

When contamination does occur, it will typically result in an apparent nonmatch (false exclusion), not a spurious match (false inclusion). Therefore, all exclusions should be carefully scrutinized. The AFDIL procedures preclude a false match due to contamination by a mtDNA blood reference. All blood specimens are tested in a different area of the laboratory. Moreover, the whole blood maternal reference is usually processed after the skeletal remains, and in most cases after the results generated from the skeletal remains have been reported. Only during databasing operations will there be the occasional situation when reference results are generated prior to the skeletal remains results.

An important quality check of the skeletal remains sequence is the independent extraction and testing of multiple bones. Multiple bones are independently extracted in most cases by the AFDIL, unless only a single specimen is available for testing. In addition, the primer sets used for amplification are overlapping, providing further confirmation of the authenticity of the skeletal sequence.

Given the inevitable random contamination inherent in mtDNA testing, redundancy, when possible, at the level of source material is key. Cautionary statements are important where replicate testing cannot be performed.

H. PCR Amplification Ambiguities

Taq polymerase, an enzyme used in current PCR amplification reactions, is known to occasionally and randomly misincorporate erroneous nucleotide bases at a rate of approximately one to ten in 10,000. Generally this is not problematic, because correct sequences will overwhelmingly predominate. However, if the starting target sequence consists of only a few copies, then the chance of a false result from misincorporation during the first few rounds of thermal cycling becomes a possibility.

Nonspecific priming may occur, particularly where the starting conditions are not optimal. This may result in ambiguous or errant sequence when the starting DNA concentration is very low.

During PCR amplification, the polymerase enzyme may stop due to template fragmentation or damage. The partially extended sequence may then anneal to another template fragment in the next cycle. In fact, several fragments may be assembled in this process to recreate the original full length sequence. This process, known as "jumping PCR", may complicate the interpretation of diploid sequences. In mtDNA, where only one sequence is present, "jumping PCR" may have advantages. However, it may also produce errors from incorrect assembly of fragments so small that they have lost their specificity. Duplication of testing results or overlapping sequence data will allow for interpretation of these occurrences.

I. Sequencing Ambiguities

The evaluation of DNA sequencing data is extremely tedious and time consuming. Evaluation of data using manual sequencing methods may result in transcription and reading errors. Automated sequencing methods will minimize reading errors. Nonetheless, sequencing errors do occur in automated analysis; rates have been published for current instrumentation. Most instrumental error can be avoided by limiting the information read to shorter lengths, because the vast majority of errors occur near the end of a sequencing run as the resolution and strength of the base signals diminish. The automated sequencing instrument used by the AFDIL has an error rate of approximately 1 to 2% for the size templates analyzed (200-300 base pairs). This error rate increases to greater than 10% when longer templates (greater than 450 base pairs) are analyzed using current

conventional sequencers. The error rate does not represent errors in the sequence reactions, but instead represent errors in the ability of the instrument software to make an accurate base call.

In large scale sequencing operations, where a maximal quantity of sequencing information is emphasized, some errors can be tolerated. This is not the case in the forensic context. Visual evaluation of the data is mandatory to prevent read errors. Moreover, it is standard practice to confirm sequences, usually by checking complementarity of the reverse strand sequence. With the advent of new sequencing technologies, it may be possible to confirm sequences by a second technology which may not have a tendency toward the same systematic errors. At least two qualified individuals must independently evaluate the sequence in order to ensure that the results accurately represent the data, and to catch transcription errors. It should be recognized that redundancy in casework also provides an opportunity to catch errors; a case involves multiple bones and at least two family blood reference specimens where available. These procedures can virtually eliminate sequencing errors.

J. Casework Experience

Since October of 1994 when the AFDIL began full production of ten cases per month, it successfully obtained sequence information from Southeast Asia, Korean, and World War II cases representing 37 individuals. The AFDIL has achieved a success rate for obtaining DNA sequence information of greater than 95% from CILHI casework since late 1992, even for retested cases from 1991 Southeast Asia casework when the original success rate was only 40%. This does not mean that the AFDIL obtains sequence information from every bone, nor does it mean that the AFDIL obtains full sequence information in every case. The AFDIL obtains sequence results from little more than half of the bones it tests.

CILHI casework is not routine in the sense that, in virtually every case problems are encountered that require retesting, often using modified test conditions. In the first six months of production mtDNA sequencing operations, an average of 3.7 bones were tested per case, 42% of extractions had to be repeated, 37 amplification reactions were required (2.5 times the minimum), and 55.5 sequencing reactions were required (twice the minimum)

[these numbers do not include the number of controls run per case]. Although these numbers are high, they are expected efficiencies for "ancient DNA" testing. Thus, despite difficulties, the AFDIL is now able to obtain mtDNA information from the vast majority of cases. The results from recent Korean cases suggest that Korean war remains may not be any more problematic than those from Southeast Asia, presumably due to better preservation of the remains.

MtDNA identifications from ancient remains is a scientific tour de force, at the cutting edge of today's capabilities. As single DNA molecules can be detected, some degree of cross-contamination is inevitable, especially in a scaled up production facility. The Task Force nevertheless concludes that with appropriate control measures (redundant testing and meticulous lab hygiene) these problems are surmountable, and a good record has been presented in the currently on-going work.

The Task Force finds that the present probability of coincidental matches between mtDNA control region sequences is no more than a few percent. Once sequences from 500 members of a population have been determined, precise statements about the chance of a false association of a set of remains with a family will be able to be made. Published data may be of value, but samples will be needed from Southeast Asian populations. **The precision is expected to suffice in the vast majority of cases, given other non-DNA evidence, to effect the mtDNA identification of unassociated Korean remains. It will not be possible to identify every bone. A great deal can be done with anatomical and historical evidence alone.**

The Task Force finds that control of contamination is essential to PCR-based laboratory testing. Some contamination is unavoidable, particularly in mtDNA testing of ancient remains, but it does not preclude reliable casework testing where redundancy, good laboratory practices, and appropriate cautionary language are used and constant oversight is maintained.

The Task Force finds that casework experience demonstrates capability to type Korean skeletal remains.

III. RELIABILITY

DSB TOR: To evaluate current and emerging scientific evidence concerning the reliability of the techniques when compared with other current and evolving methodologies. *[Is the current MtDNA sequencing identification method reliable? Are quality assurance efforts satisfactory? What further measures would enhance reliability? What studies are necessary to validate a new DNA typing methodology on ancient skeletal remains? What continuing scientific oversight or advisory body should monitor these DNA identification efforts? What quality assurance mechanisms or measures should be implemented?]*

The scientific community believes that current mtDNA sequence identification technology is reliable. Forensic laboratories in the United States and Great Britain have begun to use it in casework, recognizing they must be able to defend the technology in court, if needed.

The component technologies of DNA extraction, PCR amplification, and DNA sequencing are all validated, having been used in research and service orientated molecular biology laboratories throughout the world for many years. Furthermore, their application to so-called ancient DNA is well established in the scientific literature.

The studies generally accepted by the forensic community to validate a new DNA testing technology have been articulated as guidelines by the Technical Working Group on DNA Analysis Methods (TWGDAM), sponsored by the Federal Bureau of Investigation (FBI). Such studies include the following: optimized standard source studies (such as studies on fresh body tissues and fluids, stored tissues and fluids, and samples from different tissues from the same individual); variance analysis (studies on measurement precision from known DNA controls); population studies (studies of population frequency distribution in different racial/ethnic groups); preservation studies (studies on tissues and fluids as would be typically found at a scene investigation, eg. dried stains); time/temperature studies (studies on samples incubated at various time and temperature); environmental exposure studies (studies on effects of various commonly encountered substances); evidentiary source studies (studies on nonprobative evidence from typical crime scenes); non-human studies (studies on tissues and fluids from common non-human sources); on-site evaluation

(studies to evaluate methodology transfer to a working forensic laboratory setting). The FBI is completing their validation studies on mtDNA typing in preparation for casework.

The TWGDAM Quality Assurance Guidelines address basic considerations in DNA analysis--such as analyst training, reagent quality control, evidence handling, analytical procedures, proficiency testing requirements, and method validation. Requirements for standard cell line positive controls and extraction controls are included. Requirements specific to RFLP analysis (such as precision of fragment size measurements) and PCR analysis (such as negative reagent blanks and separation of pre and post amplification areas) are also covered. TWGDAM has a subcommittee devoted to the forensic use of mtDNA sequencing and has recently adopted modifications to their guidelines for mtDNA sequence analysis in forensic laboratories. However, they do not concern themselves with application to ancient DNA.

It is not generally appropriate or possible to exhume remains of known individuals which have been buried for 30 years to validate the utility of a DNA typing system. However, indirect evidence of the validity can be found in the contexts presented elsewhere in this text.

Corroboration will ensue from the internal consistency of the findings. In some cases, corroborative evidence for or against an identification may appear from other documentary or physical evidence acquired after the DNA tests have been concluded.

Evidence of the efficacy of skeletal remains identification by mtDNA sequencing can be found in the identification efforts performed to date by the AFDIL. Repeatedly, bones from the same case yield the exact same sequence. The sequences obtained from the bone samples have matched those of the putative family members. Often these sequences are unique, never having been seen before. The AFDIL processes the skeletal specimens before testing the family reference(s), eliminating the possibilities of bias or cross contamination. Controls are run with all cases and the results are always checked against the mtDNA sequence of the staff processing the specimens. In one case, an exclusion was found by mtDNA testing; when CILHI received the test results, they suggested a second name association which was then confirmed by a match of the case sequence to that of the second putative family. The AFDIL duplicates the mtDNA sequencing casework of

new personnel as they begin actual casework--this data has not revealed any discrepancies to date.

The quality assurance requirements for adequate mtDNA testing of ancient DNA are demanding. The molecular anthropology community has reiterated many of the concerns in this area, particularly that of contamination and adequacy of controls.

Proficiency surveys do not yet exist specifically for mtDNA sequence identifications, although the AFDIL personnel have been submitting mtDNA sequence data to the College of American Pathologists (CAP) and Cellmark proficiency surveys. However, interlaboratory exchanges have been initiated by TWGDAM members. Data, to date, demonstrate the reliability of this method. The National Institute of Standards and Technology (NIST) is currently developing mtDNA reference material.

The greatest single potential for a mistaken DNA sequence result is from contamination. Some contamination events during laboratory testing of ancient remains are inevitable. However, contamination can be minimized through proper laboratory design and sample handling procedures and can be detected through the use of appropriate controls. The molecular anthropology community advocates the use of multiple negative controls, including a DNA carrier control. TWGDAM recommends negative extraction and amplification controls as well as a known positive amplification control.

The discrepancy case, "X-6", between the AFDIL and the outside laboratory, reinforces the critical importance of quality assurance measures. The AFDIL's ability to blindly replicate their testing results on the case two subsequent times and independent confirmation of AFDIL's results by the British FSS, demonstrate the reliability of results when proper quality assurance measures are employed.

Ambiguities and errors can occur during amplification and sequencing. Nonspecific priming can be reduced through well designed primer sets used under optimized conditions. Sequences should be visually checked and interpreted cautiously, in accordance with appropriate protocols. Sequences should be confirmed. Sequencing results should be independently verified by a second qualified analyst (double reading).

Redundant specimens or replicate extractions should be tested, where possible, to ensure that a given finding is not a chance result of a contaminant. Where redundancy is not possible, cautionary language should be used in the reporting of the results.

A principle of forensic testing that is applied in the military's urine drug testing program is that, after a test result is obtained, it should be confirmed by a second test employing another method. There is not a second method by which current mtDNA testing can be confirmed. However, new technologies are being developed which, when they become available, could have application in testing skeletal remains.

Mutations in the mitochondrial sequence are not common, but do occur. Mutations raise the possibility of a false exclusion. A single base change should therefore be considered "indeterminate", rather than an exclusion or an inclusion. The AFDIL methodology of requiring two (2) family references has already proven to be useful to resolve mutational events.

Continuing scientific oversight is a significant part of the military's quality assurance program. The AFIP is formally reviewed twice a year through a Scientific Advisory Board (SAB). The SAB reports are distributed to the AFIP Director and its Board of Governors, which includes the Army Surgeon General, the executive agent for the AFIP, the Surgeon Generals of the Navy and the Air Force, the Public Health Service, the Medical Director of the Veteran's Administration, and the ASD(HA), who has policy oversight of the AFIP. One SAB member is dedicated to review of the Department of Defense DNA Registry. Currently, this Board member is Maimon Cohen, Ph.D., Chairman of the Division of Genetics, University of Maryland Medical School.

The ASD(HA) has ultimate responsibility for oversight of quality assurance of all identifications of human remains performed on behalf of the military services. A plan has been published of the minimum standards for military or military-contract laboratories to follow in performing mtDNA sequence analysis of ancient skeletal remains (Annex C). These standards were submitted to the American College of Medical Genetics, the College of American Pathologists, TWGDAM, the American Society of Crime Laboratory Directors and the Department of Defense Clinical Laboratory Improvement Office for review and comment before

promulgation.

The ASD(HA) Quality Assurance Program calls for laboratory accreditation by an outside professional organization, specified additional mtDNA-specific standards, and oversight by an independent board. The Quality Assurance Oversight Committee, composed of civilian technical consultants, will perform reviews of casework, reports of inspections, proficiency test results, and accreditation materials.

There is currently no framework for mandatory proficiency testing and accreditation of independent laboratories, a circumstance beyond the mandate of this Task Force. Substantial voluntary efforts are nevertheless underway, with important leadership from the TWGDAM.

The Task Force finds that appropriate measures must be taken to prevent and control possible contamination in the testing laboratory.

The Task Force finds that current AFDIL protocols, if diligently performed, are capable of generating quality mtDNA sequence identifications.

The Task Force finds that adequate quality assurance requires accreditation and an oversight board.

The Task Force finds that the program for mtDNA sequencing quality assurance promulgated by the Assistant Secretary of Defense (Health Affairs) for identification of ancient remains is adequate and responsive.

IV. OTHER DNA TARGETS

DSB TOR: To evaluate the possibility of obtaining useful mtDNA information from skeletal remains through mtDNA outside the control region (D-Loop) or possibly nuclear DNA. *[What other DNA typing could potentially be used to augment current DNA identification efforts of ancient skeletal remains, e.g. noncontrol region mtDNA sequence, multicopy nuclear DNA regions, etc.? Should the government expend funds to investigate other DNA typing possibilities?]*

The discriminatory potential of sequencing the hypervariable regions of the mtDNA control region is powerful, with most sequences in current samples being unique. In particular cases, the sequence found to match between remains and putative family members may not be found in general population samples. In that case, the probability of a coincidental match will be less than one percent. Stronger statements may not be possible because the proximity of the variable sites in the control region may mean that they are correlated.

Polymorphisms in the mitochondrial genome at sites outside the control region may be used to enhance the discriminatory power of current mtDNA identifications. Polymorphisms at other sites have been well-described and may have significant additional discriminatory power.

The discriminatory power could particularly be enhanced if nuclear DNA could be recovered from ancient skeletal remains. Although to date, typable nuclear DNA has not been obtained from remains from Vietnam and Korea, the fact that mtDNA itself can be recovered lends to the theoretic possibility of recovering nuclear DNA. In fact, short tandem repeat (STR) analysis of nuclear DNA was performed with limited success from the 73 year old remains of Tsar Nicholas II. However, in that case the mtDNA was more intact than in remains recovered from Vietnam and Korea, presumably due to the favorable environmental conditions of Russia.

The primary reason given for the ability to recover mtDNA but not nuclear DNA is the hundreds to thousands fold increase in copy number. The circular nature of the mtDNA, its particular sequence, the presence of a mitochondrial membrane, and the absence of intra-organelle degradative enzymes may confer some

additional longevity to mtDNA; however, there is no experimental data to argue for or against such theoretical possibilities.

Some nuclear sequences are present in many copies within the same individual. These multicopy sequences may be particularly likely to be recovered. Ribosomal DNA repetitive regions may also be of significance.

If nuclear DNA testing is possible, then the children and spouse of deceased servicemen could donate a reference DNA sample; where using mtDNA, they cannot usefully today. Thus, a greater pool of potential volunteers for family reference DNA specimen donations could be created. Given the aging population of reference family members for mtDNA markers, this is of very important consequence. Furthermore, nuclear DNA markers, if they could be obtained, would also provide a greater discriminatory power.

The Task Force recommends that the AFDIL investigate the potential to perform DNA typing outside the mtDNA control region.

V. FAMILY REFERENCE DATABASE

DSB TOR: To determine the degree to which mtDNA matching could be accomplished with reference donors (family members of up to 8,100 unaccounted for in the case of Korea, adequacy for discrimination of individual from such a database, and what alternatives exist if such family donors are deceased. *[Can the remains without a name association from Southeast Asia and/or Korea be identified using a database of family reference DNA typing information? To what extent could DNA be used in isolation and in combination with other identification evidence to individually identify servicemembers from a panel of over 8100 U.S. family DNA types? To what extent are identification efforts hampered by the lack of family reference specimens from all families? To what extent are identification efforts hampered by the presence of foreign nationals among the unidentified remains? What reference specimens should be sought from families, e.g. from 2 family members, nuclear DNA relations, consanguinity within 3 generations? Could reference specimens be obtained from the exhumed bodies of family members?]*

Since mtDNA is maternally inherited, maternal kindred can be used as sources of reference material. Specifically, references include the biological mother, siblings, maternal grandmother, maternal aunts and uncles, children of sisters, or children of deceased female servicemembers. Unless nuclear DNA can be used in these cases (see above section IV., Other DNA Targets, p. 27), children of deceased servicemen are not useful for reference specimens.

Since mtDNA is not inherited from both mother and father and does not undergo recombination in the same way that nuclear chromosome pairs do, an exact match of mother to son or daughter is expected. A consequence of finding an exact match in the kindred, is that relatives far removed generationally from the deceased member, may be an appropriate reference. For example, Tsar Nicholas II was identified through his great grandniece.

Unlike the situation with servicemembers from Southeast Asia, the U.S. military does not have a comprehensive database for the approximately 8,100 families from the Korean conflict. A substantial effort will be needed to identify and contact eligible family members for mtDNA reference specimen collections. Many families may come forth quickly upon news of an undertaking

to recover and identify remains of soldiers from Korea. Yet these interested family members are now an aging population and may no longer be available to provide this reference.

The consensus of the military staff working this endeavor is the majority of the families are interested in resolution of these cases. The issue of genetic privacy may be raised in this context in the same way that it is raised with the current DNA collections program for Active and Reserve Component servicemembers. It will require sensitivity to the issue, adoption of safeguards, and recognition of the need for educational of the prospective donors.

The POW/MIA Affairs Division, U.S. Total Army Personnel Command, has developed the framework for an outreach program to identify and contact persons authorized to make a decision on the disposition of remains and acquire reference blood specimens from the families of Americans whose remains were not recovered from the Vietnam, Korean, and Cold War incidents (Annex G). It is estimated that half of the families will be contactable and provide blood reference specimens. The outreach program will: 1) identify and contact persons authorized to decide on the disposition of remains; 2) identify and contact prospective DNA donors; and 3) raise public awareness of government efforts to repatriate, and identify the remains of Americans lost in Southeast Asia, Korea, and the Cold War. A military-supervised, contractor-operated operation is contemplated. The contractor will proactively solicit persons authorized to decide on the disposition of remains and eligible mtDNA donors for 2 years or when 70% of the donors/persons authorized to direct disposition of remains are contacted. The cost of the outreach program will include family notification, a computerized database, blood collections, travel, and staffing.

Where a family reference cannot be obtained from living relatives other theoretical possibilities exist. Possibilities, from either deceased kindred or from the deceased servicemember himself, include biopsy samples maintained in a hospital repository, saved deciduous teeth, and locks of baby hair. Reference samples could conceivably be obtained from exhumed familial remains. A mtDNA match has been performed by the AFDIL between skeletal remains from Vietnam and baby hair from 1927. The AFIP itself may have slides and paraffin-embedded specimens stored from servicemembers who served in the Korea conflict.

As with all methods of identification, comparison data are necessary. Only where family sequences for the prospective remains are known or the ambiguities accounted for, can a match provide certainty of identification. Even a complete family reference database will not address the possibility that given remains may be those of other UNC or KPA troops. The power of a mtDNA sequence match can be stated with increasing confidence as the database expands. Databases of 500 population group (ie. Caucasian, Black, Hispanic, Vietnamese, Korean) are desirable. Other evidence of identification will also confer increasing confidence in a mtDNA sequence match. A given sequence may provide certain identification where all sequences are known within a sufficiently characterized subgroup such as the case of those personnel known to be buried at a particular site. With each identification, the overall pool of potential reference sequences will decrease; thus the power of individuation will increase with completion of the database.

The database of family reference sequence information must be largely established before identification of the unassociated remains can proceed, unless non-DNA evidence allows a restriction on the possible families in a specific case. Moreover, the collection of family references is time sensitive; there is some urgency to contact appropriate family members as many of them are now quite elderly. For each family, it will be necessary to identify the members most likely to allow identification, paying attention to the possibility of the future use of nuclear DNA information.

The Task Force finds that, with a reasonable effort, a sufficient proportion of families are expected to provide mtDNA samples to allow identification of many of unassociated remains from Korea through mtDNA testing, and to attach meaningful probability statements in those cases.

The Task Force recommends consideration be given to the collection of DNA reference specimens from maternal and paternal family members in case future technology permits nuclear DNA testing. Collections from nonmaternal kindred members (e.g. children) should be made with full disclosure of realistic expectations.

VI. STATISTICAL DATABASE

DSB TOR: To recommend a statistical database to be used in calculating the statistical information. *[What database(s) should be used in casework for determination of statistical inferences? What are the minimal technical requirements to be met before inclusion in the database? Should databases for racial and ethnic populations be constructed and if so which ones? Should reports include the frequency estimates for indigenous Vietnamese populations as well as for Americans? Is the counting method the best and only form of statistical inference for mtDNA? What degree of independence between polymorphisms exists within and without the mtDNA control region?]*

If a remains are from a closed population group or subgroup, the mtDNA sequence and ancillary data need only distinguish among the population pool without reference to statistical inference. If mtDNA is to be used for identification in situations other than those of closed populations, then estimates are needed of frequencies of specific variants to assess the weight attached to matches of sequences from difference sources. Account must be taken of mutation rates to calculate the likelihood of differing sequences being from the same maternal lineage.

The full discriminatory power of mitochondrial sequencing for identification purposes can be determined only through a database larger than currently exists. However, preliminary evidence from combined statistical bases indicate that a very high discriminatory value can be achieved.

Typing both the skeletal remains and family blood references will result in the largest known DNA sequence databases. These databases will provide the means for attaching probability statements to identifications in situations other than those arising from the Korean Conflict. Further studies are needed to establish the extent to which other mitochondrial and nuclear information may be combined with mitochondrial control region information. A database of 500 individuals per population group will be large enough to ensure (within a 99% confidence limit) that a variant with a frequency of 1% will be seen.

Before sequences should be entered into a common database, certain criteria should be met. These criteria include confirmed sequence of at least a certain length, using a given

nomenclature, and with certain demographic data attached.

Moreover, to allow for the possibility of non-U.S. remains being typed, and to allow for the genetic differences within the U.S. population, sequences should be collected from the major groups; Caucasian, African-American, Hispanic and Asian. As data accumulate, it will become evident if there are substantial mitochondrial differences among different Asian countries, or within the racial groups.

The Task Force finds that the existing databases, and those anticipated from the family collections, provide an adequate basis for the current mtDNA sequencing efforts.

The Task Force recommends an expansion of the current AFDIL database.

VII. LARGE-SCALE OPERATIONS

DSB TOR: To ascertain what effects a large volume of remains could have on the identification process utilizing DNA technologies. *[What are the consequences and potential problems in scaling up to high volume mtDNA typing operations, e.g. risks and preventive measures of contamination, quality of typing services, resource implications, etc.? Are sufficient qualified personnel available for hire?]*

If the military decides to perform mtDNA testing on skeletal remains from the Korean conflict, an order of magnitude in the scale-up of operations will be required. However, it may not be advisable to greatly scale-up operations. A more sensible approach may be to spread the casework over a longer period of time.

The scaling up of operations is not necessarily a simple unit expansion of current efforts. Fixed costs and fixed assets must be differentiated from incremental costs and resources. Infrastructure support is often a discontinuous step function. Economies of scale may come into play. Larger operations may permit restructuring and work to be performed in ways not possible by smaller operations.

The AFDIL assessment is that scale-up for a higher volume operation is possible within the existing AFDIL facility. With existing facilities and current technologies, the AFDIL expects to achieve a 25% increase in efficiency by FY 96 and a 50% increase by FY 97, after further experience is gained by new personnel and through structural changes in the way in which specimens are processed. More fully implementing and improving the computerization of operations would significantly impact the throughput. Future technologies could potentially increase throughput by several fold. Assuming the administrative space shortage could be resolved, the doubling of laboratory personnel could be accommodated within the existing AFDIL facility; and an even larger staff, if a double shift mode is used. Hence, the military could take advantage of existing facilities to achieve the needed throughput.

A concern for a high volume operation is the increased potential for cross-contamination within the laboratory. Existing facilities are capable of meeting this need.

The forensic DNA community is small. Nationally, the number of academic programs conferring degrees in the forensic sciences is a mere handful. However, only approximately one half of the States have DNA testing capabilities and the turnover among current forensic DNA Analyst positions is low. Consequently, a significant although not large group of forensic laboratorians exists from which to draw. The molecular biology community is much larger and represents an alternative pool of personnel which could be tapped for large scale operations. Medical Technologists are specially educated, trained, and certified in laboratory techniques, practices, principles and quality assurance and represent a further pool of potential analysts.

Each of these three major disciplines has its strengths and weaknesses in regard to the backgrounds for performing this testing. Most analysts will not possess great expertise in mitochondrial DNA or ancient DNA analysis at the time of hire. On-the-job training would be anticipated. Nonetheless, AFDIL has not had great difficulty in recruiting high quality staff with significant relevant backgrounds and capable of being trained in the particular discipline. Shortages of qualified personnel should not prevent scale-up efforts, although some hiring delays, significant training, and a substantial learning curve should be anticipated.

Due to the large number of potential cases, the military may have to consider contracting the work to civilian laboratories. The mtDNA QA testing standards are in place in anticipation of contract requirements. Contracting options have been favorably viewed by the Federal Government in recent times as generally more expeditious and less costly than in-house programs. They are particularly useful for projects of a temporary nature or of a defined lifespan. When a project requires resources greater than that in government or would require additional facilities, the contracting option may provide a solution.

However, services involving confidential matters and matters which are of particular seriousness are usually kept within Government. Government contractors cannot be held accountable to the public in the same way as Government agencies. In the case of apparent problems, government contractors may simply declare bankruptcy, dissolve, or otherwise go out of business leaving the Government without recourse. Ultimately, the Government will be held accountable to the public whether work is performed in-house

or out. Due to the special status of Government functions, laws, regulations, judicial and agency scrutiny have developed such that nongovernment entities are not covered and not under the same constraints as the Government. Accordingly and significantly, laboratories for law enforcement are overwhelmingly in-house Government laboratories.

In this particular context, quality needs and reliability may override cost economies which might be achievable from the private sector. Further, families may be uneasy to trust such private information to private sources. Thus, for reasons of accountability and sensitivity, the Federal Government should play a large role in any large scale program of mtDNA testing of remains from Korea.

Economies of scale are important to mtDNA sequencing operations, as they are elsewhere. Due to the low volume and high cost of laboratory space and equipment in current testing operations, fixed costs are disproportionately high; thus, economies of scale are particularly striking. A centralized laboratory would be significantly more cost effective than contracting multiple small private laboratories. It is easier to oversee and control the quality of one or a few large laboratories than several smaller laboratories.

The Task Force finds that current mtDNA testing efforts could be augmented for large scale operations. There are strong arguments for a centralization of the laboratory work for the sake of vigorous oversight, quality control, and accountability.

VIII. RESOURCE REQUIREMENTS

DSB TOR: To determine the scientific and other resource implications of undertaking large scale mtDNA testing for identification of unassociated remains. *[What are reasonable resource estimates of mtDNA sequencing identifications as currently performed? What are reasonable and likely projected cost estimates associated with performing DNA identifications for Korean War remains? What laboratory personnel are available to perform large-scale mtDNA typing operations?]*

MtDNA sequencing of ancient remains is resource intensive. This is due primarily to the slow and tedious nature of mtDNA sequencing from poor source material, poor quality DNA template, amplification difficulties, and sequencing reactions that must be optimized and repeated numerous times. Furthermore, the forensic nature of evidential testing demands greater care and documentation.

Current cost estimates for DNA sequencing generally are simply not applicable to the mtDNA sequencing operations necessary for CILHI casework, because they are based on estimates of high volume sequencing operations, acceptance of low levels of base miscalls, non-forensic DNA testing standards, and on optimal DNA template. Typical costs from subsidized genome project operations are between $1 and $2 per "finished" base sequenced.

The AFDIL currently sequences 613 bases as two sets of overlapping fragments (a total of 1,046 bases), and confirms the sequence by also sequencing in the reverse direction. An openly bid contract for sequencing the same region for population studies resulted in a open market figure of nearly $1,000 per blood specimen. On the other hand, when performing testing on actual casework (including multiple bone and blood specimens), the British FSS charged the Department of Defense over $100,000 for mtDNA analysis of two cases. This is not unreasonable when considering the cost of salaries for several analysts and Ph.D. molecular biologists over a ten month period.

The DCSPER has funded the AFDIL to perform mtDNA sequencing for remains identification from Southeast Asia. Projected estimates are that 500 cases would require mtDNA testing over a five year period. A case unit for workload projections is four bones and two blood reference specimens. Current mtDNA sequencing

operations cost approximately $17,500 per case at 120 cases per year, excluding the lease cost of the facilities (approximately $9,000 per case). The largest cost components consist of a new laboratory facility, staff (primarily of one DNA Analyst and one DNA Technician per case per month), and equipment and supplies. Fixed costs are quite high for this labor-intensive equipment-driven operation. Retesting accounts for a substantial proportion of the laboratory testing. Recent AFDIL casework has averaged 3.7 bone fragments tested, 5.2 extractions, 37 amplification reactions, and 55.5 sequencing reactions to obtain 1,480 bases of polished/confirmed sequence per case. In a recent AFDIL workload study, approximately 25% of the labor hours were spent in laboratory testing, 50% in data analysis, and 25% in reviewing and reporting the data (Figure 8). Full casework production operations, at ten cases per month, began in October of 1994. Due to the lack of historical data, actual operational costs have not yet been fully established, but appear to be close to projected costs. Due to the immaturity of the program, full operational efficiency has not yet been achieved.

However, the decision to perform DNA testing for unassociated remains from North Korea has yet to be made. There are over 8,100 servicemembers whose remains were not recovered and identified from the Korean conflict. The best information available indicates that no more than 3,000 of the 6,000 remains could be recovered from North Korea. However, it is anticipated that DNA testing would be performed on most cases. Preliminary returns suggest that commingling of remains will be frequent, unless perhaps, joint recovery with U.S. teams is permitted; thus the number of bones tested per case may be increased. Furthermore, a family reference database of mtDNA sequences for all families would need to be constructed for comparison purposes. The total costs for the entire operation would have to include the costs of: 1) recovery and repatriation; 2) documentation review; 3) primary identification processing by CILHI; 4) a family outreach program (Annex G); 5) the family reference mtDNA sequence database; and 6) mtDNA testing of remains.

The Korean workload could be phased in as the Southeast Asia workload ends (Figure 9).

Phase I (Annex H) of a program of DNA testing of Korean remains would require the creation of a database of family reference

mtDNA sequences. It is estimated that 40 to 70% of the 8,100 families of Korean servicemembers from the Korean conflict would be contactable and willing to provide a blood sample for this purpose. The AFDIL protocol normally requires two maternal reference specimens per family. Consequently, the creation of the family reference database is anticipated to require mtDNA sequencing of between 6,480 to 11,340 blood specimens. Some skeletal remains testing could be performed without the full generation of the family database. Blood samples can be sequenced at significantly less cost and higher production rate than ancient skeletal remains. This can be performed over a two to three year period, in current facilities, at a cost of $1.6 to $2.6 million.

Phase II (Annex I) of a program of DNA testing of Korean remains would involve the mtDNA sequencing of up to 3,000 skeletal remains. Given that one laboratory analyst with technical support can process four bone specimens per month, and an average of four bones are tested per set of skeletal remains with the marginal cost for each bone being approximately $4,400 at current efficiency and staff, the projected cost of this program is approximately $51 million in FY 98 dollars over 10 years; the annual projected cost of $4.9 million per year consists of $1.3 million in fixed costs and $3.8 million in marginal costs. Some difficult cases are far more consuming of resources than other cases. The duration of the program could be shortened or lengthened, but no less than eight years if all work is performed in the current AFDIL facility and using current technology. Recovery and preprocessing of remains by CILHI may limit the number of specimens to be tested per year.

This projection assumes a reasonably expected increase of efficiency of 50%, but efficiencies could well be much higher with the development of new technologies. Projected increases in efficiency anticipate functional and organizational shifts from case to bone sequence reporting. Improved computer assisted analysis should be in place by the time of this phase of the program; decreasing analysis time by 50% will increase throughput by 25%. The better preserved bones from Korea along with improvements in extraction and amplification could half the rate of retesting now performed in casework. One particular primer set of the four primer pairs used by the AFDIL performs substantially better than the other three, suggesting that improvements could be made to the other three systems. These

improved efficiencies are foreseeable without the development of new automated DNA technologies, which will surely come to bear within the time frames projected.

New DNA sequencing and other DNA typing technologies could achieve an order of magnitude faster turn around time for testing, and decreased testing cost. However, it will not eliminate the very labor intensive job of extraction of DNA from bone nor the analysis time.

The most speculative variable significantly affecting the cost estimate is the assumption that four bones are tested per skeletal remains. This projection does not assume all bones are DNA tested, but rather that CILHI is able to successfully segregate some skeletal remains. The number could be substantially higher given the potential for commingling.

The Task Force finds that current mtDNA testing efforts at AFDIL are funded appropriately for the Southeast Asia mission.

The Task Force concurs with the projections that analysis of Korean War remains could be accomplished over the next twelve years with an increase of funding of approximately $2 million per year over the cost of current operations.

IX. NEW TECHNOLOGIES

DSB TOR: To evaluate other technologies to assist in the automation and reduction of costs associated with DNA testing. *[What alternative technologies might be brought to bear that may assist current DNA identification efforts? Could new technologies improve DNA typing efforts by improving discriminatory power or enhancing recovery? Could new technologies speed DNA typing efforts? Could new technologies bring down the cost of DNA identifications? Would new technologies replace or confirm current technologies? Would new technologies permit ancillary studies to improve DNA typing efforts, e.g. assessment of original DNA template damage, quantitation of human and other DNA present, etc.? Should the U.S. military fund any investigations, research, or technology development which might enhance cost effective DNA identifications? If so, what are they?]*

Advances in biotechnology are progressing rapidly, particularly as part of the Human Genome initiative. Many will undoubtedly apply to the DNA identification efforts of the U.S. military. Over the next 3 to 10 years, even without expending funds specifically to develop applicable technologies, off-the-shelf technologies are expected to permit far more rapid testing at a fraction of the cost of current testing. However, investment in technologies now will help to accelerate advances in the area of identification and particularly with respect to military efforts to identify ancient remains, such that long term cost savings would surely be realized. A relatively small investment by the military could be leveraged to achieve significant gains applicable to their needs.

Significant improvements could be achieved over the currently employed DNA typing methods and technologies. New methods and technologies could potentially improve the success of DNA typing efforts as well as substantially reduce the turn-around-time and cost of testing. Possible areas for exploration and investment include, but are not limited to, the following:

A. Specimen Preparation and DNA Extraction

The first steps in mtDNA testing of skeletal remains are sample preparation and extraction of the DNA. Significant improvements could be made in enhancing the success, speeding processing, and

making it more cost efficient.

1) Recovery Improvement

The success of mtDNA identification efforts primarily rests on the ability of the laboratory to get useful mtDNA information from the samples of skeletal remains tested. The AFDIL was successful in only half of the cases initially, but has been more successful in most recent cases from Southeast Asia. Several re-extractions and re-amplifications were necessary to accomplish these successes. Not only does this demonstrate significant progress has already been made, but it also suggests that current testing is pushing the limits of technology. It is anticipated that less well-preserved remains will be returned from Southeast Asia. Early testing on Korean remains suggests that the samples will be more challenging. Even marginal improvements in mtDNA extraction will significantly impact the ability of the military to effect mtDNA identifications. Whole genome amplification through a random primer technique, use of alternative polymerases, improved buffer systems and reaction conditions are obvious directions for possible improvements.

2) Automated Specimen Preparation and Extraction

Preparation and extraction of mtDNA from bone specimens are currently labor-intensive and time-consuming processing. Even if subsequent sequencing is speeded, this step may act as a significant bottleneck. Automation of sample preparation and extraction may be possible. As already performed by some in the ancient DNA community, the surfaces of bone samples could be cleaned chemically rather than by grinding, as currently performed at the AFDIL. Then instead of mechanical breakdown of the bone samples, enzymatic dissolution of the bone may be possible, particularly when accompanied by ultrasonic agitation. Cleaned bone samples could be placed into troughs for incubation, and the dissolved extract subsequently robotically manipulated through the next stages of processing.

Automation could also apply to reference blood samples. Typing family reference samples to create a database for identification of unassociated remains will require literally thousands of mtDNA typings. Automated extraction would greatly benefit these efforts. The AFDIL currently transfers liquid blood samples to bloodstains on cards for ease of use and storage. Small punches

from the cards are then used in analyses. The development of a robotic instrument to punch bloodstain cards and then extract the card punches could dramatically speed the process.

Automation will speed processing and free labor to perform other tasks. Labor is by far the greatest cost in mtDNA typing of ancient skeletal remains. Accompanying substantial decrease in operational costs should be realized after the capitalization of the equipment. Automation may eliminate many potential manual errors in repetitive sample handling. Moreover, automation may further decrease the chance of cross contamination. The automated equipment may also be a boon to other AFDIL service casework, which has call for high volume and rapid sample processing.

3) Repair

Ancient DNA is not only degraded but damaged. Strand nicks, cross-links, extraneous molecular attachments, and other sorts of damage may prevent a polymerase from reading through the length of a DNA fragment. If the DNA reparative machinery that is normally found in a cell can be used to repair small defects in DNA target templates, then the ability to amplify such ancient DNA may be greatly enhanced. Enzymes known as ligases are commonly used in molecular biology labs to splice strands together and could repair a nick in one strand of a double-stranded fragment. A polymerase may replace a missing base by matching the complement of the opposite strand. Bacterial cloning maybe useful to repair damaged DNA using the full set of repair machinery of the host organism. Some evidence suggests that repair of the DNA before amplification can indeed assist recovery of sequence information from ancient DNA. A marginal improvement in the success of the recovery of mtDNA information from bone specimens will significantly impact current identification efforts.

B. Specimen Evaluation

Evaluation of the DNA template from which the DNA information is to be derived may permit accurate sample loading, tailored

amplification conditions, an assessment of likelihood of success, and an indication of how hard to try to recover information from the sample.

1) mtDNA Quantitation

Accurate quantitation of the sample DNA target is an important aspect of various methods of mtDNA testing. Quantitation allows accurate amounts of the sample extract to be added to an amplification reaction mixture. An absence of quantitated mtDNA should result in a determination that further testing would be an unwarranted waste of time. Quantification of source mtDNA would permit inferences of susceptibility to sampling and enzymatic errors. Unfortunately, current methods of quantitating mtDNA are not human mitochondria specific and are too insensitive to be of great value in these old skeletal remains cases. A semi-quantitative sample gel electrophoresis is performed which is virtually always negative unless a significant bacterial DNA content is present. A sensitive human mitochondria specific assay could be developed to assist in this casework. Competitive PCR assays, serial dilution assays, and kinetic assays are potential methods, among others, for quantitation.

2) DNA Damage Assessment

The DNA from ancient skeletal remains is severely degraded and damaged. An assessment of the damage to the DNA to be analyzed would be valuable as an indicator of how best to scientifically approach the sample for analysis and whether or not to continue to expend time and effort in reanalysis of the sample. Such an assessment may be useful to determine whether to attempt repair of the damage before amplification.

Environmental damage of DNA takes the form of certain predictable classes of chemical reaction resulting in the formation of certain specific DNA adducts. A new sensitive, rapid, inexpensive, and specific method that has recently been developed involves matrix-assisted laser desorption followed by mass spectrographic analysis.

C. DNA Amplification

Any mtDNA analysis of skeletal remains will require amplification of the few mtDNA template strands present. This is a critical step in the mtDNA testing. The basic PCR amplification process has not changed substantially since it was first introduced.

1) Alternative Polymerases

The PCR reaction involves the polymerase enzyme produced from a microbe which lives in waters of hot springs called Thermus Aquaticus and known as the Taq polymerase. While Taq polymerase is an excellent polymerase for most DNA amplifications, it does have its limitations. The knowledge of DNA polymerases has recently increased and continues to improve. Newer alternative polymerases offer the possibility of greater fidelity, processivity, and ability to process damaged template or tolerate Taq inhibitors.

2) Inhibition

Inhibition of the PCR amplification reaction is often encountered in the processing of old skeletal remains. Currently, an enormous effort is expended in overcoming a variety of technical hurdles encountered with AFDIL cases. Improvements in overcoming inhibition are likely through better optimization of reaction conditions, better DNA extraction methods, and through alternative polymerases, as previously mentioned. Because inhibition directly prevents the success of DNA typing, even marginal increases are likely to be significant.

D. DNA Analysis

Currently mtDNA analysis has been performed using conventional sequencing on conventional equipment. Due to the Human Genome Initiative advances in DNA analysis are in rapid development. Several possibilities exist that could dramatically increase throughput and decrease the cost mtDNA analysis. Complete automation with integration of sample preparation, DNA

extraction, DNA amplification, and analysis is an achievable and desirable goal.

1) Automation of Sequencing Reactions

Current sequencing reactions are performed manually. Automated methods for performing sequencing reactions are available. Automation not only can achieve higher throughput and reduce overall costs, but also may achieve better results. The large number of steps involved in current testing creates the potential for manual errors to be made. Automation will allow dye primer and T4-sequenase chemistries that are too tedious for manual production sequencing casework efforts.

2) Capillary Electrophoretic Arrays

Current sequencing technology is based on gel electrophoresis in which DNA is pulled across a slab of polyacrylamide gel using an electric current. The speed of this system is determined by the electric voltage applied, but the voltage is limited by the heat generated by the current in the gel. If the electrophoresis is carried out in a very fine capillary then a substantially higher voltage can be applied and a faster run time achieved. Run times are measured in minutes instead of hours. Due to the higher resolution, a greater number of bases can be read per run. Furthermore, a capillary can be sample loaded in an automated fashion, obviating the need for pouring and loading a gel. As a result, many DNA samples (up to 36 or 48) can be run simultaneously on a slab gel, but only one DNA sample can be loaded in a capillary per run. Instruments that employ arrays of capillaries are currently in development and could dramatically speed DNA sequencing using conventional sequencing chemistries, with an accompanying reduction in cost.

3) Sequencing Microchips

Another major new technology under development is sequencing by hybridization (SBH) performed on microchips. Microlithography is an engineering tool used in the manufacture of computer microchips. Microlithography will permit the cost-effective generation of arrays of thousands of DNA probes which enables SBH technology. One of the first prototype "DNA chips" will sequence the D-Loop of mitochondria. This microchip can perform sequencing at a fraction of the time and cost of current DNA

sequencing methods. These microchips should be first used to confirm current sequencing technology and perhaps later employed as the primary method of sequence information.

4) Single-Point Polymorphism Assays

An alternative technique to sequencing which captures most sequence data, developed by Dr. Mark Stoneking in conjunction with Dr. Henry Erlich's laboratory, uses hybridization of sequence specific oligonucleotide (SSO) probes to amplified segments of HV1 and HV2 that have been immobilized on a membrane. These SSO probes are designed to detect variation at specific polymorphic nucleotide positions to produce a set series of yes/no results known collectively as a "mitotype". This system is more limited in its discriminatory power than full sequencing; but it offers the advantages of being quick, inexpensive, and does not require great technical skills. Preliminary results suggest that it is more sensitive than current sequencing methods. An analysis of mtDNA SSO-type variation in 142 U.S. whites using 23 SSO-probes revealed 99 different types, and an average probability of identity of 1.9%. A commercial dot/blot strip is under development for the typing of the mitochondrial D-Loop. Moreover, a similar method could be used to interrogate sites outside the D-Loop which are so spread out as to make the cost of sequencing for these polymorphisms prohibitive.

Dot/blot technology could be reduced to microchip probe arrays. Site polymorphisms could also be interrogated through the use of oligonucleotide ligation assays and detection systems, which offer the advantages of greater specificity and large scale multiplexing. These various systems for detecting point polymorphisms are highly amenable to automation. Also denaturing gel electrophoresis could be used as a method to screen for polymorphisms. These technologies could be employed as a screening tool to sort bones with cases, as a method of confirming sequencing results as well as a primary mtDNA typing method to increase the discriminatory potential.

5) Computer Software

Currently, far greater time is spent in the analysis of the sequencing results than in sequencing itself. A significant portion of the time is spent appropriately formatting the data and then checking the raw sequencing data and reading. Computer

software could be developed which could perform these functions more quickly, more accurately, and without tedium. A preliminary neural network system already demonstrates a several fold diminution in the rate of instrumental miscalls from the raw output of the Applied Biosystems, Incorporated, model 373a instrument currently used by the AFDIL.

A laboratory information management system (LIMS) could be developed which could greatly improve the efficiency of current efforts. A LIMS could integrate robotic systems, capture instrument data and export it into analysis software, generate reports, improve case tracking and perform quality assurance functions.

The idea that new technologies should be developed and explored, should not be construed as implying current technologies are in any way inaccurate or not worthwhile. Current technology is, however, costly and time consuming, and mitigation of these burdens would be welcome.

The Task Force finds that new technologies should be surveyed for the best prospects for increasing the success and efficiency of DNA identification of ancient skeletal remains. Future investments should be guided by the progress of technology and appropriate advice.

X. SCIENTIFIC BASIS

DSB TOR: To determine the degree of scientific experience and expertise available to support the Task Force findings. *[Is the current scientific foundation for DSB findings to these questions adequate or is it premature to answer the foregoing questions? Is further investigation and experience necessary or helpful in determining the answers to the above questions? Should the Task Force be reconvened or should the military otherwise revisit the above issues in the future?]*

The foregoing Task Force Report findings have been based on substantial scientific evidence. While it is always true that more information can and will be produced bearing on important scientific questions, these findings do not seem premature. We note that a Quality Assurance Oversight Committee has been created by direction of the ASD(HA), composed of civilian consultants, which will act to ensure that efforts will continue to be performed with the greatest scientific integrity.

The Task Force supports the ASD(HA) in the creation of a scientific advisory board composed of civilian technical consultants.

The Task Force finds that current DNA identification efforts are supported by sufficient scientific evidence to proceed, in particular with application of mtDNA sequencing to ancient remains from the Korean conflict.

CONCLUSION

The Defense Science Board Task Force find that ancient skeletal remains can be accurately and successfully identified through mtDNA testing. The current methods of mtDNA sequencing are scientifically sound and can be performed in a reliable manner. Furthermore, mtDNA can be used to identify unassociated remains through the creation of a database of family reference specimens.

1. FEASIBILITY:

The Task Force finds that identification of so-called ancient skeletal remains by a program of mtDNA testing is possible, particularly in association with other information. A few specimens may remain unresolved. Although contenders may emerge; at this time, mtDNA sequencing technology is the most appropriate technology.

2. FACTORS:

The Task Force finds that the present probability of coincidental matches between mtDNA control region sequences is no more than a few percent. Once sequences from 500 members of a population have been determined, precise statements about the chance of a false association of a set of remains with a family will be able to be made. Published data may be of value, but samples will be needed from Southeast Asian populations. The precision is expected to suffice in the vast majority of cases, given other non-DNA evidence, to effect DNA identification of unassociated Korean remains. It will not be possible, however, to identify every bone. A great deal can be done with anatomical and historical evidence alone.

The Task Force finds that control of contamination is essential to PCR-based laboratory testing. Some contamination is unavoidable, particularly in mtDNA testing of ancient remains, but it does not preclude reliable casework testing where redundancy, good laboratory practices, and appropriate cautionary language are used and constant oversight is maintained.

The Task Force finds that casework experience demonstrates capability to type Korean skeletal remains.

3. RELIABILITY:

The Task Force finds that appropriate measures must be taken to prevent and control possible contamination in the testing laboratory.

The Task Force finds that current AFDIL protocols, if diligently performed, are capable of generating quality mtDNA sequence identifications.
The Task Force finds that adequate quality assurance requires accreditation and an oversight board.

The Task Force finds that the program for mtDNA sequencing quality assurance promulgated by the Office of the Assistant Secretary of Defense (Health Affairs) for identification of ancient remains is adequate and responsive.

4. OTHER DNA TARGETS:

The Task Force recommends that the AFDIL investigate the potential to perform DNA typing outside the mtDNA control region.

5. FAMILY REFERENCE DATABASE:

The Task Force finds that, with a reasonable effort, a sufficient proportion of families are expected to provide DNA samples to allow identification of many of the unassociated remains from Korea through mtDNA testing, and to attach meaningful probability statements in those cases.

The Task Force recommends consideration be given to the collection of DNA reference specimens from maternal and paternal family members in case future technology permits nuclear DNA testing. Collections from nonmaternal kindred members (e.g. children) should be made with full disclosure of realistic expectations.

6. STATISTICAL DATABASE:

The Task Force finds that the existing databases, and those anticipated from the family collections, provide an adequate basis for the current mtDNA sequencing efforts.

The Task Force recommends an expansion of the current AFDIL

database.

7. LARGE-SCALE OPERATIONS:

The Task Force finds that current mtDNA testing efforts could be augmented for large scale operations. There are strong arguments for a centralization of the laboratory work for the sake of vigorous oversight, quality control, and accountability.

8. RESOURCE REQUIREMENTS:

The Task Force finds that current mtDNA testing efforts at AFDIL are funded appropriately for the Southeast Asia mission.

The Task Force concurs with the projections that analysis of Korean War remains could be accomplished over the next twelve years with an increase of funding of approximately $2 million per year over the cost of current operations.

9. NEW TECHNOLOGIES:

The Task Force finds that new technologies should be surveyed for the best prospects for increasing the success and efficiency of DNA identification of "ancient" skeletal remains. Future investments should be guided by the progress of technology and appropriate advice.

10. SCIENTIFIC BASIS:

The Task Force supports the Assistant Secretary of Defense (Health Affairs) in the creation of a scientific advisory board composed of civilian technical consultants.

The Task Force finds that current DNA identification efforts are supported by sufficient scientific evidence to proceed, in particular with application of mtDNA sequencing to ancient remains from the Korean conflict.

The Task Force notes that the technology for mtDNA sequencing will continue to improve and become less costly due to activities both within and outside the military. Military efforts to identify skeletal remains using DNA will greatly impact law enforcement and contribute to the science of molecular biology, and molecular anthropology.