NCJRS
National Criminal Justice Reference Service

Office of Justice Programs
BJA BJS NIJ OJJDP OVC SMART

The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	Assessing the Effects of DNA Damage on Next Generation mtDNA Sequencing
Author(s):	Mitchell M. Holland
Document Number:	252862
Date Received:	April 2019
Award Number:	2015-DN-BX-K025

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Agency: National Institute of Justice

Award number: 2015-DN-BX-K025

Project Title: Assessing the Effects of DNA Damage on Next Generation mtDNA Sequencing

PI: Mitchell M. Holland
Associate Professor BMB-Forensics
mmh20@psu.edu
814-865-5286

Submitting official: Mitchell M. Holland
Associate Professor BMB-Forensics

Submission date: 09/27/2018

Recipient Organization: The Pennsylvania State University – Univ Park
Office of Sponsored Programs, 110 Technology Center Building,
University Park, PA 16802-7000

Award Period: 01/01/2016 to 12/31/2018 (no cost extension)

Reporting Period End Date: 12/31/2018 Final Summary Overview

Signature of Submitting Official:

____

Mitchell M. Holland

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

PURPOSE

The ability to resolve, report, and leverage the discrimination potential of heteroplasmy has significantly enhanced the value of mitochondrial (mt) DNA analysis in forensic casework [1]. A massively parallel sequencing (MPS) approach allows the community to capture this information on a routine basis. The current study was a multifaceted assessment of the impact of DNA damage on the interpretation of MPS mtDNA heteroplasmy from a variety of perspectives. The first part of the study focused on methods development (an mtDNA qPCR assay to assess both quantity and quality of mtDNA; amplification schemes for production of mtDNA products for MPS analysis; and bioinformatic pipelines and software development for analysis of MPS mtDNA data), along with initial considerations of DNA damage. These accomplishments were critical for successful completion of the project, and resulted in multiple publications [2-4]. The newly developed and commercially available software has been adopted by key laboratories in the forensic community (for example, the FBI Laboratory and the Ohio Bureau of Criminal Investigation). The second part of the study has focused on the evaluation of different approaches to modeling deamination, depurination, and oxidation-based DNA damage, including whether approaches of biological repair can reverse the damage. These studies have led to targeted experiments on mock-evidence-type samples (for example, ammunition components with different metallic surfaces, ancient skeletal remains from the 9th to 20th centuries, and DNA damage in human hairs), and to modeling experiments assessing the impact of low template samples on the damage process. The findings have already resulted in one publication [4], with another manuscript submitted to Forensic Science International: Genetics in August [5], and at least two more manuscripts in preparation. Scholarly articles have included recommendations for developing best practices when conducting mtDNA MPS analysis in forensic laboratories.

This resource was prepared by the author(s) using Federal funds provided by the U.S.
Department of Justice. Opinions or points of view expressed are those of the author(s) and do not
necessarily reflect the official position or policies of the U.S. Department of Justice.

EXPERIMENTAL DESIGN & METHODS

Project findings, including methodologies associated with an mtDNA qPCR assay, amplification schemes for production of mtDNA products for MPS analysis, and bioinformatic pipelines and software development for analysis of MPS mtDNA data, can be found in cited publications [2-4]. Unpublished project design elements and methods are summarized below.

Passive Damage Study

a. Sample Source: In collaboration with the OCME in NYC, we have evaluated methods for storing DNA extracts that do not involve cold storage, particularly -20°C storage. The goal of this study was to assess whether the storage method had an impact on the MPS results, including the introduction of DNA damage, with levels of damage evaluated over the first six months of storage. The study involved three storage methods: 1) 25 uL DNA extracts in water stored at -80°C at a concentration of 0.5 ng/uL, 2) DNA extracts (25 uL of 0.5 ng/uL in water) deposited and dried on Whatman paper and stored at room temperature, and 3) 25 uL DNA extracts (25 uL of 0.5 ng/uL in water) lyophilized and stored at room temperature. Sets of 30 sample were stored for 0, 3, and 6 months. This was a modification of the experimental design from the original proposal (through discussions with our consultants), as the focus shifted primarily to active damage assessments rather than passive damage studies.

b. Data Generation: At each time point the following amount of each sample was added to an amplification reaction using the PowerSeq™ CRM Nested System kit from Promega: 1) 2 uL of the -80°C samples, 2) 2 uL of a consistent cutting (~0.5 cm^2) from the center of the filter paper soaked in 25 uL of low TE, and 3) 2 uL of the lyophilized sample reconstituted in 25 uL of water. Following amplification, samples were taken through library preparation and

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

run on the Illumina MiSeq using the v3 600-cycle kits from Illumina in 300x300 paired end read mode. Each time point included negative and positive controls.

c. Data Analysis: Sequence data was mapped to the revised Cambridge Reference Sequence (GenBank ID NC_012920.1) using the MiSeq Reporter integrated computer software platform (MSR; v2.1.43 and v2.2.29), which operates on a Burrows-Wheeler Aligner (BWA) and the Genome Analysis ToolKit (GATK) for variant calling of single nucleotide polymorphisms (SNPs) and insertions and deletions (indels). Secondary analysis of the MSR generated FASTQ data (sequence and quality scores) was performed using GeneMarker® HTS software. Additional secondary analysis was achieved using pipelines developed by our laboratory using a combination of UNIX line commands and the R environment.

Active Damage Study

a. Sample Source: The general sample source was extracted DNA at a concentration suitable for active damage studies stored in low-TE at -20°C.

b. Data Generation: Extracts were subjected to chemical agents or conditions that cause specific types of DNA damage; deamination, depurination, and oxidative-based damage. Amplification was conducted using the PowerSeq™ CRM Nested System kit or a 1 kb in-house method. Sample sets were run through the library preparation and MiSeq process to assess levels of existing heteroplasmy, and the production of false positive heteroplasmic sites caused by DNA damage.

c. Analysis: See above.

Low-Template Damage Study

a. Sample Source: Extracted DNA was stored in water for two weeks at room temperature (~22-25°C). We conducted numerous studies to identify these conditions as ideal for

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

facilitating hydrolytic DNA damage; typically resulting in deamination and depurination events. Two approaches were employed to assess damage. First, samples from five donors were diluted to 2,500, 25,000, and 100,000 copies of mtDNA/uL prior to storage at room temperature. This first dimension assessed the impact of damage as the quantity of DNA in a sample decreases. Second, samples from five donors were diluted to 100,000 copies of mtDNA/uL prior to storage at room temperature, and then a portion of the samples were diluted to 2,500 and 25,000 copies/uL after the two week period. This second dimension assessed the impact of diluting a previously damaged sample. An undamaged control with 100,000 copies of mtDNA/uL and stored at -20°C was run with the five damage samples, for a total of six samples per donor.

b. Data Generation: The extracts were amplified in duplicate. This approach allowed for an assessment of whether differences should be expected depending on when the damage occurs. Amplification was conducted using the PowerSeq™ CRM Nested System kit and a 1 kb in-house method to assess the difference between a short or long amplicon approach when observing damage. Sample sets were run through the library preparation and MiSeq process.

c. Analysis: See above.

Mock Forensic Evidence Damage Study

a. Sample Source: Evidence containing biological cells from known DNA contributors was exploded to mimic samples typically received by the ATF Laboratory. Hair samples were tested from multiple contributors to assess the impact of damage. Skeletal remains of decades to centuries in age were examined for patterns of DNA damage, including sets of remains associated with a nun from Croatia who is being considered for Sainthood.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

b. Data Generation: Amplification was conducted using the PowerSeq™ CRM Nested System kit, a 1 kb in-house method, and a capture method developed at AFDIL [6]. Sample sets were run through the library preparation and MiSeq process.

c. Analysis: See above.

Repair of DNA Damage

a. Sample Source: Samples were consistent with the Active Damage experiments, as these studies were run in parallel.

b. Data Generation: Amplification was conducted using the 1 kb in-house method. Sample sets were run through the library preparation and MiSeq process.

c. Analysis: See above.

DATA ANALYSIS

A critical ingredient necessary for the adoption of an MPS approach in forensic laboratories is the availability of a suitable software package for data analysis. We collaborated with SoftGenetics, Inc, to develop a software package for forensic researchers and practitioners; GeneMarker® HTS [3]. Careful examination of mtDNA MPS data is important, as illustrated by the publication of several high-profile reports that have been deemed in error due to an inability to distinguish between heteroplasmy and other sources of mixed data, including those associated with software alignment anomalies [7]. The MPS data for this study were analyzed using GeneMarker® HTS at a 1% analytical threshold and a 2% reporting threshold for minor sequence variants.

We used a conservative approach to assess error rates, considering MPS data with nucleotide calls observed in <50% of the sequencing reads as assumed error. An example of this can be found in a recent publication that reports on differentiation of maternal relatives using MPS

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

mtDNA analysis of heteroplasmy [8]. Overall methods for data analysis can be found in the previous publications from our laboratory.

FINDINGS

Project findings, including methodologies associated with an mtDNA qPCR assay, amplification schemes for production of mtDNA products for MPS analysis, and bioinformatic pipelines and software development for analysis of MPS mtDNA data, can be found in cited publications [2-4]. The passive study was reported to the OCME, with details provided in a previous semi-annual report to the NIJ. The active damage and repair studies are being completed, with oral and poster presentations given, and a publication to be submitted in the near future. The low template study provided a valuable perspective on the influence of low template samples when considering the impact of DNA damage on the interpretation of mtDNA MPS data. In addition, the findings confirmed previous observations that as the amplicon size being evaluated for damage increases, the amount of damage decreases. This is an important outcome, as most of the studies in forensic laboratories involve the analysis of smaller amplification products. Final data analysis has been completed and a manuscript is in preparation.

An example of the impact of DNA damage on mock case samples is illustrated through the experiments conducted on ammunition components. In particular, this was an assessment of oxidative damage facilitated by copper ions. This work has been submitted for publication [5], and is summarized below.

In an effort to assess the impact of recovering mtDNA from the surface of metallic ammunition casings and projectiles, three separate sets of 36 samples containing copper, brass, and aluminum were used in three independent experiments. Recovery of mtDNA (measured as copies of mtDNA/uL) from the surface of casings made of aluminum was, in most cases,

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

significantly higher when compared to copper projectiles or brass casings. In addition, mtDNA yields from brass casings significantly exceeded values for copper projectiles when considering touched components. This is consistent with the expectation that as copper ion concentration increases, so does DNA damage and degradation. The yields for samples collected from aluminum casings with swabs moistened with water versus 0.5M EDTA were 94.3% and 90.1%, respectively. In comparison, the yields for copper projectiles were 3.9% and 13.0% of expected, and for brass casings were 1.7% and 18.8% of expected. These data illustrate that yields are likely to be relatively high when liquid DNA is deposited on the surface of ammunition components made of aluminum, whether water or EDTA is used to collect the DNA, and that collection with EDTA will result in greater recovery of DNA, on average, when working with copper or brass components.

Naked DNA in an aqueous form when deposited on a surface containing copper ions will rapidly facilitate DNA damage, leading to degradation and an inability to amplify larger DNA templates, with reasonably good improvements in yield when using 0.5M EDTA for the collection process. The driving force for DNA damage and degradation when exposed to copper and brass surfaces is an aqueous environment. As such, individuals who place cartridges in their mouth before loading ammunition in a firearm will deposit cellular material in liquid form, resulting in an expected acceleration of the damage and degradation processes when associated with copper and brass surfaces.

As expected, degradation indices (ratios) were highest for copper projectiles and brass casings, with aluminum producing relatively low levels of degradation; *p*-values of 1.36×10^{-6} for copper compared to aluminum and 8.95×10^{-5} for brass compared to aluminum, with copper to brass at 3.56×10^{-1}. When considering the collection method used to lift DNA from the

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

surfaces of the casings and projectiles (i.e., the double swab method with water versus 0.5M EDTA) there was little effect on the degradation ratios for aluminum casings. While there was a small reduction in the degradation ratios for brass casings when collection was performed with 0.5M EDTA, the opposite effect was observed with copper projectiles, but neither difference was significant, most likely due to extensive scatter in the data. As expected, the vast majority of DNA degradation observed for copper and brass samples was associated with DNA deposited in liquid form. Far less degradation was observed when touch DNA was deposited on the surfaces. In addition, degradation improved when touch DNA was collected with 0.5M EDTA for both copper and brass surfaces, approaching degradation levels seen in aluminum samples. This correlated well with the data for liquid samples, as the aqueous environment for those samples appears to have facilitated early damage and degradation to the DNA, lessening the impact of recovery with EDTA. It is important to note that degradation was still relatively high for touch DNA samples on copper and brass when collected with water, suggesting that addition of EDTA to the collection process is impactful.

All samples from aluminum casings produced full MPS mtDNA profiles. In contrast, success rates for copper and brass were significantly lower; 44.4% and 43.1% full profiles, respectively, or 32 and 31 profiles from the 72 samples of each type, with *p*-values of 1.74×10^{-8} for copper compared to aluminum and 9.01×10^{-9} for brass compared to aluminum. Overall, the ability to generate profile information from copper and brass was 59.7% and 57.0%, respectively, when including partial profiles. While clearly lower than aluminum, these values are high enough to warrant performing mtDNA analysis on copper and brass ammunition on a routine basis; or any other surface made of copper or brass.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Overall, 44 sites of damage were observed in the entire dataset; 23 sites associated with copper samples, 14 for brass samples, and seven for aluminum samples. Of the 44 sites, 33 (75%) were present in 1-2% of the total reads for a sample, below the reporting threshold of 2%, but above the analytical threshold of 1%. Therefore, these sites would not impact the reporting of low-level heteroplasmy. However, the remaining 11 sites were above the threshold, with five sites above 5% of the total reads, but no site was observed in greater than 10% of the total reads. In addition, no site was replicated in the duplicate amplification and MPS analysis conducted for the copper and brass samples. This is consistent with previous studies looking at the impact of DNA damage on the reporting of heteroplasmy [2], and further illustrates that replicate analysis is an important approach for mitigating the impact of DNA damage when conducting forensic mtDNA analysis on an MPS platform.

A broad study on the impact of sample type and MPS approach (amplification scheme, library preparation, etc) on the noise in the MPS system is being conducted. Data from the current study is included in that assessment, with a manuscript in preparation.

CRIMINAL JUSTICE IMPACT

Thus far, we have presented our work on this project approximately 18 times over the course of 30 months. For example, oral presentations were given at the Pennsylvania Research Exchange in State College, PA in April 2018, the annual NEAFS meeting in November 2017, Hobart & William Smith Colleges in April 2017, Centre of Forensic Sciences in Toronto, Canada in May 2017, ISABS meeting in June 2017, ISHI meeting in September 2016, a Promega Workshops in July 2016, and a Bode Mid-Atlantic/Illumina Conference in November 2016. Poster presentations were given at the 29th ISHI meeting in September 2018, the AAFS meeting in February 2018, and the 28th ISHI meeting in September 2017. We also organized two

This resource was prepared by the author(s) using Federal funds provided by the U.S.
Department of Justice. Opinions or points of view expressed are those of the author(s) and do not
necessarily reflect the official position or policies of the U.S. Department of Justice.

consultant workshops, one on site at Penn State and one through a webinar. These were attended by members of our three consultant laboratories; AFDIL, Mitotyping, and the ATF Laboratory. Lastly, a well-attended and received one-day workshop was held at the 29th ISHI conference in Phoenix, AZ, entitled *The Future is Now for MPS mtDNA Analysis*. Guest speakers included scientists from the FBI Laboratory, NIST, CA DOJ, OH BCI, Battelle, AFDIL, and the Institute of Legal Medicine, Medical University of Innsbruck.

The interactions with our consultants have ensured that the project is visible and relevant in high profile forensic laboratories. Charla Marshall and Erin Gordon are members of the Emerging Technologies section of AFDIL. Dr. Marshall is the leader of that section and has been associated with numerous publications covering various topics of MPS mtDNA analysis. AFDIL is one of the forces leading the charge for advancement and implementation of MPS mtDNA analysis. Gloria Dimick is the Technical Leader of Mitotyping Technologies, one of the most respected commercial forensic mtDNA laboratories in the world. Todd Bille and Steve Weitz were the least experienced with mtDNA analysis, so gained the most by learning about its capabilities and the impact of the project. They were keenly interested in the outcomes of the study associated with ammunition components.

Ultimately, the most important outcome of our project will be publications. We already have three publications in press [2-4], have a fourth publication submitted [5], and are in the process of writing at least two additional manuscripts. Therefore, we anticipate that at least six publications will result from this project. Each of our papers provides recommendations that the forensic community can use to develop best practices as they implement mtDNA MPS methods in their laboratories. Therefore, the outcomes of this project should have a significant impact on forensic mtDNA casework and the criminal justice system.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

APPENDIX

BIBLIOGRAPHY

1. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of the remains of Tsar Nicholas II. Nat Genetics. 1996;12:417-20.
2. Rathbun MM, McElhoe JA, Parson W, Holland MM**.** Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data. Forensic Sci Int: Genetics. 2017;26:1-11 (https://doi.org/10.1016/j.fsigen.2016.09.008).
3. Holland MM, E Pack, JA McElhoe. Evaluation of GeneMarker® HTS for improved alignment of mtDNA MPS data, haplotype determination, and heteroplasmy assessment. Forensic Sci Int: Genetics. 2017;28:90-98 (https://doi.org/10.1016/j.fsigen.2017.01.016).
4. Gallimore J, Holland CA, McElhoe JA, Holland MM. Assessing heteroplasmic variant drift in the mtDNA control region of human hairs using an MPS approach. Forensic Sci Int: Genetics. 2018;32:7-17 (https://doi.org/10.1016/j.fsigen.2017.09.013).
5. Holland MM, Bonds RM, Holland CA, McElhoe JM. Recovery of mtDNA from unfired metallic ammunition components with an assessment of sequence profile quality and DNA damage through MPS analysis. Forensic Sci Int: Genetics. 2019;39:86-96 (https://doi.org/10.1016/j.fsigen.2018.12.008).
6. Marshall C, Sturk-Andreaggi K, Daniels-Higginbotham J, Oliver RS, Barritt-Ross S, McMahon TP. 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: Genetics. 2017;31:198-206.

This resource was prepared by the author(s) using Federal funds provided by the U.S.
Department of Justice. Opinions or points of view expressed are those of the author(s) and do not
necessarily reflect the official position or policies of the U.S. Department of Justice.

7. Just RS, Irwin JA, Parson W. Questioning the prevalence and reliability of human mitochondrial DNA heteroplasmy from massively parallel sequencing data. Proc Natl Acad Sci. 2014;111:e4546-e47.
8. Holland MM, Makova KD, McElhoe JA. Deep-coverage MPS analysis of heteroplasmic variants within the mtgenome allows for frequent differentiation of maternal relatives. Genes. 2018;9, 124:1-21 (https://doi.org/10.3390/genes9030124).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.