

Article

# Routine Mitogenome MPS Analysis from 1 and 5 mm of Rootless Human Hair

Lauren C. Canale <sup>1</sup>, Jennifer A. McElhoe <sup>1</sup> , Gloria Dimick <sup>2</sup>, Katherine M. DeHeer <sup>3</sup> , Jason Beckert <sup>4</sup> and Mitchell M. Holland <sup>1,2,\*</sup>

<sup>1</sup> Forensic Science Program, Department of Biochemistry & Molecular Biology, Eberly College of Science, Pennsylvania State University, University Park, PA 16802, USA

<sup>2</sup> Mitotyping Technologies, 2565 Park Center Blvd., Suite 200, State College, PA 16801, USA

<sup>3</sup> Indiana State Police, 100 North Senate Avenue, Indianapolis, IN 46204, USA

<sup>4</sup> Microtrace 790 Fletcher Drive, Suite 106, Elgin, IL 60123, USA

\* Correspondence: mmh20@psu.edu

**Abstract:** While hair shafts are a common evidence type in forensic cases, they are often excluded from DNA analysis due to their limited DNA quantity and quality. Mitochondrial (mt) DNA sequencing is the method of choice when working with rootless hair shaft fragments due to the elevated copy number of mtDNA and the highly degraded nature of nuclear (n) DNA. Using massively parallel sequencing (MPS) of the mitochondrial (mito) genome, we studied the impact of hair age (time since collection) and physical characteristics (hair diameter, medullary structure, and length of hair tested) on mtDNA recovery and MPS data quality. Hair shaft cuttings of 1 and 5 mm from hairs less than five years to 46 years of age from 60 donors were characterized microscopically. Mitogenome sequences were generated using the Promega PowerSeq™ Whole Mito System prototype kit and the Illumina MiSeq instrument. Reportable mitogenome sequences were obtained from all hairs up to 27 years of age (37 donors), with at least 98% of the mitogenome reported for more than 94% of the 74 hair samples analyzed; the minimum reported sequence was 88%. Furthermore, data from the 1 and 5 mm replicates gave concordant haplotypes. As expected, mtDNA yield decreased, mtDNA degradation increased, and mitogenome MPS data quality declined as the age of the hair increased. Hair diameter and medullary structure had minimal impact on yield and data quality. Our findings support that MPS is a robust and reliable method for routinely generating mitogenome sequences from 1 and 5 mm hair shaft samples up to 27 years of age, which is of interest to the forensic community, biological anthropologists, and medical geneticists.

**Keywords:** next generation sequencing; forensic science; hair



**Citation:** Canale, L.C.; McElhoe, J.A.; Dimick, G.; DeHeer, K.M.; Beckert, J.; Holland, M.M. Routine Mitogenome MPS Analysis from 1 and 5 mm of Rootless Human Hair. *Genes* **2022**, *13*, 2144. <https://doi.org/10.3390/genes13112144>

Academic Editors: Erin K. Hanson and Claire L. Glynn

Received: 29 October 2022

Accepted: 14 November 2022

Published: 18 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hairs are one of the most common evidence types found at crime scenes [1–3] and are a useful sample type for medical research [4–6] and archaeological investigations [7,8]. Nuclear (n) and mitochondrial (mt) DNA can both be recovered from hairs. For example, hairs in the anagen and catagen growth stages present follicular tissue, especially when forcibly removed, allowing for nDNA analysis when roots are present. Telogen hairs are easily shed, but due to the keratinization process leading to the degradation of cellular organelles and chromosomal nucleic acids, nDNA is generally too degraded for conventional analysis [9–11]. However, under certain circumstances, nDNA can be recovered and analyzed from hair shafts [10,12–14]. Along the shafts of hairs in all phases, nDNA is present in higher quantities than mtDNA, but due to its extensive fragmentation is typically not of sufficient quality for amplicon-based nDNA analysis; for example, short tandem repeat (STR) profiling [10]. Therefore, mtDNA sequencing remains the method of choice when analyzing shed hairs or small fragments of hair shafts.

Historically, mtDNA testing has focused on recovering the control region (CR) of the mitogenome using electrophoretic techniques such as Sanger type sequencing (STS) [15]. Haplotypes generated from the CR are highly informative, however, expanding to the entire mitogenome significantly increases the discriminatory power [16]. Unfortunately, sequencing the mitogenome using STS is labor-intensive, requires large quantities of DNA, and is a low-throughput method. Contrary to this, with advancements in massively parallel sequencing (MPS), analysis of the mitogenome can be performed relatively quickly with higher throughput and on samples that are routinely encountered in forensic cases, including hair shafts [17].

Microscopical analysis of hair has historically been an integral part of forensic investigations and remains an important component of hair analysis [18,19]. Hair characterization typically involves an assessment of color, size, distribution and pattern of pigment granules, the appearance and structure of the cuticle, cortex, and medulla, indications of chemical treatments, the physical condition of the shaft including diameter, and other microscopical characteristics [20,21]. Comparison to reference hairs and assessment of inclusion or exclusion remains a relatively subjective process [18]. Currently, hairs are assessed microscopically prior to DNA analysis, noting if a root is present and the growth phase of the root to determine whether nDNA profiling or mtDNA sequencing should be performed. Crime laboratories that do not perform mtDNA sequencing will either not move forward with these hairs or send them to a private laboratory that does sequence analysis.

Previous research has focused on either the mtDNA CR or on small sample sizes when performing mitogenome analysis [11,22]. In addition, the majority of the tested hairs have been longer and/or recently collected with a little background on microscopic characteristics or hair treatments. Gallimore et al. [11] observed heteroplasmy in the CR using an MPS approach in 75 hair shafts from five donors that were all tested shortly after collection and were 2 cm cuttings. Parson et al. [22] successfully recovered complete mitogenomes from hair shafts and shed hairs but tested a small sample size of ten hairs. A survey study assessing 691 casework hairs analyzed using an STS approach over many years drew comparisons between any noted physical characteristics with the resulting CR (HV1/HV2) sequence profiles [1]. They concluded that with increasing age, the likelihood of generating a full sequence profile decreased, while with increasing natural hair color and diameter, the likelihood increased. To our knowledge, a similar study has not been done using an MPS approach on the entire mitogenome.

Microscopical analysis of hairs prior to sequencing is typically not performed to assess whether morphological features contribute to the outcome. In addition, forensic laboratories often receive small hair fragments for analysis (0.1–2+ cm), with 2 cm being the typical sample size tested. Hair evidence can also be stored for years if part of a cold or historical case. The goal of the current study was to assess the quantity and quality of mitogenome MPS data from hair samples representative of those encountered in crime laboratories when considering age (since time of collection) and physical characteristics of the tested hairs (for example, hair diameter, medullary structure, and length of hair tested).

## 2. Materials and Methods

### 2.1. Sample Collection and Preparation

Head hair samples from 60 donors were collected from two operating forensic laboratories (Mitotyping Technologies, LLC, a Division of SoftGenetics, State College, PA, USA and Microtrace, LLC, Elgin, IL, USA) and an academic research laboratory (Pennsylvania State University, University Park, PA, USA). The donated hairs fell into three age categories for time since removal from head: 13 recent (R) hairs (<5 years), 24 old (O) hairs (5–27 years), and 23 older (VO) hairs (41–46 years). Twelve of the R hair samples were closer to five years from collection. The 24 O hairs had an average age of almost 14 years. Table S1, Supplementary Materials, provides the age for each hair sample.

Twelve of the R hair samples were donated for a previous hair study [11], under the Penn State University internal review board (IRB) approved protocol STUDY00000970,

and one was collected under STUDY00014305. These hairs were stored in paper envelopes at room temperature. Twelve of the O hair samples were provided by Mitotyping and were stored at  $-20^{\circ}\text{C}$  as part of proficiency testing kits. Once received from Mitotyping, the hairs were stored in paper envelopes at room temperature for approximately one month before MPS analysis. The remaining hair samples (12 O hairs and 23 VO hairs) were provided by Microtrace and were stored in plastic bags at room temperature. Once received from Microtrace, the hairs were stored in plastic bags at room temperature for approximately five days before MPS analysis.

Two cuttings were sampled from each hair, 1 mm and 5 mm in length. Cuttings were taken 1 cm from the root end when a root was present (45 of the 60 hairs) and from the proximal end at an unknown distance from the root when it was not present. The 120 cuttings were then subjected to microscopical analyses.

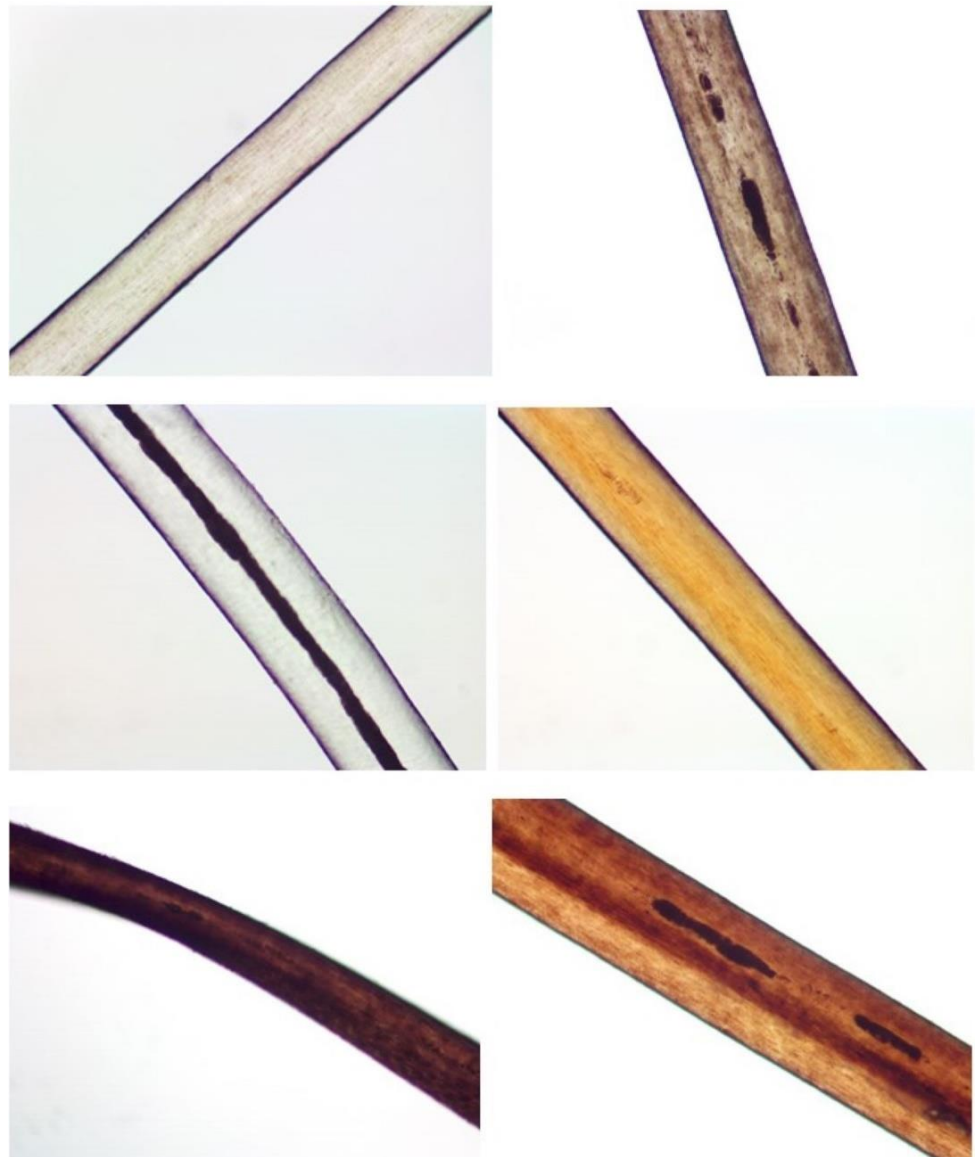
## 2.2. Microscopy

The 1 and 5 mm cuttings were analyzed using a Leica FS4000 microscope at  $200\times$  magnification. The hairs were mounted on glass microscope slides using deionized water, and photographs were taken using an MU 1000 camera and AmScope software version 3.7. Figure 1 provides representative pictures of hairs from the study. Microscopic characteristics noted were the width, the presence or absence of a medulla, and the pigmentation. The width of the hairs was measured by calibrating the AmScope software with a micrometer. If the width of the hair changed, two measurements were taken, one at the largest diameter and one at the smallest diameter, and the average was taken and used for data analysis. If a medulla was present, it was categorized as either trace, discontinuous, or continuous. Hair treatment information was mostly unknown for the hair samples, except for the hairs received from Microtrace. All microscopy data can be found in Supplemental Table S1, Supplementary Materials.

## 2.3. DNA Extraction

Extraction of DNA from each hair cutting was conducted using the method reported by Gallimore et al. [11]. Cuttings were first cleaned in 1 mL of 5% Terg-a-zyne for 20 min with sonication. Duplicate washes of 1 mL 100% ethanol and 1 mL MGW water were performed to remove the Terg-a-zyne. A 300  $\mu\text{L}$  aliquot of ATL buffer (Qiagen; Hilden, Germany), 20  $\mu\text{L}$  of 1 M DTT, and 20  $\mu\text{g/mL}$  Pro K were added to the clean cuttings and then incubated in a thermomixer (Eppendorf; Hamburg, Germany) at  $56^{\circ}\text{C}$  and 900 RPM for one hour, followed by the addition of 300  $\mu\text{L}$  of AL buffer (Qiagen; Hilden, Germany) and a second incubation at  $70^{\circ}\text{C}$  and 900 RPM for 10 min. Lysates were purified using the PrepFiler<sup>®</sup> Forensic DNA Extraction Kit (Applied Biosystems; Foster City, CA, USA) according to the manufacturers protocol with the following exceptions: 300  $\mu\text{L}$  of isopropanol and 40  $\mu\text{L}$  of elution buffer according to Gallimore et al. [11]. Modifications to the Gallimore et al. protocol included a second water wash during cleaning and the addition of a second thermomixer for room temperature incubation. Twenty hair cuttings were extracted at a time with one reagent blank.

Twenty-one (21) of the 120 hair cuttings (two R, four O, and 15 VO) were re-extracted and the testing was repeated due to suspected contamination based on mixed sites at major SNP (single nucleotide polymorphism) sites. Eleven (11) of the hairs had presumed contamination from laboratory personnel and the remaining 10 hairs had possible contamination from an unknown source. The additional testing addressed any suspected contamination concerns, and as a result, there was no effect on the findings and conclusions presented in this study.



**Figure 1.** Pictures of representative hair samples. Moving down from left to right: R\_03F\_A, O\_16\_A, O\_24\_A, R\_13\_A, R\_08F\_A.3, and O\_20\_A. All pictures were taken at 200 $\times$  magnification. R\_08F\_A.3 has an undulation and the donor of O\_24\_A has albinism.

#### 2.4. *mtq*PCR Assay

A custom mtDNA quantitative PCR assay (*mtq*PCR) described in Gallimore et al. [11] was used to quantify copies of mtDNA recovered per  $\mu$ L of DNA extract. The eight concentrations for the standard curve were 200,000, 20,000, 2000, 200, 150, 100, 50, and 10 mtDNA copies/ $\mu$ L. Copies of mtDNA per unit volume in each hair extract were quantified immediately following extraction. All quantifications were performed on an Applied Biosystems<sup>®</sup> 7500 instrument (Foster City, CA, USA) and performed in duplicate.

#### 2.5. Amplification of the Mitogenome, Library Preparation, and MiSeq Conditions

Each hair extract was amplified using the PowerSeq<sup>™</sup> Whole Mito System prototype kit from Promega (Madison, WI, USA), targeting an input of 40,000 copies of mtDNA or the maximum extract volume of 15  $\mu$ L for low-yield samples, with partial MPS results obtained from as few as 110 copies of mtDNA template (Table S1, Supplementary Materials). The kit provides for a single or dual multiplexed PCR reaction with 161 amplicons covering the mitogenome. The single multiplexed approach was used for the samples in this study.

A maximum of 23 samples were included per amplification and library preparation run (20 hair extracts, one reagent blank, one negative amplification control, and one positive amplification control) except for the final run, which included 21 hair extracts for a total of 24 samples. Amplification, purification, quantification, and normalization were according to the manufacturer's protocol. Library preparation was performed with the Illumina® TruSeq DNA PCR-Free HT kit and IDT for Illumina® TruSeq DNA UD Indexes (Illumina; San Diego, CA, USA). If samples did not meet the recommended 4 nM library concentration, a concentration of 2 nM was used for pooling. Samples that did not meet the 2 nM were not diluted prior to pooling. Two runs were pooled at 1.5 nM due to most samples being below 2 nM. The pooled libraries were spiked with 10% PhiX with a final loading concentration of approximately 7.0 pM. The prepared libraries were run on the Illumina® MiSeq with a 600-cycle v3 kit (Illumina; San Diego, CA, USA), 276 × 276 paired read ends.

## 2.6. MPS Data Analysis

The fastq files generated from the MiSeq Reporter software were analyzed using GeneMarker® HTS (SoftGenetics; State College, PA, USA), version 2.2.5.0 [23]; referred to as GM HTS. The sequences were aligned to the revised Cambridge Reference Sequence (rCRS) [24] with a custom motif file to ensure phylogenetically correct calls. SNPs present in more than 50% of the reads were considered major variants and referred to as the haplotype. Minor variants are SNPs that are in less than 50% of the reads. When assessing the minor variants, all indels (insertion/deletion) were excluded along with the following variants commonly seen when using the PowerSeq Whole Mito System: 523A, 524C, 663G, 709G, 2706A, 3010G, 7028C, 8019T, 8860A, 13760T, and 15326A. The filter settings included a minimum read coverage of 200 for the haplotype, a minimum of 40 reads for reporting minor variants, a minimum threshold of 2% for reporting minor variants, ≤10 for the allele score difference, ≤2.5 for the SNP balance ratio, and ≤5.0 for the indel balance ratio. Haplogroups were generated for each sample using HaploGrep2 (version 2.2). Sequences generated from the VO hairs contained significantly more noise in the MPS data (see Section 3.3); measured as the number of minor sites reported. As a result, a detailed analysis of the VO hairs will be the subject of a future study.

## 2.7. Statistical Analysis

Normality was tested using the Shapiro–Wilk test [25], which showed the dataset was non-normal. Therefore, the non-parametric Kruskal–Wallis test [26] with a Dunn's post hoc test [27] utilizing the Holm correction for multiple comparison testing [28] was used to evaluate the statistical significance of differences in the analyses of mtDNA yield, the MPS data quality, and the microscopical correlations (critical  $p$ -value < 0.05). The analysis of the total error rate per sample was performed using a custom R script generated for a previous study [29] and significance was also calculated using the Kruskal–Wallis test with a Dunn's post hoc test adjusted with the Holm correction (critical  $p$ -value < 0.05). A simple linear regression model was run to determine if the mtDNA yield could be predicted by the width of the hairs and if the MPS data quality could be predicted by the mtDNA yield [30]. All tests were conducted using RStudio (2022.07.1+554, RStudio Team, 2022).

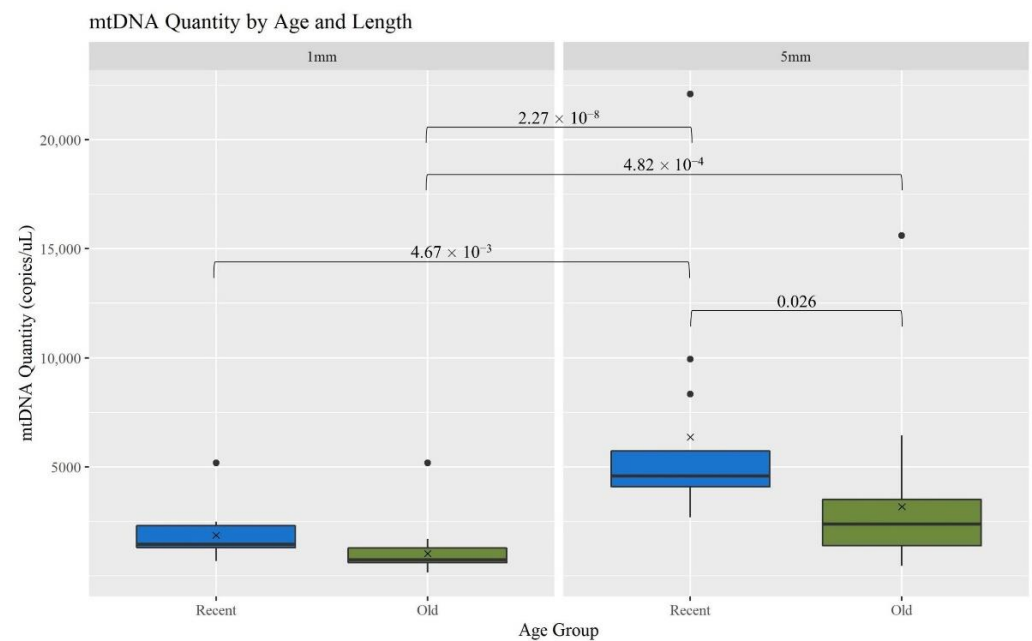
# 3. Results and Discussion

## 3.1. mtDNA Yield from Hair Shafts

The recovery of mtDNA from hair shafts (measured as copies of mtDNA/μL) was compared between two age categories [recent (R, <5 years) and old (O, 5–27 years)], both within and across the two hair cutting lengths (1 and 5 mm). Overall, mtDNA yield decreased significantly with age and length (Figure 2). When considering the 5 mm hairs, the R hairs had significantly more mtDNA recovered than the O hairs;  $p$ -value of 0.026 (Table S2, Supplementary Materials provides a list of  $p$ -values for mtDNA recovery comparisons). Comparing across the lengths, the R 5 mm hair extracts have significantly more mtDNA than all the 1 mm extracts;  $p$ -values of  $4.67 \times 10^{-3}$  for R hairs and  $2.27 \times 10^{-8}$

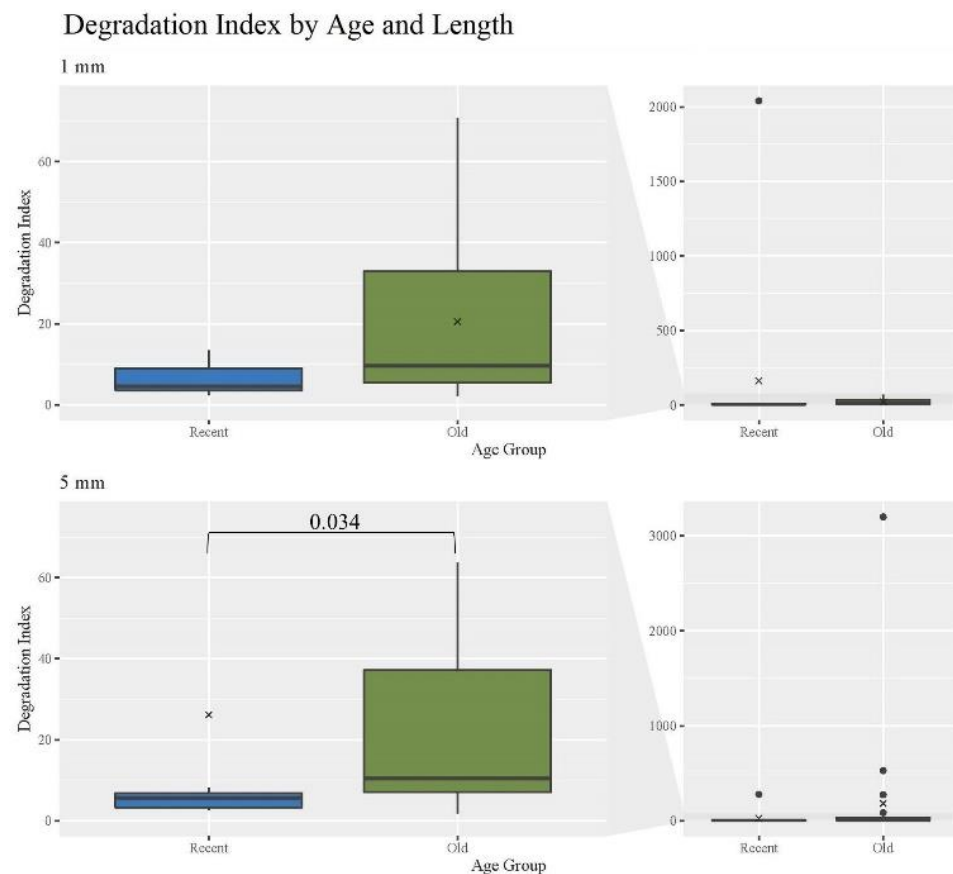


for O. The 5 mm O hair extracts also had a significantly higher mtDNA yield than the 1 mm O extracts;  $p$ -value of  $4.82 \times 10^{-4}$ . This is consistent with the expectation that as more hair sample is extracted, more mtDNA is recovered and as samples age, less DNA is available due to factors such as degradation and damage [1].



**Figure 2.** Boxplots for the comparison of mtDNA quantity for each age group and length ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ). The 'x' on the boxplots represents the mean and the line represents the median. Brackets show comparisons with significant  $p$ -values ( $p < 0.05$ ).

Various types of DNA damage can lead to degradation and the inability for targets of interest to be amplified [31] impacting the analysis of mtDNA from hair shafts. The degradation index (DI) for hair extracts was determined by taking the ratio of the small to large target quantities of each quantification replicate and calculating the mean value for each individual hair sample, with statistical comparisons performed on the mean values. For samples where the large target value was undetermined, the quantity value was artificially set to the lowest datapoint in the standard curve of 10 copies/ $\mu$ L to allow a DI to be calculated. As expected, DIs increased significantly with age (Figure 3). Consistent with expectations, the O hairs had a significantly higher DI than the R of the same length;  $p$ -value of 0.034 for R 5 mm hairs (Table S2, Supplementary Materials provides a list of  $p$ -values for DI comparisons). No other comparisons regarding age or length were significant. In addition, of particular note, the O hairs have relatively low average DIs, suggesting that hairs up to 27 years of age do not generally experience significant DNA degradation. Contrary to expectations, when comparing different cutting lengths, the O 5 mm hairs have a higher DI than all 1 mm hairs, noting that the difference is not significant. With more than double the average DI when comparing the O 1 mm and 5 mm, this suggests that the range of template lengths may exhibit greater levels of degradation as mtDNA yields increase, while lower yields reflect templates of generally the same length. Overall, these findings are consistent with the expectation that older samples will have greater levels of degradation than recent samples, but that degradation levels may increase with increasing yields of mtDNA and hairs up to 27 years of age will still yield fairly good quality DNA.



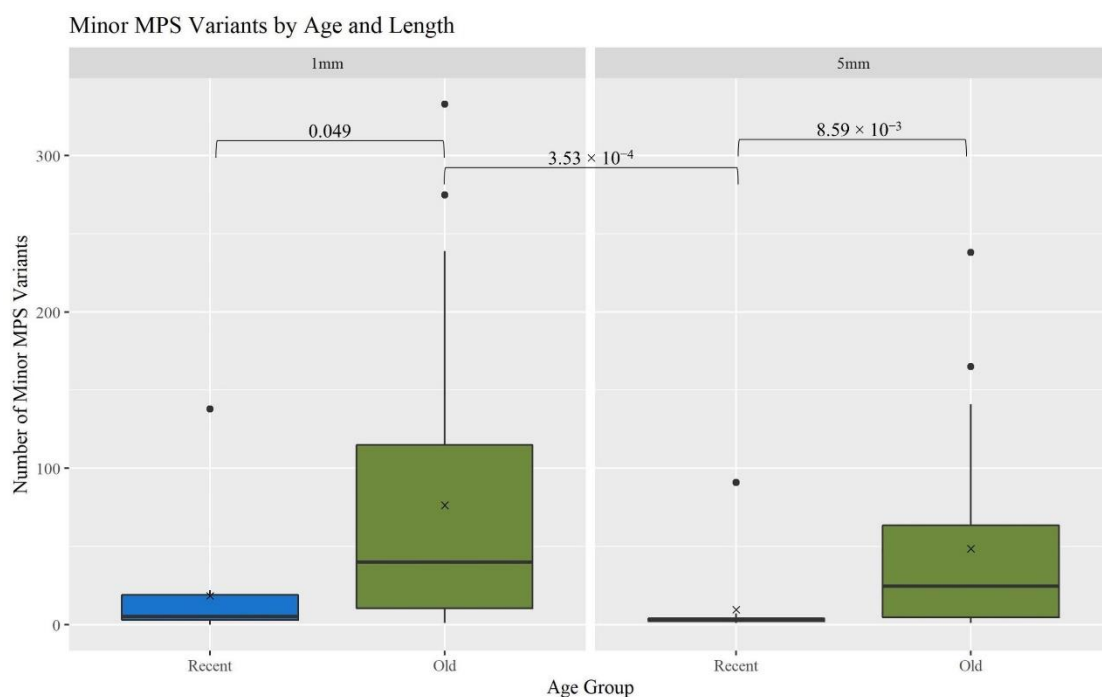
**Figure 3.** Boxplots for the comparison of degradation index for each age group and length ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ). The 'x' on the boxplots represents the mean and the line represents the median. Bracket shows comparison with significant  $p$ -value ( $p < 0.05$ ).

The quantity and quality of mtDNA is highly variable in shed hairs from different individuals and even from hairs of the same individual [10,11,32,33]. In addition, previous studies have found that mtDNA content decreases distally from the root end of a hair [10]. Most of the hair cuttings in the current study were taken approximately 1 cm from the root end and should, therefore, reflect the greatest possible mtDNA yield when compared to the remaining hair shaft. However, a number of the hairs (approximately 29% of the R and O hairs) did not have a root or were re-extracted with additional cuttings taken at greater lengths from the proximal end. While the distance from the root may impact yield and DI, hair extracts in the current study yielded more copies of mtDNA than previous studies using similar and different extraction methods regardless of where the cutting was taken. For example, Brandhagen et al. [10] recovered 43,992 copies of the mitogenome/cm from a recent hair segment taken approximately 1 cm from the root end, while the average yield for the current study was almost 405,360 copies/cm, including a R hair cutting taken at an unknown distance from the root which yielded 333,523 copies/cm, and with an overall range of yields of 36,187 to greater than two million copies/cm (Table S1, Supplementary Materials, R and O hairs). Gallimore et al. [11] utilized the same extraction method except for the modifications made (Materials and Methods Section 2.3) and generated on average between 200–300 copies/ $\mu$ L for 5 mm cuttings of recently collected head hairs. In comparison, the same size cutting yielded on average 4289 copies/ $\mu$ L in the current study; Table S1, Supplementary Materials provides the quantities for all hairs in both copies/ $\mu$ L and copies/cm. These results demonstrate that the modified extraction method is highly robust. While many factors, including natural variability, the environment, and chemical or physical changes to the hair can impact

mtDNA recovery, these results support that the age of the hair and the length of the cutting have the greatest impact on mtDNA yield.

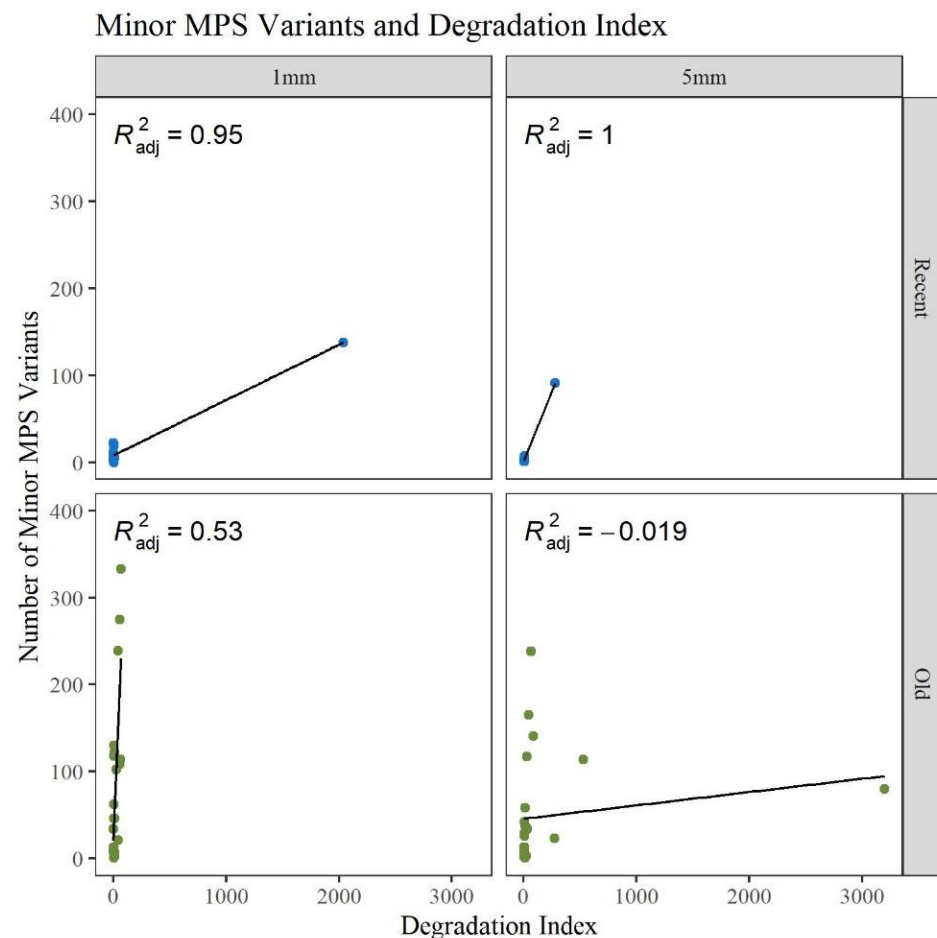
### 3.2. Assessment of MPS Sequence Data Generated from Hair Shafts

A gross assessment of MPS data quality was linked to the number and types of minor sequence variants reported by the analysis software (GM HTS), the generation of complete haplotypes, and the coverage of reads across the mitogenome. Minor variants consist of mixed sites due to heteroplasmy, damage sites, possible contamination, NUMTs (nuclear inserts of mitochondrial DNA), or random error [17]. An elevated number of minor variants may indicate the presence of DNA damage. The number of minor variants called by GM HTS increased significantly as the age of the hairs increased (Figure 4). The O hairs have significantly more minor variants when compared to the R hairs within the two lengths (1 mm:  $p$ -value of 0.049; 5 mm:  $p$ -value of  $8.59 \times 10^{-3}$ ) and when comparing across the two lengths ( $p$ -value of  $3.53 \times 10^{-4}$ ) (See Table S2, Supplementary Materials for all  $p$ -values). Given that the number of minor variants and the DI relate to the amount of DNA damage in a sample, Figure 5 provides a comparison of the number of minor variants reported to the DI. For the R hairs, the DI and the number of minors are both low and cluster near the origin as they have less apparent damage. As age increases, the number of minor variants and the DI increase. For the O hairs, the number of minor variants increases more than the DI, resulting in a cluster along the y-axis. The adjusted  $R^2$  values in Figure 5 along with the residual standard error and F-statistic values presented in Table S3, Supplementary Materials, indicate that the DI is a good predictor for the number of minor variants for the longer, younger (R) hairs, but not the older (O) hairs. It was expected that as the DI increases, the number of minor variants would also increase as low template and low-quality samples have higher error and base substitution rates [29,31,34]. Due to the O hairs not showing much of an increase in the DI in comparison to the increase in minor variants, the amount of template DNA added to the reaction was assessed.



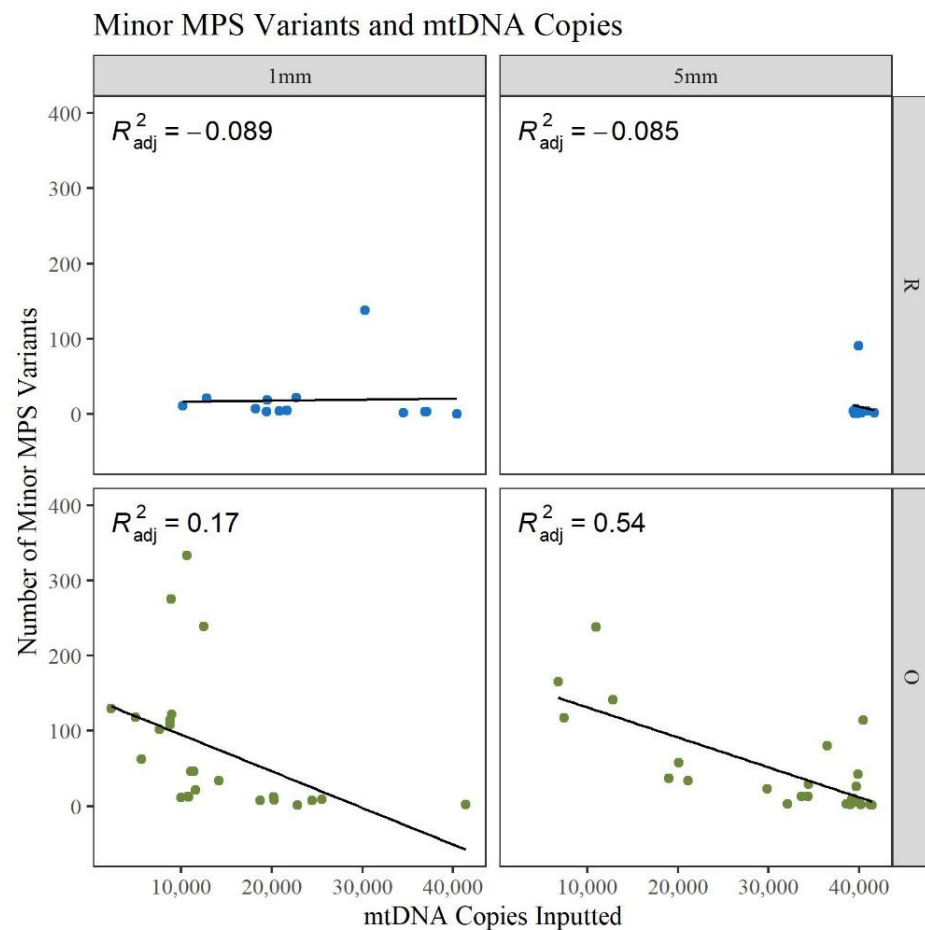
**Figure 4.** Boxplots for the comparison of the number of reported minor MPS variants for each age group and length ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ). The 'x' on the boxplots represents the mean and the line represents the median. Brackets show comparisons with significant  $p$ -values ( $p < 0.05$ ).





**Figure 5.** Linear regression of the number of reported minor MPS variants compared to the calculated degradation index for each age group and length ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ).

During amplification, the maximum volume of template DNA, 15  $\mu$ L, was often used to get as close to 40,000 copies of input mtDNA. Due to the lower yield from the O 1 mm hairs, fewer copies were added to the reaction. Figure 6 shows that for the 5 mm O hairs, a higher DNA input resulted in fewer minor variant calls. The 1 mm O hairs follow a similar trend but with more hairs clustered at lower input amounts and a wider spread in the number of minor variants. The 5 mm R hairs are clustered near 40,000 copies and close to zero minor variants while the 1 mm R hairs are at a relatively constant number of minor variants with a widespread template input amount. It has been previously shown that damage sites tend to appear more when lower template amounts are used because those minor variants have more opportunities to be amplified when poorer quality DNA is added to the reaction [31]. The data in Figure 6 and Table S3, Supplementary Materials, illustrates that template amount may impact the number of minor variants when working with older samples; however, it is not a strong predictor. In general, inputting more DNA template does not mean the quality of the sequence results will be of better quality. If the DNA is damaged or degraded, adding more to the amplification reaction will still result in noisy sequencing results. When analyzing mtDNA MPS data, it is important to assess and understand noise or error rates. This will help to distinguish heteroplasmy from noise, as heteroplasmic sites are useful for increasing the discrimination potential of mtDNA sequences and can be used to distinguish between maternal relatives [29,35,36]. With the high number of minor variants being observed in the O hairs, assessment of low-level heteroplasmic sites becomes more challenging.

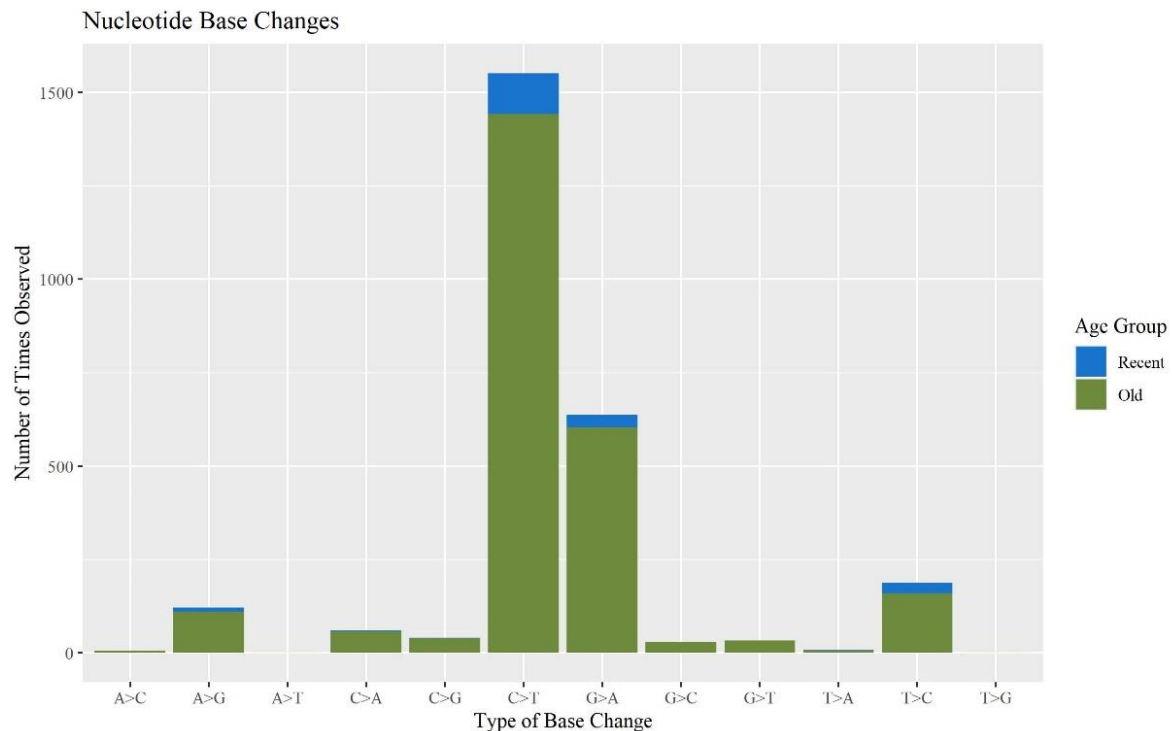


**Figure 6.** Linear regression of the number of reported minor MPS variants compared to the number of mtDNA copies inputted to the amplification reaction ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ).

As stated in Section 2.1, the hairs provided by Microtrace were stored in plastic bags at the time they were collected (seven to 46 years ago). The hairs were dry when they were collected, put in polyethylene bags, and were then kept in file cabinets at ambient temperature. When comparing the MPS data generated from the hairs provided by Mitotyping (stored frozen) and Microtrace, the 12 O hairs from Mitotyping had significantly fewer minor variants than the 12 O hairs from Microtrace ( $p$ -value of  $2.68 \times 10^{-3}$ ). On average, the O Mitotyping hairs had 33.125 minor variants, the O hairs from Microtrace had 91.833 minor variants, and the 23 VO hairs from Microtrace had 207 minor variants called in GM HTS. Based on this comparison, it appears that the long-term storage in plastic impacted the MPS data quality in combination with the increasing age of the hairs.

When assessing the types of base changes for possible DNA damage observed in the MPS data, apparent deamination damage was identified. Cytosine deamination has been observed in forensic samples in previous studies [31,34,37,38]. Deamination events typically appear as C to T transitions on one strand or G to A transitions on the other. These base changes accounted for 57.9% and 23.6% of the minor base changes, respectively. The majority of these sites occurred in the O hairs (Figure 7), which is consistent with previous studies showing that older samples exhibit more DNA damage [39]. Chemical treatments to hair can also contribute to DNA damage. Microtrace provided hair treatment information for the donated hairs. There was a significant difference between hairs that were treated and not treated with respect to the percent of the whole mitogenome reported ( $p$ -value of  $8.58 \times 10^{-3}$ ), but no other aspects of the sequence quality or mtDNA yield experienced a significant difference (Table S4, Supplementary Materials). We predict that treated hairs will yield more damaged DNA, possibly caused by oxidative damage from

bleaching and coloring and, therefore, partial mitogenome sequences. However, due to the low sample size of treated hairs ( $n = 8$ ) and the high number of hairs where treatment is unknown ( $n = 48$ ), more research is needed on the effects of hair treatment on mtDNA MPS data. An error assessment was performed on all 120 hair samples in the current study and is discussed below in Section 3.3.

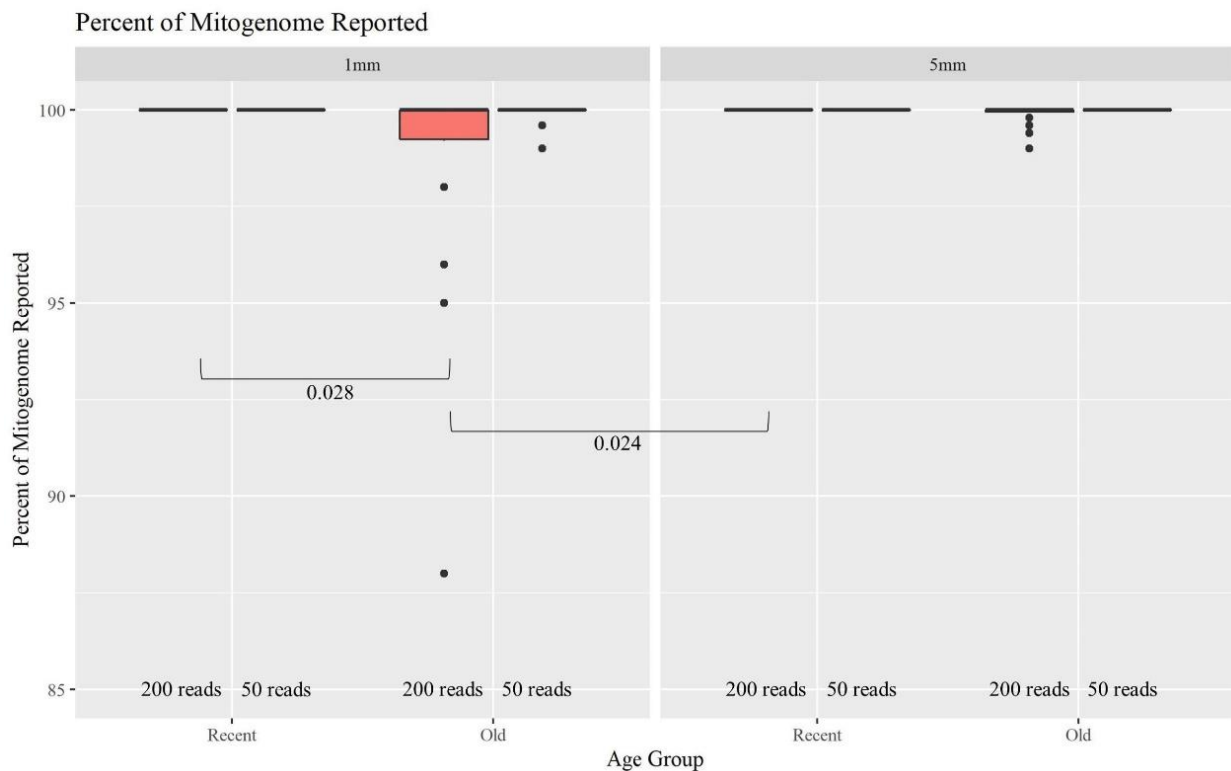


**Figure 7.** The number of times each base change is observed in the minor variants for each sample separated by age group ( $n = 74$ ).

In general, correct haplotypes were generated for all hair samples. While the actual haplotypes were unknown, as the donors were anonymous, no haplotype was observed more than once and when comparing the 1 and 5 mm cuttings from the same individual, the haplotypes matched over the comparable regions of sequence. Any differences were due to low coverage causing SNPs to drop out by not meeting the 200-read filter, being called minor SNPs, or not meeting the allele balance ratio filter (see Materials and Methods). Each haplotype was also run through Haplogrep2 to determine the haplogroup as extra quality control. The haplogroups matched for the 1 and 5 mm cuttings from the same individual except when key SNPs dropped out due to low coverage. The haplotypes and haplogroups are listed in Table S1, Supplementary Materials.

The minimum number of copies of mtDNA recommended to generate full mitogenome sequences is 20,000 copies. Therefore, hair extracts with at least 1333 copies of mtDNA/ $\mu\text{L}$  are expected to produce complete MPS profiles when considering a maximum input volume of 15  $\mu\text{L}$ . Many of the older and shorter hairs yielded fewer mtDNA copies (Figure 2). Furthermore, the O hairs were also slightly more degraded (Figure 3). This combined data suggests that those hairs would have a lower success rate for generating full mitogenome sequences. Figure 8 further supports this expectation, illustrating that as the age of hairs increases, the percentage of the mitogenome reported at a read depth of greater than 200 reads per nucleotide position decreases significantly. Nonetheless, 44 of the 48 O samples resulted in full or almost full mitogenome sequences (98–100%). In turn, when the required read depth was lowered to 50, all 48 O hairs reached the 98–100% threshold, resulting in no significant differences due to age or length. At a read depth of 200, for the 1 mm, a higher percentage of the mitogenome is reported for the R hairs compared to

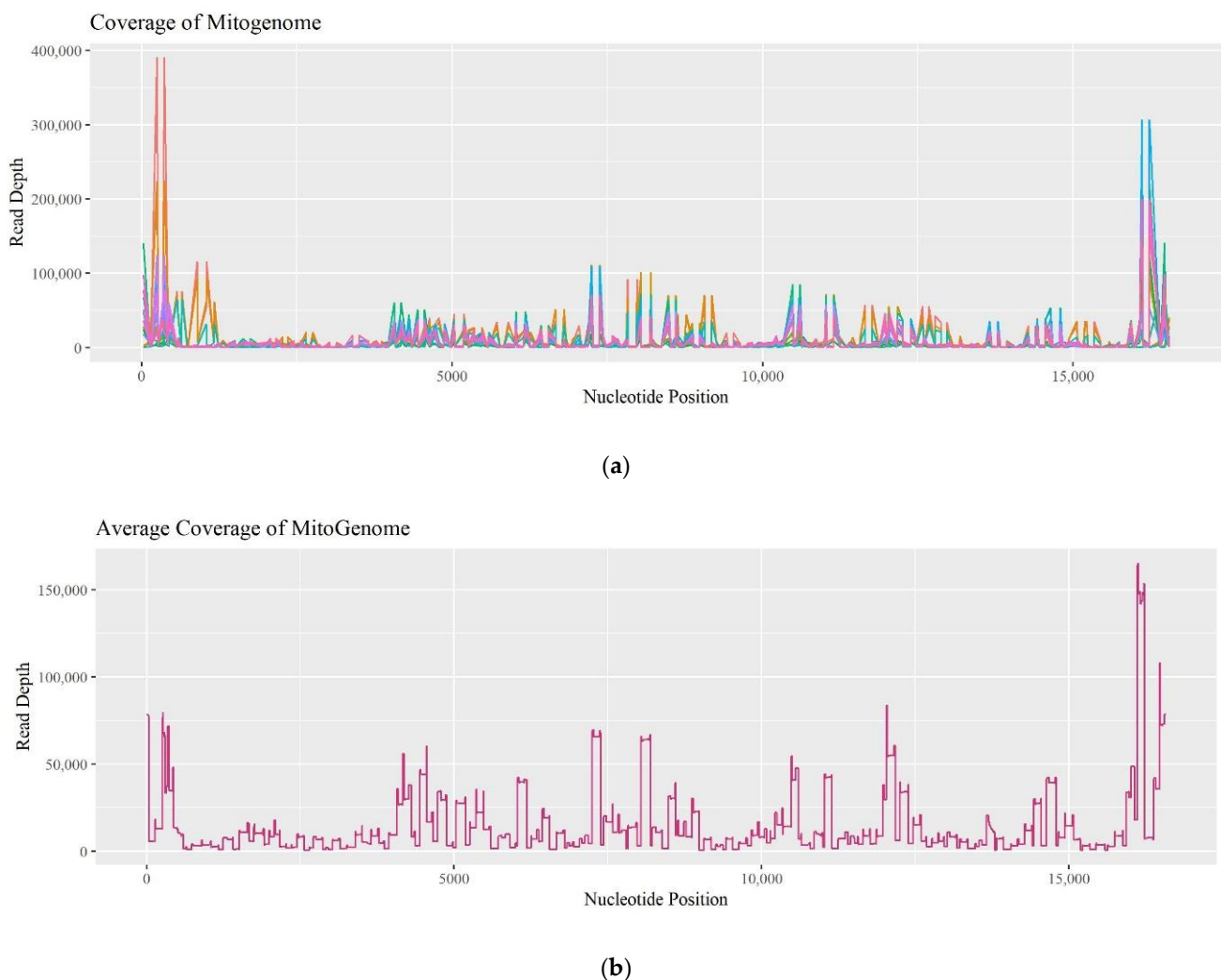
the O hairs ( $p$ -value of 0.028). Additionally, the 1 mm O hairs had a significantly lower percentage reported than the 5 mm R hairs ( $p$ -value of 0.024; see Table S2, Supplementary Materials for all  $p$ -values). Overall, hair length does not appear to significantly impact the percentage of the mitogenome reported as much as the age since hairs of the same age, but different lengths did not significantly differ in the percentage of the whole mitogenome reported. On average, >99.5% of the mitogenome was reported at a read depth of greater than 200 reads per nucleotide when working with as little as 1 mm of hair. That average increases to 99.9% for a read depth of greater than 50 reads; noting that the haplotype does not change, it simply becomes more complete due to drop-out occurring when setting the filter to 200 reads. Therefore, crime labs can use a less stringent filter when reporting a haplotype if the profile is not complete. While this data does follow the trend of older samples not having as high of a success rate, full whole mitogenome profiles were still produced from almost all of the O 1 mm hairs. This supports that the MPS method is robust and can generate mitogenome sequences on a routine basis from as little as 1 mm of hair shafts up to 27 years of age.



**Figure 8.** Boxplots for the comparison of the percent of the whole mitogenome reported for each age group and length at 200 reads and 50 reads ( $n = 74$ ;  $n_R = 26$ ,  $n_O = 48$ ). The line on the boxplots represents the median. Brackets show comparisons with significant  $p$ -values ( $p < 0.05$ ).

Since the commercially available PowerSeq<sup>TM</sup> Whole Mito System kit remains a prototype, the coverage of individual amplicons was evaluated to determine if there was a pattern of drop-out. Fifteen R and O samples experienced coverage dropping below 200 reads. Figure 9a plots read depths across the mitogenome for these 15 samples, while Figure 9b plots data for the entire set of 74 R and O samples tested. The data follows a similar trend in coverage and is comparable to previous assessments when sequencing the mitogenome [16]. Low-coverage zones are observed from nucleotide position (np) 1250–4000, np 7500–8000, np 9000–11,000, and np 15,400–16,000. Furthermore, the amplicon that spans np 2914–3034 dropped below 200 reads in 40% of the 15 samples, and amplicons spanning np 15,163–15,239 and np 8976–9073 dropped out in 47% of the 15 samples. This data suggests that certain regions are experiencing drop-out due to amplification and/or

sequencing challenges, and as a result, consistently lower coverage can be expected in those regions.

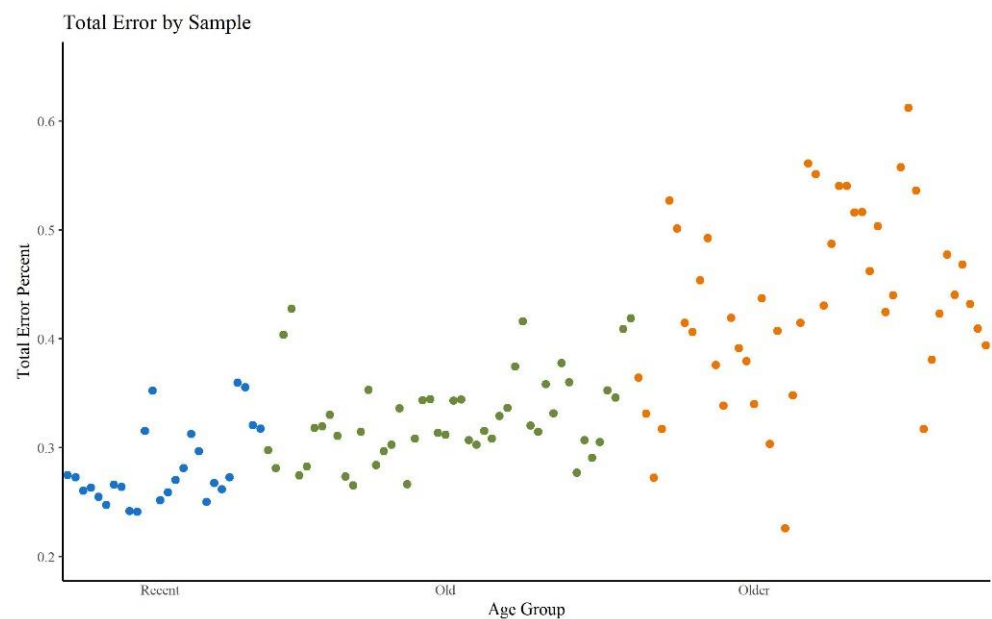


**Figure 9.** (a) A linear plot of the read depth across the mitogenome for the 15 samples that experienced amplicons dropping below 200 reads. (b) A linear plot of the average read depth for all 74 recent and old samples across the mitogenome.

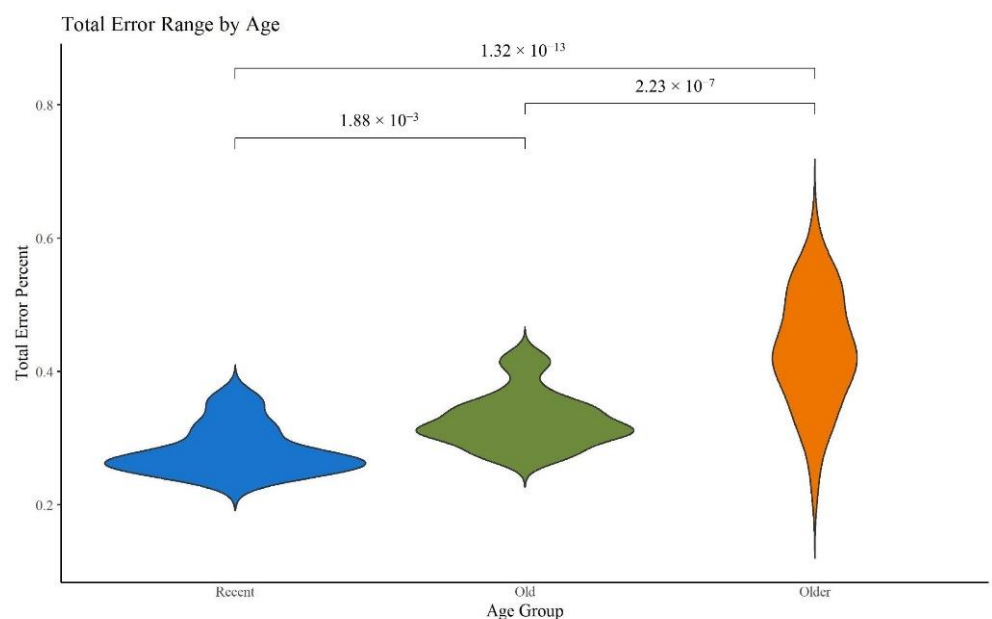
### 3.3. MPS Error Assessment

As stated in the Materials and Methods section, 23 hairs ranging from 41 to 46 years old were also subjected to mitogenome MPS analysis. While the coverage and read depth were sufficient to generate at least partial haplotypes for each hair, they resulted in noticeably noisier sequences. McElhoe and Holland [29] previously generated an R script to characterize the background noise in MiSeq MPS data. This method was used to calculate the total error rate per sample and the top 25 highest positions of error per nucleotide for all 120 hair samples tested. Figure 10a supports the expectation that higher error rates are seen in older samples. This increase in error rate is statistically significant among all three age categories (Figure 10b; Table S5, Supplementary Materials) and aligns with expectations given qualitative assessments of the sequencing data. While the average error rate for each age category is relatively low, falling well below the minor variant reporting threshold of 2%, a higher average may translate into a greater number of individual nps where the noise exceeds the threshold.





(a)

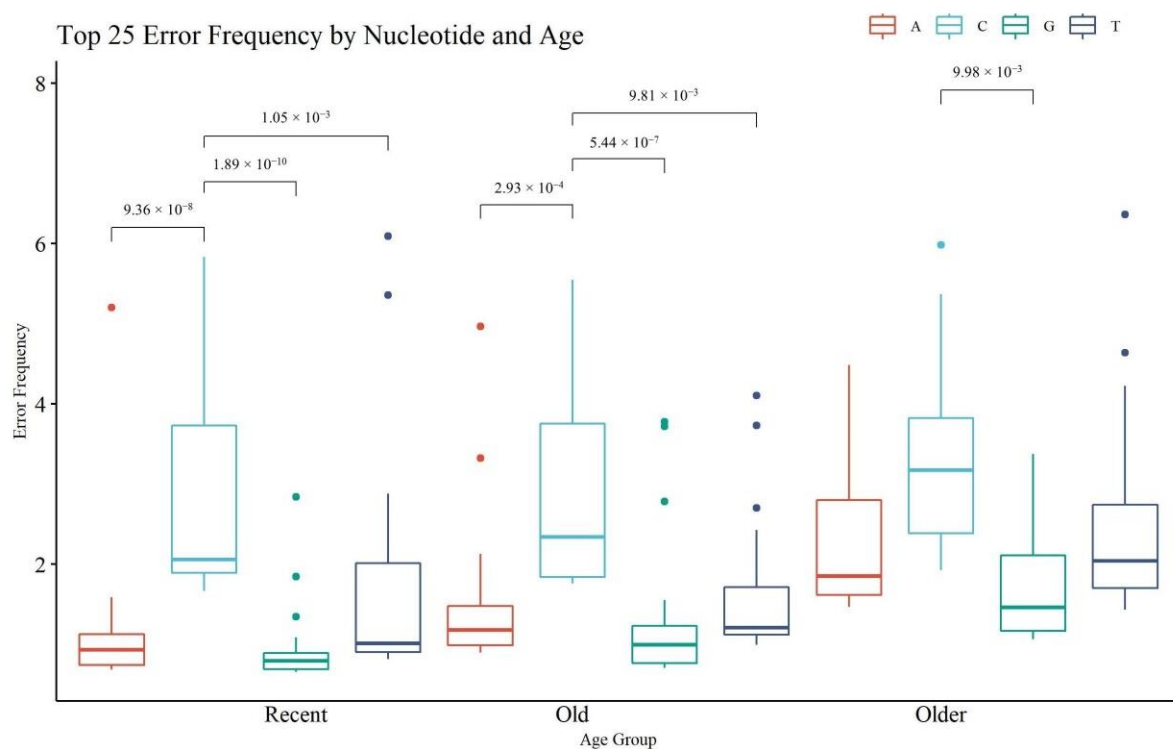


(b)

**Figure 10.** (a) Scatter plot of the total error for each sample. (b) Violin plot of the total error per age group. Brackets show comparisons with significant  $p$ -values ( $p < 0.05$ ).

Figure 11 provides an assessment of the top 25 nps with the highest rate of noise or error, on a per nucleotide basis (i.e., A, C, G, T) (see Table S6, Supplementary Materials for the top 25 positions and their frequencies). Statistical significance was calculated within each age category and across the same nucleotide (significant  $p$ -values for within each age category only are presented in Figure 11). The R and O hairs had significant differences in the frequency of noise between C and any other nucleotide; the VO hairs had significantly different error frequencies between C and G calls. A, G, and T calls were significantly higher in VO hairs compared to R hairs (see Table S7, Supplementary Materials for all

*p*-values). These trends align with the paradigm that older hairs contain more damaged DNA which in turn is associated with noisier MPS data. The rate of A, G, and T error generally fell below the 2% reporting threshold, with certain sites above the threshold in the older hairs that appear as reportable mixed sites. The rate of C substitutions is generally the highest in all age groups and was above 2% in all age categories. The T error was the next highest, followed by A and then G error. The elevated C error rate can be attributed to a combination of the filtering limitations of the MiSeq with respect to A/C detection and the alignment and sequencing challenges associated with homopolymeric C stretches in the control region [26,40]. The higher rate of T error compared to A and G can be explained by the observed deamination damage (C > T mutations), which has been previously noted as a predominate error type in hair samples [29].



**Figure 11.** Boxplots of noise for the top 25 nucleotide positions, per nucleotide type (A, C, G, T), and per age. Brackets show comparisons with significant *p*-values (*p* < 0.05) within each age category.

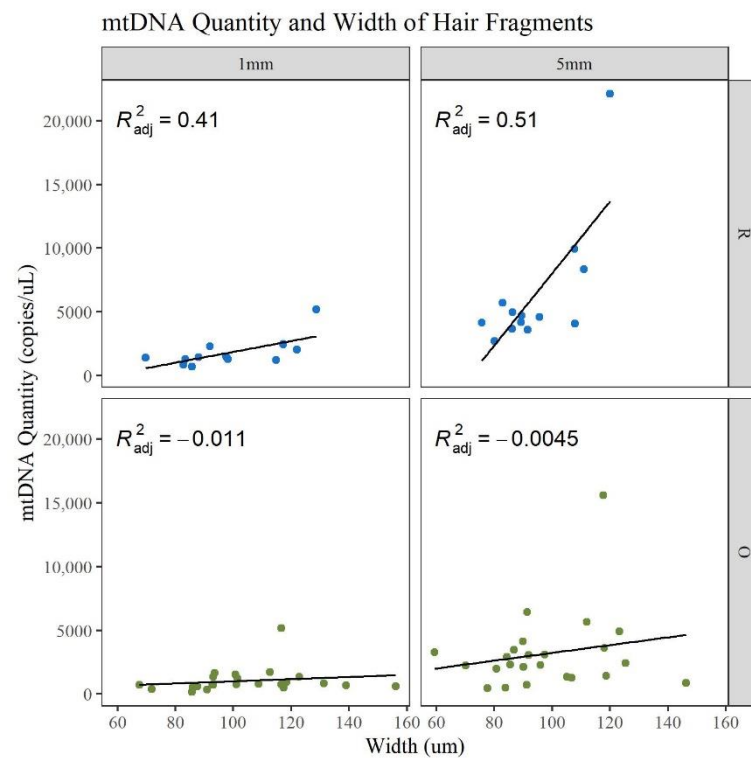
When sites of error exceed the minor variant reporting threshold, they complicate the interpretation of heteroplasmy. Substitutions caused by contamination, random error, and damage are typically random and will not be duplicated across replicates of the same sample [11,34,38]. Any duplicated sites when running replicates are more likely to be true heteroplasmy, excluding damage hotspots. Of the 300 nps represented in Figure 11 (25 nps per boxplot), 121 were observed once (40%) with 179 positions observed multiple times (60%). Of the 121 positions observed once, 30 sites (24.8%) were located in the control region of the mitogenome which is non-coding and contains the hypervariable regions [16]. The control region is 1122 nucleotides long and only represents 6.7% of the entire mitogenome (16,569 base pairs). The overrepresentation of error in the control region is expected given that mutations are more likely to occur in that region. Of the 300 error sites, 105 (35%) reached a frequency above the set 2% reporting threshold and 61 of those 105 (58%) were reported in at least one sample. A little more than half of the sites that were reported were in the VO hairs (55 sites; 52%). The remaining sites were not reported due to failing the other filter parameters including the read balance filter. The read balance filtering parameter is important as most hotspot error sites will not be reported due to imbalance [29].

When comparing the top error sites in the current study to the sites in hair samples from a previous error assessment [29], some of the same C error sites exceeded 2% (np 499, np 538, and np 545), but the other error type sites were not in common. Some of the sites for each nucleotide were duplicated across the different age groups, either being repeated in all three age groups or only in R and O or O and VO. Most of the duplicated sites were not reported, with only 31 sites exceeding all the filtering parameters. Therefore, as with other studies [29,31,38], increasing age and DNA damage will elevate the noise in MPS data, making it more difficult to interpret true heteroplasmy. Fortunately, recent hairs, which should be the majority of hairs received in crime laboratories, have low noise levels allowing for heteroplasmy to be detected and reported.

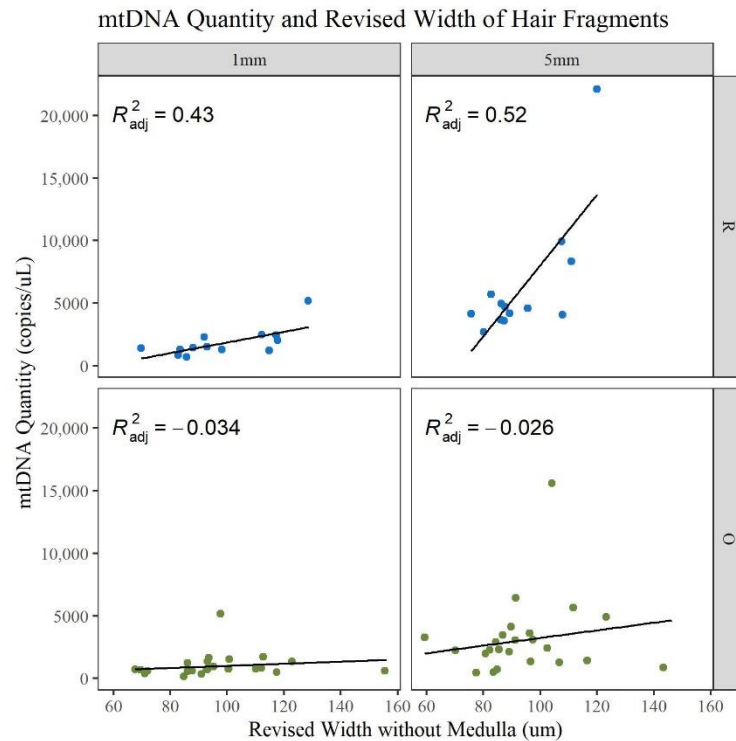
### 3.4. Microscopic Correlations

Microscopic characteristics such as hair width, medulla, and pigment were noted during examination of the hairs. It is expected that larger-diameter hairs will generally yield more DNA and more of a complete MPS profile [1,11]. In this data set there is an increase in mtDNA quantity with an increase in width, but it is not strongly correlated (Figure 12a; Table S3, Supplementary Materials). The linear regression model shows that there is a slight correlation for the longer and younger hairs and a poor correlation for the older hairs. Since the medulla is generally air space, if a medulla was present, its width was also subtracted from the width to determine if this changed the correlation to mtDNA yield. Certain data points shifted to the left (decrease in width), resulting in an increase in the  $R^2$  values for the 1 and 5 mm R hairs. However, the other  $R^2$  values decrease and the ability of the width to predict the mtDNA quantity remains the same (Figure 12b; Table S3, Supplementary Materials). To further assess this characteristic, the medulla was classified into four categories during microscopy: continuous, discontinuous, trace, and absent. The type of medulla did not significantly impact mtDNA yield, the DI, the number of minor MPS variants called, or the percent of the whole mitogenome reported (Table S8, Supplementary Materials). Therefore, while mtDNA yield increases with the hair's width slightly, the presence or absence of a medulla does not appear to impact the mtDNA recovery or quality of the MPS result. This is contrary to previous contentions that the presence of a medulla would reduce the DNA content in a hair shaft [41].

The pigment level observed in the tested hairs included no pigment (albino), blonde, light brown, medium brown, dark brown, red, and dark auburn (Table S1, Supplementary Materials). When considering only the hair's pigment, no trend was observed for the quantity and quality of recovered mtDNA or the quality of the MPS sequence data once the Holm correction was applied (Table S9, Supplementary Materials). The age and length of the hair had a more significant impact on the sequence results (See Sections 3.1 and 3.2). A previous study found that hairs with darker pigments are more likely to produce a full mtDNA profile [1], which are not consistent with the results from this study. One of the hair samples from this study had no pigment (albino) and produced a complete mitogenome. Overall, data from the current study do not suggest that microscopical characteristics significantly impact mtDNA MPS results. Nonetheless, microscopical examination of hairs is still a necessary step in the analysis process to document the physical characteristics of the hair, including whether the hair is of human origin, when applicable, and whether the hair has been damaged. These data can be used to exclude hairs for DNA analysis, provide information on the individual's ancestry, note any physical damage to the hair that could impact sequencing or provide details to a forensic investigation, and determine chemical treatment information such as bleaching or coloring.



(a)



(b)

**Figure 12.** Linear regression of the mtDNA quantity compared to the width of the hair. (a) Width including the medulla. (b) Width excluding the medulla.

#### 4. Conclusions

Based on the findings of the current study, mitogenome sequences can be generated on a routine basis from as little as 1 mm of human hair shaft up to 27 years of age since time of hair collection, illustrating that the applied MPS method is highly effective and robust. Furthermore, data from 1 and 5 mm replicates gave the same haplotype illustrating that the applied MPS method is reliable. As expected, as age increases, yield decreases, degradation increases, and sequencing quality declines. While the total error rate was low, older hairs with greater DNA damage or random error resulted in a higher rate of minor MPS variants, especially for the 1 mm hairs, jeopardizing the ability to assess heteroplasmy in those hairs.

The time is now for employing MPS-based methods for mtDNA analysis of DNA recovered from human hairs collected in forensic cases. If implemented, this approach will result in complete mitogenome profiles for samples historically deemed too challenging to analyze and will contribute to solving more forensic cases. In addition, generating mitogenome profiles from hair samples will assist the work of anthropologists and medical geneticists when studying ancestral origins, migration patterns, and health-related conditions.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13112144/s1>, Table S1: Sample information, extraction data, and sequencing results from all 120 hair samples. All VO hairs were excluded from the analyses for this publication due to significantly high sequencing error rates that require additional analysis. The VO hairs will be the subject of a future publication. The abbreviation ‘unk’ stands for ‘unknown’. Table S2: A listing of *p*-values for all possible comparisons of age and length for mtDNA quantity, degradation index, number of minor variants, and percent of the genome reported. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values. Table S3: Results of the linear regression models. When assessing how well the model fits the data, a residual standard error close to zero, an R-squared close to one, and a high F-statistic indicate a good model; Table S4: A listing of *p*-values for all possible comparisons of hair treatment for mtDNA quantity, degradation index, number of minor variants, and percent of the genome reported. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values. Table S5: A listing of *p*-values for all possible comparisons of the total error in the three age groups. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values. Table S6: The top 25 error positions for each nucleotide and age group. The sites that are colored in were observed multiple times within each nucleotide and across the age groups (green for A, blue for C, yellow for G, and red for T). Table S7: A listing of *p*-values for all possible comparisons of the top 25 error nucleotides for the three age categories. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values. Table S8: A listing of *p*-values for all possible comparisons of medulla type for mtDNA quantity, degradation index, number of minor variants, and percent of the genome reported. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values. Table S9: A listing of *p*-values for all possible comparisons of hair pigment for mtDNA quantity, degradation index, number of minor variants, and percent of the genome reported. Comparisons resulting in significant differences are in bold font. The \* denotes corrected *p*-values.

**Author Contributions:** Conceptualization, M.M.H.; Writing—original draft, L.C.C., J.A.M. and M.M.H.; Writing—review and editing, K.M.D., G.D. and J.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by grant 2019-DU-BX-0045 from the National Institute of Justice (NIJ). The points of view in this document are those of the authors and do not represent the official position or policies of the U.S. Department of Justice.

**Institutional Review Board Statement:** The study was approved by the Institutional Review Board of Penn State University under STUDY00000970 and STUDY00014305.

**Informed Consent Statement:** For samples collected at Penn State, informed consent was obtained.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to restrictions regarding consent.



**Acknowledgments:** The authors wish to thank Walther Parson (Institute of Legal Medicine, Medical University of Innsbruck) and Adam Garver (Ohio BCI) for their service as committee members on this project. Thanks to Shelby Bain (The Pennsylvania State University) for his assistance throughout the study. Lastly, thanks to the Forensic Science Program, Department of Biochemistry & Molecular Biology, Eberly College of Science at Penn State for the support and opportunity to perform this study, and to the Penn State Genomics Core Facility for running the samples on the MiSeq.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Melton, T.; Dimick, G.; Higgins, B.; Lindstrom, L.; Nelson, K. Forensic mitochondrial DNA analysis of 691 casework hairs. *J. Forensic Sci.* **2005**, *50*, 73–80. [CrossRef] [PubMed]
- Nilsson, M.; Norlin, S.; Allen, M. Sequencing of mtDNA in shed hairs: A retrospective analysis of material from forensic cases and a pre-screening method. *Open Forensic Sci. J.* **2012**, *5*, 13–22. [CrossRef]
- Cuenca, D.; Battaglia, J.; Halsing, M.; Sheehan, S. Mitochondrial sequencing of missing persons DNA casework by implementing Thermo Fisher’s precision ID mtDNA whole genome assay. *Genes* **2020**, *11*, 1303. [CrossRef] [PubMed]
- Kao, S.H.; Liu, C.S.; Wang, S.Y.; Wei, Y.H. Ageing-associated large-scale deletions of mitochondrial DNA in human hair follicles. *Biochem. Mol. Biol. Int.* **1997**, *42*, 285–298. [CrossRef] [PubMed]
- Schembri, K.; Scerri, C.; Ayers, D.; Pfannkuche, K.; Wideltz, R.B. Plucked Human Hair Shafts and Biomolecular Medical Research. *Sci. World J.* **2013**, *2013*, 620531. [CrossRef] [PubMed]
- Zheng, Y.; Luo, X.; Zhu, J.; Zhang, X.; Zhu, Y.; Cheng, H.; Xia, Z.; Su, N.; Zhang, N.; Zhou, J. Mitochondrial DNA 4977 bp deletion is a common phenomenon in hair and increases with age. *Bosn. J. Basic Med. Sci.* **2012**, *12*, 187–192. [CrossRef] [PubMed]
- Ricaute, F.X.; Bellatti, M.; Lahr, M.M. Ancient mitochondrial DNA from Malaysian hair samples: Some indications of Southeast Asian population movements. *Am. J. Hum. Biol.* **2006**, *18*, 654–667. [CrossRef]
- Gilbert, M.T.P.; Djurhuus, D.; Melchior, L.; Lynnerup, N.; Worobey, M.; Wilson, A.S.; Andreasen, C.; Dissing, J. mtDNA from hair and nail clarifies the genetic relationship of the 15th century Qilakitsoq Inuit mummies. *Am. J. Phys. Anthropol.* **2007**, *133*, 847–853. [CrossRef]
- Holland, M.M.; Melton, T.; Holland, C. Forensic mitochondrial DNA analysis: Current practice and future potential. In *Forensic DNA Analysis Current Practices Emerging Technologies*; CRC Press: Boca Raton, FL, USA, 2013; pp. 249–278.
- Brandhagen, M.D.; Loreille, O.; Irwin, J.A. Fragmented nuclear DNA is the predominant genetic material in human hair shafts. *Genes* **2018**, *9*, 8–12. [CrossRef]
- Gallimore, J.M.; McElhoe, J.A.; Holland, M.M. Assessing heteroplasmic variant drift in the mtDNA control region of human hairs using an MPS approach. *Forensic Sci. Int. Genet.* **2018**, *32*, 7–17. [CrossRef]
- Shih, S.Y.; Bose, N.; Goncalves, A.B.R.; Erlich, H.A.; Calloway, C.D. Applications of probe capture enrichment next generation sequencing for whole mitochondrial genome and 426 nuclear SNPs for forensically challenging samples. *Genes* **2018**, *9*, 49. [CrossRef] [PubMed]
- Grisedale, K.S.; Murphy, G.M.; Brown, H.; Wilson, M.R.; Sinha, S.K. Successful nuclear DNA profiling of rootless hair shafts: A novel approach. *Int. J. Legal Med.* **2018**, *132*, 107–115. [CrossRef] [PubMed]
- Gorden, E.M.; Sturk-Andreaggi, K.; Marshall, C. Capture Enrichment and Massively Parallel Sequencing for Human Identification. *Forensic Sci. Int. Genet.* **2021**, *53*, 102496. [CrossRef] [PubMed]
- Forsythe, B.; Melia, L.; Harbison, S. Methods for the analysis of mitochondrial DNA. *WIREs Forensic Sci.* **2021**, *3*, e1388. [CrossRef]
- King, J.L.; Larue, B.L.; Novroski, N.M.; Stoljarova, M.; Seo, S.B.; Zeng, X.; Warshauer, D.H.; Davis, C.P.; Parson, W.; Sajantila, A.; et al. High-quality and high-throughput massively parallel sequencing of the human mitochondrial genome using the Illumina MiSeq. *Forensic Sci. Int. Genet.* **2014**, *12*, 128–135. [CrossRef]
- Canale, L.C.; Parson, W.; Holland, M.M. The time is now for ubiquitous forensic mtMPS analysis. *WIREs Forensic Sci.* **2022**, *4*, e1431. [CrossRef]
- Koch, S.L.; Liebowitz, C.; Shriver, M.D.; Jablonski, N.G. Microscopical discrimination of human head hairs sharing a mitochondrial haplogroup. *J. Forensic Sci.* **2021**, *66*, 56–71. [CrossRef]
- Wilkinson, L.; Gwinnett, C. An international survey into the analysis and interpretation of microscopic hair evidence by forensic hair examiners. *Forensic Sci. Int.* **2020**, *308*, 110158. [CrossRef]
- Deedrick, D.W. FBI—Hairs, Fibers, Crime, and Evidence, Part 1. Available online: <https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/july2000/deedric1.htm> (accessed on 8 June 2021).
- Houck, M.M.; Budowle, B. Correlation of Microscopic and Mitochondrial DNA Hair Comparisons. *J. Forensic Sci.* **2002**, *47*, 964–967. [CrossRef]
- Parson, W.; Huber, G.; Moreno, L.; Madel, M.B.; Brandhagen, M.D.; Nagl, S.; Xavier, C.; Eduardoff, M.; Callaghan, T.C.; Irwin, J.A. Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples. *Forensic Sci. Int. Genet.* **2015**, *15*, 8–15. [CrossRef]
- Holland, M.M.; Pack, E.D.; McElhoe, J.A. Evaluation of GeneMarker® HTS for improved alignment of mtDNA MPS data, haplotype determination, and heteroplasmy assessment. *Forensic Sci. Int. Genet.* **2017**, *28*, 90–98. [CrossRef] [PubMed]

24. Turnbull, D.M.; Andrews, R.M.; Kubacka, I.; Chinnery, P.F.; Lightowlers, R.N.; Howell, N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA—Nature Genetics. *Nat. Genet.* **1999**, *23*, 147.
25. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [[CrossRef](#)]
26. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 585–621. [[CrossRef](#)]
27. Dunn, O.J. Multiple comparisons using rank sums. *Technometrics* **1964**, *6*, 241–252. [[CrossRef](#)]
28. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **1979**, *6*, 65–70.
29. McElhoe, J.A.; Holland, M.M. Characterization of background noise in MiSeq MPS data when sequencing human mitochondrial DNA from various sample sources and library preparation methods. *Mitochondrion* **2020**, *52*, 40–55. [[CrossRef](#)]
30. Chambers, J.M.; Hastie, T.J. Linear Models. In *Statistical Models in S*, 1st ed.; Chambers, J.M., Hastie, T.J., Eds.; Wadsworth: Belmont, CA, USA; Brooks/Cole: Pacific Grove, CA, USA, 1992; pp. 61–87.
31. Holland, C.A.; McElhoe, J.A.; Gaston-Sanchez, S.; Holland, M.M. Damage patterns observed in mtDNA control region MPS data for a range of template concentrations and when using different amplification approaches. *Int. J. Legal Med.* **2021**, *135*, 91–106. [[CrossRef](#)]
32. Desmyter, S.; Bodner, M.; Huber, G.; Dognaux, S.; Berger, C.; Noël, F.; Parson, W. Hairy matters: MtDNA quantity and sequence variation along and among human head hairs. *Forensic Sci. Int. Genet.* **2016**, *25*, 1–9. [[CrossRef](#)]
33. Andréasson, H.; Nilsson, M.; Budowle, B.; Lundberg, H.; Allen, M. Nuclear and mitochondrial DNA quantification of various forensic materials. *Forensic Sci. Int.* **2006**, *164*, 56–64. [[CrossRef](#)]
34. Rathbun, M.M.; McElhoe, J.A.; Parson, W.; Holland, M.M. Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data. *Forensic Sci. Int. Genet.* **2017**, *26*, 1–11. [[CrossRef](#)] [[PubMed](#)]
35. Holland, M.M.; Makova, K.D.; McElhoe, J.A. Deep-coverage MPS analysis of heteroplasmic variants within the mtgenome allows for frequent differentiation of maternal relatives. *Genes* **2018**, *9*, 3. [[CrossRef](#)] [[PubMed](#)]
36. Marshall, C.; Sturk-Andreaggi, K.; Gorden, E.M.; Daniels-Higginbotham, J.; Sanchez, S.G.; Bašić, Ž.; Kružić, I.; Anđelinović, Š.; Bosnar, A.; Čoklo, M.; et al. A forensic genomics approach for the identification of Sister Marija Crucifiksa Kozulić. *Genes* **2020**, *11*, 938. [[CrossRef](#)] [[PubMed](#)]
37. Gorden, E.M.; Sturk-Andreaggi, K.; Marshall, C. Repair of DNA damage caused by cytosine deamination in mitochondrial DNA of forensic case samples. *Forensic Sci. Int. Genet.* **2018**, *34*, 257–264. [[CrossRef](#)] [[PubMed](#)]
38. Holland, M.M.; Bonds, R.M.; Holland, C.A.; McElhoe, J.A. Recovery of mtDNA from unfired metallic ammunition components with an assessment of sequence profile quality and DNA damage through MPS analysis. *Forensic Sci. Int. Genet.* **2019**, *39*, 86–96. [[CrossRef](#)] [[PubMed](#)]
39. Gilbert, M.T.P.; Hansen, A.J.; Willerslev, E.; Rudbeck, L.; Barnes, I.; Lynnerup, N.; Cooper, A. Characterization of genetic miscoding lesions caused by postmortem damage. *Am. J. Hum. Genet.* **2003**, *72*, 48–61. [[CrossRef](#)]
40. Schirmer, M.; Ijaz, U.Z.; D’Amore, R.; Hall, N.; Sloan, W.T.; Quince, C. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* **2015**, *43*, e37. [[CrossRef](#)]
41. Linch, C.A.; Whiting, D.A.; Holland, M.M. Human hair histogenesis for the mitochondrial DNA forensic scientist. *J. Forensic Sci.* **2001**, *46*, 844–853. [[CrossRef](#)]