Quantitative PCR and Sanger Sequencing of Mitochondrial DNA Recovered from Waterlogged Bone

Kailey Babcock
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/frsc_projects

Part of the Biology Commons, and the Molecular Biology Commons

© The Author(s)

Downloaded from
https://scholarscompass.vcu.edu/frsc_projects/20

This Directed Research Project is brought to you for free and open access by the Dept. of Forensic Science at VCU Scholars Compass. It has been accepted for inclusion in Master of Science in Forensic Science Directed Research Projects by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
Quantitative PCR and Sanger Sequencing of Mitochondrial DNA Recovered from Waterlogged Bone

Kailey Babcock

Semesters of Enrollment: Spring 2020, Fall 2020, Spring 2021

Host Laboratory: Dr. Baneshwar Singh’s Laboratory

Research Mentor: Dr. Baneshwar Singh

Date of Submission: March 21\textsuperscript{st}, 2021

Date Approved: Pending

\textit{Partial fulfillment statement:} A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Forensic Science at Virginia Commonwealth University.

© Kailey Babcock 2021
All Rights Reserved
Acknowledgements

Dr. Baneshwar Singh

Dr. Claire Cartozzo

Dr. D’Arcy P. Mays

Dr. Tal Simmons
Abstract

In forensic contexts, samples containing heavily fragmented DNA are commonly encountered. Compromised biological samples are especially prevalent in instances where human remains have been submerged in an aqueous environment for extended periods of time. Nuclear DNA is particularly vulnerable to the prolonged exposure to heat, moisture, and bacterial degradation that are prevalent in aquatic settings. Paired with the difficulty of recovering DNA from skeletal remains, which are often the only remaining component after the soft tissues have been stripped away, mitochondrial DNA (mtDNA) analysis serves as an invaluable alternative. In this multifaceted study, mtDNA analysis was performed on waterlogged bone samples to determine the capabilities and limitations of a proposed mtDNA workflow. The results originating from two bone sample types, scapula and rib, were compared, in addition to a comparison between a freshwater lentic system, Henleys Lake, and a freshwater lotic system, the James River. To compare the effectiveness of two extraction methods, the samples were extracted once using an organic phenol-chloroform method and once using the magnetic-bead based ChargeSwitch® gDNA Plant Kit. The bone specimens used in this study were collected approximately every 500 ADD, up to 4500 ADD, for the purpose of determining at which ADD interval mtDNA sequence recovery was no longer possible. After using gel electrophoresis as a preliminary qualitative assessment of the total DNA extracts, quantitative PCR (qPCR) was then used as both a quantitative and qualitative assessment of the mtDNA. Using the qPCR data, the analyses of covariance (ANCOVA) demonstrated that the lentic location and the rib bone type exhibited more extensive mtDNA degradation than their counterparts, and that mtDNA quantitation values significantly decreased as ADD increased. Additionally, ChargeSwitch proved to be most effective at recovering high quantities of mtDNA. Using only the ChargeSwitch extracts, mitochondrial sequencing was performed, demonstrating that while location did not affect the ability of the workflow to recover a sequence, rib bones were more likely to yield a sequence than scapulae. It was also found that with each 500 ADD increase, the likelihood of recovering a mitochondrial sequence decreased by 40%. In its entirety, this study and its proposed workflow can assist forensic analysts with recovering mitochondrial profiles from degraded, waterlogged bone samples originating from multiple aquatic settings. Ultimately, a successful yield of a mitochondrial sequence will help with identification efforts for unidentified skeletal remains.

Keywords: Mitochondrial DNA, Sanger sequencing, quantitative PCR, submerged skeletal remains
1. Introduction

Compromised biological samples containing degraded DNA are frequently encountered in forensic contexts. DNA fragmentation becomes a particularly complicating factor in the identification of human remains that have been recovered from aqueous environments. Environmental factors, such as heat, can facilitate bacterial growth and DNase activity, which accelerate this DNA destruction\textsuperscript{1}. In aqueous environments, water can further perpetuate DNA fragmentation, mainly through the hydrolytic disruption of the chemical bonds that uphold molecular structure\textsuperscript{2}. Heightened temperatures only serve to increase the reaction rates of the destructive chemical reactions that take place in aqueous environments, making for a substantially diminished likelihood of recovering adequate DNA profiles. In addition to the detrimental effects of heat, moisture, and environmental factors, using bone as the source of DNA can be an added challenge.

Because human remains that have been submerged in aqueous environments for extensive lengths of time are often stripped of their soft tissue components by the time they are recovered, DNA analysis of the skeletal remains is commonly relied upon for identification efforts. Although the hard tissue structure of skeletal remains provides them with a higher resiliency to degradation, DNA recovery from bone can still be impaired when long-term heat exposure, extensive submersion times, and surrounding environmental factors are at play\textsuperscript{3}. These factors particularly pose a threat to the integrity of nuclear DNA (nDNA), due to its low copy number per cell, linear structure, and cellular location\textsuperscript{4}. Alternatively, mitochondrial DNA (mtDNA) serves as a more reliable option for human identification because of its high copy number per cell, circular structure that makes it less susceptible to strand cleavage, and more protected cellular location\textsuperscript{4}. The
robustness of mtDNA allows it to be recoverable in circumstances where nDNA recovery is no longer a possibility.

Although mtDNA has a lower power of discrimination than nDNA due to its matrilineal inheritance, meaning the same mtDNA profile belongs to all members of a maternal lineage, it is still an exceptionally useful human identification tool when able to be compared alongside familial reference profiles. Given the usefulness of mtDNA, it is worthwhile to establish an effective method of mtDNA recovery, specifically from compromised bone samples that have been subjected to aqueous environments and high cumulative temperatures. Although previous studies have explored the retrieval of DNA from dry bone, there is limited research on techniques geared towards bone samples that have been subjected to long-term underwater submersion. This study aims to contribute to existing knowledge by establishing an effective methodology for mtDNA recovery from waterlogged, degraded bone samples through a comparison of aquatic systems, bone types, extraction methods, and accumulated degree day (ADD) intervals.

The effects of environmental exposure on submerged bone samples are better represented by ADD intervals as an independent variable than time or temperature alone. ADD values, which are the sum of the average temperature of each day, can be tracked in order to collect submerged bone samples at specific ADD increments. The collection of bone specimens throughout a wide range of ADD intervals allows for an estimate of the maximum ADD value at which mtDNA profile recovery is still possible. Prior to determining whether the mtDNA sequence can be successfully profiled, the DNA must be extracted from the bone samples and the extract quality subsequently assessed. By comparing the efficacy of two different extraction methods, the method that best retrieves high quantities of high quality DNA products can be determined. Selecting the best extraction method is a crucial part of developing an effective mtDNA recovery protocol, given
that the method must be able to work effectively with various severities of damaged DNA, in addition to removing the PCR inhibitors commonly seen with waterlogged samples\(^9\). Organic, phenol-chloroform extraction methods are widely used for their ability to yield DNA from complicated samples, including bone\(^10\). However, organic extraction methods are not always able to remove the PCR inhibitors that can render DNA profile recovery unsuccessful\(^11\). Alternatively, magnetic bead-based methods, such as the ChargeSwitch\(^\text{®} \) gDNA Plant Kit, have also shown success with bone and with effectively removing PCR inhibitors\(^12\). Determining which extraction technology works best with this unique sample type can be achieved through a comparison of these two methods.

Because submerged bone samples are recovered from a variety of aquatic environments, it would also be worthwhile to compare the utility of these extraction methods using two different freshwater aquatic locations. Assessing mtDNA yield for samples submerged in a lentic environment, where the water is stagnant, as well as a lotic environment, which contains running water, will reveal whether the proposed DNA recovery methods can be applied to samples associated with both ecosystems\(^13\). As part of ensuring that a proposed methodology could be used with a variety of bone specimens, a comparison could also be made between bone types. For submerged human remains, prior research shows that the bones of the torso, which include the ribs and scapulae, are last to disarticulate from one another\(^14\). A comparison between rib and scapula bone samples would not only demonstrate whether the mtDNA recovery method is able to be used with both bone types, but it would also help determine which bone type has the greatest mtDNA yield and should therefore be prioritized for analysis.

In this study, mtDNA yield will be assessed in a comparison of organic and ChargeSwitch methods, rib and scapula bone types, freshwater lotic and lentic environments, and across a wide
range of ADD intervals. As one objective of this research, quantitative PCR (qPCR) will be used to assess mtDNA quantity and quality for each group of the independent variables and at each ADD increment. This quantitation method offers the advantage of only reporting the amount of amplifiable DNA present, which is important for downstream PCR applications\(^\text{15}\). Quantitative PCR can also be utilized for a more extensive estimation of DNA quality, by calculating the degradation index. To calculate the degradation index, the quantitation value of smaller amplicons generated using one primer set is divided by the quantitation value of larger amplicons generated using a different primer set\(^\text{16}\). When this ratio is calculated at each ADD interval, it can be used to determine at which ADD increment extensive DNA degradation occurs, marked by the absence of the larger amplicon as a result of strand cleavage and fragmentation\(^\text{16}\).

After utilizing qPCR for assessing mtDNA quantity and quality for both extraction methods, only the samples from the most effective extraction method will be utilized during mitochondrial sequencing. During the sequencing portion of this study, a region of the mitochondrial displacement loop (D-loop) will be targeted. The D-loop is a favorable portion of the mitochondrial genome to use for identification purposes because it contains hypervariable regions, which can differ greatly between individuals, thereby serving as an identifying marker\(^\text{17}\). This identifying marker becomes useful upon comparison with a familial reference profile, from which an identification of the remains can be made. Mitochondrial sequencing using the D-loop specific primers will be considered successful if there is a full or partial recovery of a sequence, whereas the complete absence of a sequence will be considered unsuccessful.

In its entirety, this multifaceted study will demonstrate which of two different extraction methods is most effective at recovering high quantities of high quality DNA, whether the proposed method works best with lotic or lentic aqueous environments, whether the proposed method works
best with scapula or rib bone specimens, and at which ADD interval mtDNA sequence recovery is no longer possible. To draw these conclusions, the DNA extracts of the bone samples will first be subjected to agarose gel electrophoresis as a preliminary qualitative assessment, followed by qPCR as both a quantitative and qualitative assessment, and lastly mitochondrial sequencing of only the extracts from the most effective DNA extraction method. The ultimate goal of this research is to develop a workflow that will assist with identifying unidentified human remains by recovering mitochondrial profiles that are suitable for comparison alongside familial reference profiles.

2. Materials & Methods

2.1 Sample collection and preparation

In preparation for this study, 225 porcine scapulae and 225 ribs were acquired from a butcher. Beginning in November 2016, 19 cages, each containing five ribs and five scapulae, were submerged in Henleys Lake, a freshwater lake in Crozet, VA. In November 2017, 24 additional cages, each containing five ribs and five scapulae, were submerged in the James River, a freshwater river at the Rice Rivers Center, Charles City, VA. Each cage represented one collection at both of the locations, making for 19 total collections at the Henleys lake location and 24 total collections at the Rice Rivers location, which were collected approximately every 250 ADD. In this study, only collections from every 500 ADD up to 4500 ADD were used. The bone samples taken from each collection were extracted using two different extraction methods: a magnetic bead-based ChargeSwitch® gDNA Plant Kit (Life Technologies, Grand Island, NY, USA) extraction method, and an organic phenol-chloroform extraction method used in accordance with Iyavoo, et al.,
2013\textsuperscript{12}. With all 88 Henleys Lake samples being extracted twice, once with the ChargeSwitch method and once with the organic method, and with all 89 James River samples being extracted twice in the same fashion, a total of 354 DNA extracts were created for analysis. Samples exhibiting low DNA yields in previous applications were purified using the DNeasy\textsuperscript{®} PowerClean Pro Cleanup Kit (QIAGEN\textsuperscript{TM}, Hilden, Germany).

2.2 Gel electrophoresis of DNA extracts

As a preliminary assessment of extract quality, all 354 extracts were subjected to agarose gel electrophoresis using a 0.8\% agarose gel and 1X TAE buffer. Purple 6X gel loading dye (VWR\textsuperscript{TM}, Radnor, PA, USA) and ethidium bromide were utilized for band visualization. For each reaction, 2 µL of the purple 6X gel loading dye and 5 µL of the 1X TAE buffer were combined with 5 µL of the DNA extract. The GeneRuler 1kb DNA Ladder (Thermo Fisher Scientific\textsuperscript{TM}, Waltham, MA, USA) was used with all gels. The appearance of each of the resulting bands was used to assess the quality of each DNA extract, through the use of a five-point scale. A quality score of “1” was indicative of the poorest DNA quality, represented by a complete absence of a band, while a quality score of “5” was characteristic of the highest quality DNA quality, exhibited by a bright band or smear with a short range of fragment sizes, suggestive of minimal strand fragmentation. Quality scores of “2,” “3,” or “4” were used to denote varying degrees of intermediate quality, characterized by smears of various lengths, with large fragment sizes ranges being suggestive of extensive strand fragmentation. Chi-square tests were used to compare the quality scores between the variables of bone type, aqueous location, and extraction method.

2.3 Quantitative PCR of the cytochrome B and D-loop mitochondrial regions
For a more comprehensive assessment of the quantity and quality of mtDNA, specifically, the 354 sample extracts were subjected to two rounds of qPCR: one round using a primer set that resulted in a smaller amplicon and one round using a different primer set that resulted in a larger amplicon. Both primer sets were custom ordered from Integrated DNA Technologies™ (IDT™, Coralville, IA, USA) and diluted to a working concentration of 10µM. Each extract was first quantified using a primer set specific for a region on the cytochrome B mitochondrial protein sequence, which generated a small amplicon of 148 bp in length. These primer sequences were obtained from and used for analysis in accordance with Lopez-Oceja et al., 201618. The forward and reverse sequences for the cytochrome B primer set are as follows:

L15601 5′-TACGCAATCCTACGATCAATTCC-3′

H15748 5′GGTTGTCCTCCAATTCATGTTAG-3′

Each sample was then quantified using a primer set specific for a region within the D-loop of the mitochondrial genome, which generated a large amplicon of 531 bp in length. These primer sequences were obtained from and used for analysis in accordance with Montiel-Sosa et al., 200019. The forward and reverse sequences for the D-loop primer set are as follows:

L15592 5′-AACCCTATGTACGTCGTGCAT-3′

H16124 5′-ACCATTGACTGAATAGCACCT-3′

Regardless of primer set used, high molecular weight, high purity genomic pig DNA (SigmaAldrich™, St. Louis, MO, USA) with a stock concentration of 1 mg/mL was used to create the serial dilutions that served as the standard curve, used in duplicate (50, 20, 10, 1, 0.5, 0.1, 0.05, and 0.01 ng/µL). PerfeCTa SYBR Super Mix with ROX (VWR™, Radnor, PA, USA) was used as the fluorescence source for all of the qPCR assays. Within each qPCR reaction, 2 µL of the
sample or standard were combined with 6.25 µL of the SYBR reagent, 0.25 µL of the respective forward primer, 0.25 µL of the respective reverse primer, and 3.75 µL of water, in accordance with Seashols-Williams et al., 2018. Upon the completion of all qPCR reactions, a ratio of the quantitation value recovered using the smaller amplicon cytochrome B primer set, divided by the quantitation value recovered using the larger amplicon D-loop primer set, was then calculated for each sample. The value of this degradation index (DI) was used as an assessment of mtDNA strand integrity, and therefore DNA quality. A DI value of less than one was regarded as an indication of intact DNA, with a value between 1-10 suggesting mild to moderate degradation and a value above 10 suggesting extreme degradation. The qPCR results, consisting of the quantitation values for each sample using each primer set, as well as the degradation index for each sample, were analyzed using two statistical models. Firstly, an analysis of variance (ANOVA) was used, where the quantitative variable ADD was not included in the statistical model, followed by the use of an analysis of covariance (ANCOVA), where the quantitative variable ADD was included in the statistical model. These statistical analyses were performed using R v 4.0.3, through which mtDNA yield was compared between bone type, aqueous location, extraction method, and ADD interval; this process was also repeated with degradation index as the dependent variable.

2.4 Mitochondrial sequencing of ChargeSwitch extracts

Through determining that the ChargeSwitch extraction method was most effective at recovering high mtDNA yields during a comparison of qPCR results from each of the extraction methods, the ChargeSwitch extracts alone were selected for mitochondrial sequencing, making for a total of 176 samples that were sequenced. Prior to Sanger sequencing, the qPCR quantitation values were utilized to dilute the mtDNA quantities to an appropriate concentration of 1 ng/µL for PCR amplification. During PCR, the same D-loop specific primers and reaction volumes that were
utilized during qPCR were used, in addition to the thermalcycling protocol designed for use with this primer set, provided by Montiel-Sosa et al., 2000. Following PCR, the amplification products were subjected to 1.6% agarose gel electrophoresis to verify successful amplification, again utilizing 1X TAE buffer alongside purple 6X gel loading dye and ethidium bromide for visualization. Each sample loaded onto the gel was first combined with 7 µL of 1X TAE buffer and 2 µL of 6X purple gel loading dye, with 3 µL being the sample volume. The PCR products were then purified using the ExoSAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific, Waltham, MA, USA), during which 2.5 µL of PCR product were combined with 2.5 µL of the ExoSAP-IT reagent and subjected to the thermalcycling parameters described in Singh et al., 2011. The purified ExoSAP-IT products were used for mitochondrial sequencing with the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, Waltham, MA, USA), a sequencing kit which utilizes mock DNA alongside a universal M13 primer set as a positive control. Within each sample reaction, 1 µL of cleaned PCR product was combined with 2.25 µL of nuclease free water (Promega, Madison, WI, USA), 1 µL of 5X BigDye sequencing buffer, 0.5 µL BigDye reagent, and 0.25 µL of either the aforementioned forward D-loop specific primer or the reverse D-loop specific primer. Each sample was subjected to mitochondrial sequencing twice, once using only the forward D-loop primer, and once using the reverse D-loop primer. The thermalcycling parameters used with the BigDye sequencing protocol, in addition to the subsequent BigDye product clean-up protocol and sequencing parameters using the 3130 Genetic Analyzer (Applied Biosystems, Foster City, MA, USA), were also performed as described by Singh et al., 2011. Upon analysis of the sequencing results, a full or partial mitochondrial sequence was considered a successful recovery of a mtDNA profile whereas the complete absence of a sequence was regarded as an unsuccessful recovery of a profile. "Successful
recovery of sequence" and "unsuccessful recovery of sequence" were utilized as categorical variables in a multiple logistic regression, performed in R v 4.0.3\textsuperscript{21}. For all statistical analyses performed in this study, p-values lower than a significance value of $\alpha = 0.05$ were considered statistically significant.

3. Results & Discussion

Upon performing agarose gel electrophoresis as a preliminary quality assessment of the DNA extracts, Chi-square tests were used to analyze the quality scores that were assigned on a five-point scale (Figure 1). The results of these Chi-square tests indicated that there was no significant relationship between location and gel quality score (p-value = 0.4296). However, there was a significant relationship between gel score and bone (p-value = 1.7e-06). Rib as a bone sample type had significantly higher gel quality scores (scores of 4 and 5) than the scapula bone type, which had significantly lower gel quality scores (scores of 1 and 2). Method was also found to have a significant relationship with gel score (p-value = 0.01176). The organic extraction method had significantly higher gel quality scores (scores of 4 and 5) than the ChargeSwitch method, which had significantly lower gel quality scores (scores of 1 and 2). However, the assignment of gel quality scores is a subjective process and considers the total DNA extract, meaning it is not necessarily representative of the quality of mtDNA alone.

The quantification values obtained from performing qPCR, as well as the calculated degradation index values, were analyzed using a three-way ANOVA, followed by an ANCOVA to incorporate the numerical ADD variable (Table 1). For the D-loop quantification values, there were no significant interactions between any of the variables observed in either of the statistical models. As for the ANCOVA results using the D-loop amplicon quantitation values as the
dependent variable, method did have a significant effect on the amount of mitochondrial DNA recovered from the bone samples (Figure 2). The ChargeSwitch extraction method resulted in significantly higher quantitation values than the organic method; for this reason, only samples extracted using the ChargeSwitch method were selected for mitochondrial sequence analysis. Given that bone is largely comprised of collagen and calcium, both of which can act as PCR inhibitors, the ChargeSwitch assay may have performed more optimally than the organic method because it was able to remove these inhibitors while the organic method was not\textsuperscript{23}. Prior studies show that organic extraction methods are not effective at entirely removing PCR inhibitors, including collagen and calcium compounds, specifically\textsuperscript{11,24}. Additionally, the reagents used in organic extractions such as phenol and isopropyl or ethyl alcohol, serve as PCR inhibitors themselves, if not sufficiently removed from the final extract product\textsuperscript{25}. Alternatively, the ChargeSwitch method relies on bead-based chemistries, assays which have demonstrated a successful removal of PCR inhibitors in previous research\textsuperscript{26,27,28}. Neither location nor bone type had a significant effect on the quantity of D-loop amplicons recovered from the bone samples (Figure 3, Figure 4). Lastly, ADD interval did have a significant effect on the amount of mitochondrial DNA recovered from the bone samples; it was found that as ADD increases, the quantitation value of D-loop amplicons decreases (Figure 5). More precisely, with each one ADD unit increase, the D-loop amplicon quantitation value decreases by 0.000454 ng.

The ANOVA and ANCOVA statistical tests were repeated for the degradation index dependent variable (Table 1). Significant interactions were found to exist between all three variables and between all two-variable combinations, which was true for both the ANOVA and ANCOVA results. Of the results from the ANCOVA statistical model, it was found that method did not have a significant effect on DI value (Figure 6). However, it is possible that the significant
interactions between the variables may be masking any significant effects that method may have had on DI. Location was found to have a significant effect on the DI value, in that samples originating from Henley Lake had higher DI values (mean DI = 66.765) than the samples originating from the James River (mean DI = 5.893) (Figure 7). This could be explained by the fact that Henley Lake is a stagnant aquatic environment, where bacterial proliferation is more prominent than running bodies of water like the James River. Thus, with an increase in bacterial growth, the DNA would be subject to a more extensive process of bacterial degradation.

Bone type did have a significant effect on the DI value, in that using rib as the bone sample type resulted in significantly higher DI values (mean DI = 71.513) than the scapula bone type (mean DI = 4.651) (Figure 8). One possible explanation for rib exhibiting more extensive degradation than scapula is that rib is fragile, making it less able to preserve mtDNA quality as the hard tissue degrades. Alternatively, some regions of the scapula are much thicker and denser, allowing for better protection of the DNA within. Lastly, ADD interval was not found to have a significant effect on the DI values of the bone specimens (Figure 9). In some graphical representations of these data, there is not a visible difference between the distributions of the variable groups. When this is the case for variable groups that were shown to be significantly different from one another, such as with rib vs. scapula or Henleys Lake vs. James River for the DI variable, it is possible that outliers may be contributing to this statistical significance.

Upon performing mitochondrial sequencing, a multiple logistical regression statistical analysis was used to determine from which samples a viable sequence was able to be recovered (Table 2). It was shown that location was not a significant predictor of whether a sequence would be successfully recovered (Figure 10). Because only samples extracted using the ChargeSwitch method were analyzed, the method variable was not compared during the sequence analysis.
portion of this study. Bone type was found to be a significant predictor of whether or not a sequence would be successfully recovered, in that rib samples were more likely to yield a mitochondrial sequence than scapula samples (Figure 11). This may be attributable to rib being characteristically rich in red bone marrow, which takes part in the production of the leukocytes that mtDNA can be recovered from\(^3\). The marrow-rich composition of rib bones is likely why it is commonly selected as the sample type when recovering DNA from bone samples in similar applications\(^3\). As for the ADD numerical variable, it was found that ADD was a significant predictor of whether or not a sequence would be successfully recovered (Figure 12). It is estimated that with each increase of a 500 ADD interval, it is 40% less likely that a mitochondrial sequence will be successfully recovered. Single-base differences between the sample sequences and the \textit{Sus scrofa} mitochondrial reference genome (GenBank NC_000845.1) were observed upon alignment. Commonly observed mutations are displayed in Tables 3-5; all other mutations are not listed, as they were unique and not observed in more than one sample. An example of the sequencing analysis used to visualize single-base differences between the sample and reference genomes is shown in Figure 13.

The observed decrease in the likelihood of mtDNA sequence recovery as ADD increases can be attributed to the damaging effects of heat, moisture, and environmental conditions on DNA integrity\(^1,2\). With each prolonging of the samples’ exposure to these elements, the mtDNA will continue to become fragmented and sequence recovery will prove more difficult. Additionally, these harsh environmental factors likely explain why no mtDNA quantitation values or sequences were obtained for many of the samples. Aside from degradation, PCR inhibition may have also been a complicating factor in recovering qPCR values and in generating amplicons suitable for Sanger sequencing. A multitude of possible amplification inhibitors may have been present in the reactions, including those originating from the bone tissue itself, environmental debris, and, in the
instance of organic extractions, the reagents used\textsuperscript{25}. From bone, collagen is known to adhere to the template DNA and interfere with the polymerase’s ability to interact with the strand, while calcium ions participate in competitive inhibition with polymerase cofactors, like magnesium ions, effectively hindering the polymerase’s reaction efficiency\textsuperscript{25,35,36}. Environmental inhibitors like the humic acid characteristically found in soil, including the soils of aqueous environments, can negatively affect amplification efficiency by interfering with both the template DNA and the polymerase enzyme itself\textsuperscript{25}. Lastly, with using organic extraction reagents, alcohols such as ethanol are known PCR inhibitors, while phenol degrades the DNA polymerase and hinders the reaction\textsuperscript{25}. The absent or low mtDNA yields seen in this study are likely attributable to a combination of both DNA degradation and PCR inhibition.

4. Conclusion

Due to the multifaceted nature of this study, these experimental results will provide forensic laboratories with a thorough understanding of how best to recover mitochondrial DNA from degraded bone samples. It was determined that the ChargeSwitch extraction method was most effective at recovering significantly higher quantities of high-quality mtDNA, specifically segments of the D-loop region, which indicates that this would be a worthwhile method to implement into laboratories’ mitochondrial workflow. Additionally, the rib bone type was found to be more likely to yield a mitochondrial sequence than scapula, suggesting that using rib samples should be prioritized over scapula samples when using bone for mtDNA efforts. The effectiveness of the proposed methods was not found to significantly differ between the lentic Henley Lake and the lotic James River aquatic environments, allowing for the conclusion that this technique can be used effectively with both lentic and lotic environments. Lastly, in determining the effect of ADD
interval on the ability of these methods to successfully recover mtDNA, the knowledge that the likelihood of recovering a mitochondrial sequence decreases by 40% with each 500 ADD increase provides scientists with more realistic expectations of successfully sequencing the sample, depending on how degraded that sample may be. Although an increase in ADD did not significantly affect mtDNA quality, in that it did not alter degradation index, it did affect the quantity of mtDNA recovered, which substantially impacts the ability to recover a mitochondrial sequence. Future research using these samples will involve a comparison between nuclear, mitochondrial, and DNA recovery from the same bone types, aquatic environments, extraction methods, and ADD intervals. In its totality, the results of this research can be used to advance the sample processing capabilities of laboratories which handle mtDNA sequencing of unidentified remains. In addition to the knowledge that this study offers to forensic laboratories, it also proposes the use of a promising mtDNA workflow involving optimized extraction techniques, qPCR parameters, and mitochondrial sequencing protocols.
References


Table 1. ANCOVA p-values for D-loop amplicon quantitation and degradation index.

<table>
<thead>
<tr>
<th></th>
<th>D-loop quantitation p-value</th>
<th>Degradation index p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADD</strong></td>
<td>4.06E-10</td>
<td>0.66416</td>
</tr>
<tr>
<td><strong>Method</strong></td>
<td>0.000835</td>
<td>0.05557</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td>0.892418</td>
<td>0.04455</td>
</tr>
<tr>
<td><strong>Bone</strong></td>
<td>0.202302</td>
<td>0.04547</td>
</tr>
<tr>
<td><strong>Method:Location</strong></td>
<td>0.206468</td>
<td>0.00440</td>
</tr>
<tr>
<td><strong>Method:Bone</strong></td>
<td>0.377961</td>
<td>0.00347</td>
</tr>
<tr>
<td><strong>Location:Bone</strong></td>
<td>0.502247</td>
<td>0.00227</td>
</tr>
<tr>
<td><strong>Method:Location:Bone</strong></td>
<td>0.572289</td>
<td>0.0000463</td>
</tr>
</tbody>
</table>
Table 2. *Multiple linear regression results for sequence recovery success.*

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location: River</strong></td>
<td>1.01225</td>
<td>0.50422</td>
<td>2.0322</td>
<td>-0.0122</td>
<td>0.3541855</td>
<td>0.034</td>
<td>0.972570</td>
</tr>
<tr>
<td><strong>Bone: Scapula</strong></td>
<td>0.25903</td>
<td>0.12505</td>
<td>0.5180</td>
<td>-1.3508</td>
<td>0.3609839</td>
<td>-3.742</td>
<td>0.000183</td>
</tr>
<tr>
<td><strong>ADD</strong></td>
<td>0.99925</td>
<td>0.99896</td>
<td>0.9995</td>
<td>-0.0008</td>
<td>0.0001421</td>
<td>-5.268</td>
<td>1.38E-07</td>
</tr>
</tbody>
</table>
### Table 3. Common single-base deviations from Sus scrofa mitochondrial reference genome, pt. 1.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Samples with mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitution from T to C at 241 (241C)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR4C12CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C15CS, HLS3C18CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS3C6CS, RRS4C6CS, HLS1C12CS, HLS1C8CS, RRS5C9CS, RRS1C9CS, RRS4C2CS, RRS5C9CS, RRS1BCS</td>
</tr>
<tr>
<td>Substitution from C to T at 279 (279T)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS4C6CS, HLS1C8CS, RRS4C2CS, RRS5C9CS, RRS1BCS</td>
</tr>
<tr>
<td>Substitution from A to G at 294 (294G)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS3C6CS, RRS4C6CS, RRS5C2CS, RRS5C9CS, RRS1BCS</td>
</tr>
</tbody>
</table>
Table 4. Common single-base deviations from *Sus scrofa* mitochondrial reference genome, pt. 2.

<table>
<thead>
<tr>
<th>Change</th>
<th>Samples with Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitution from C to T at 306 (306T)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS2BCS, RRS3C4CS, RRS3C6CS, RRS4C6CS, RRS5C2CS, HLS1C8CS, RRS1C9CS, RRS4C2CS, RRS5C9CS, RRS1BCS, RRS2C6CS</td>
</tr>
<tr>
<td>Substitution from C to T at 323 (323T)</td>
<td>HLR1BCS, HLR1C4CS, HLR1C15CS, HLR2C4CS, HLR2C8CS, HLR2C10CS, HLR2C13CS, HLR2C15CS, HLR3CBS, HLR3C2CS, HLR3C4CS, HLR3C6CS, HLR3C8CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2BCS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS3C6CS, RRS4C6CS, RRS5C2CS, HLS1C8CS, RRS1C9CS, RRS4C2CS, RRS5C9CS, RRS1BCS, RRS2C6CS</td>
</tr>
<tr>
<td>Substitution from C to T at 390 (390T)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS3C6CS, RRS4C6CS, RRS5C2CS, HLS1C8CS, RRS1C9CS, RRS4C2CS, RRS5C9CS, RRS1BCS, RRS2C6CS</td>
</tr>
</tbody>
</table>
Table 5. *Common single-base deviations from Sus scrofa mitochondrial reference genome, pt. 3.*

<table>
<thead>
<tr>
<th>Change</th>
<th>Samples with Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substitution from A to G at 575 (575G)</td>
<td>HLR1BCS, HLR1C4CS, HLR2C4CS, HLR2C10CS, HLR3BCS, HLR3C2CS, HLR3C4CS, HLR3C10CS, HLR4BCS, HLR4C4CS, HLR4C10CS, HLR5BCS, HLR5C2CS, HLS1BCS, HLS2BCS, HLS3C8CS, HLS3C10CS, HLS3C12CS, HLS3C15CS, HLS4C13CS, HLS5C2CS, RRR2C4CS, RRR3C4CS, RRR4C4CS, RRS1C2CS, RRS2BCS, RRS2C2CS, RRS3BCS, RRS3C4CS, RRS3C6CS, RRS4BCS, RRS4C6CS, RRS5C2CS, HLS1C8CS, RRR1C9CS, RRS1C9CS, HLS4C2CS, HLS5C9CS, RRS1BCS</td>
</tr>
<tr>
<td>Substitution from T to C at 405 (405C)</td>
<td>RRS1C2CS, RRS3C6CS</td>
</tr>
<tr>
<td>Substitution from C to T at 452 (452T)</td>
<td>RRS1C2CS, RRS3C6CS</td>
</tr>
<tr>
<td>Substitution from T to C at 560 (560C)</td>
<td>RRS4C6CS, RRS1C9CS</td>
</tr>
<tr>
<td>Substitution from T to C at 181 (181C)</td>
<td>HLR3C12CS, HLS1C8CS, RRR1C9CS, RRS1C9CS</td>
</tr>
<tr>
<td>Substitution from G to A at 786 (786A)</td>
<td>HLR2BCS, RRS2C13CS, RRS5C15CS</td>
</tr>
<tr>
<td>Substitution from G to A at 853 (853A)</td>
<td>HLR2BCS, HLR2C6CS, RRS2C13CS, RRS5C15CS</td>
</tr>
<tr>
<td>Substitution from G to A at 870 (870A)</td>
<td>HLR2BCS, RRS2C13CS, RRS5C15CS</td>
</tr>
<tr>
<td>Substitution from G to A at 897 (897A)</td>
<td>HLR2BCS, RRS2C13CS</td>
</tr>
<tr>
<td>Substitution from A to G at 935 (935G)</td>
<td>HLR2BCS, RRS2C13CS</td>
</tr>
<tr>
<td>Insertion of C between 601 and 602 (601.1C)</td>
<td>RRR3BCS, RRS1C9CS</td>
</tr>
<tr>
<td>Insertion of C between 192 and 193 (192.1C)</td>
<td>RRR1C9CS, RRS1C9CS, RRS5C9CS</td>
</tr>
</tbody>
</table>
Figure 1. Exemplary 0.8% agarose gel image of bands paired with sample names and corresponding quality scores. (HL: Henley Lake; R#: Rib, replicate number; S#: Scapula, replicate number; C#: Collection number; OR: Organic extraction; (#): Gel quality score number.)
Figure 2. Box plots depicting the D-loop amplicon quantitation values for each method. (p-value = 0.000835).
Figure 3. *Box plots depicting the D-loop amplicon quantitation values for each location.* (p-value = 0.892418).
Figure 4. Box plots depicting the D-loop amplicon quantitation values for each bone type. (p-value = 0.202302).
Figure 5. Scatterplot with standard error of mean (SEM) bars, depicting the average D-loop amplicon quantitation values for each ADD interval. (p-value = 4.06E-10).
Figure 6. Box plots depicting the degradation index (DI) values for each method. (Note: the two highest outliers have been removed for better visibility). (p-value = 0.05557).
Figure 7. Box plots depicting the degradation index (DI) values for each location. (Note: the two highest outliers have been removed for better visibility). (p-value = 0.04455).
Figure 8. *Box plots depicting the degradation index (DI) values for each bone type. (Note: the two highest outliers have been removed for better visibility). (p-value = 0.04547).*
Figure 9. Scatterplot with standard error of mean (SEM) bars, depicting the average degradation index (DI) values for each ADD interval. (Note: the two highest outliers have been removed for better visibility). (p-value = 0.66416).
Figure 10. *Bar graph depicting percentage of mitochondrial sequences recovered for each location.*
Figure 11. *Bar graph depicting percentage of mitochondrial sequences recovered for each bone type.*
Figure 12. Scatterplot depicting the percentage of mitochondrial sequences recovered for each ADD interval, with linear equation \( y = -0.0002x + 0.768 \) and \( R^2 = 0.8709 \).
Figure 13. Screenshot of example sequence reads visualized using ChromasPro™ software. Deviations from the reference sequence can be seen at positions 294, 306, and 323.
VITA
Kailey Babcock

EDUCATIONAL INSTITUTIONS

Virginia Commonwealth University, Department of Forensic Science Graduate Program – 2019-2021
Virginia Commonwealth University, Department of Forensic Science Undergraduate Program – 2017-2019
Virginia Commonwealth University, Honors College – 2016-2019

DEGREES & HONORS

Bachelor of Science in Forensic Science – 2019
Concentration: Biology
Minor: Chemistry
Distinction of Graduation with University Honors – 2019

ACADEMIC PROJECTS

Quantitative PCR and Sanger Sequencing of Mitochondrial DNA Recovered from Waterlogged Bone – 2019-2021
Survey of Microbial Communities in Soil Associated with Porcine Remains – 2018-2021
Forensic Body Fluid Identification Using Microbiome Signature Attribution – 2018-2019
Postmortem Hair Microbiome and Its Forensic Applications – 2018-2019

TEACHING & PROFESSIONAL EXPERIENCES

Graduate Teaching Assistantship: Forensic Science Department – 2020-2021
Research Assistantship: Dr. Baneshwar Singh’s Laboratory – 2018-2019

CONFERENCE PRESENTATIONS

Postmortem Hair Microbiome and Its Forensic Applications: American Academy of Forensic Science, Baltimore, MD – 2019