Impact of DNA degradation on massively parallel sequencing-based autosomal STR, iiSNP, and mitochondrial DNA typing systems

Zavala, Elena I.; Rajagopal, Swetha; Perry, George H.; Kruzic, Ivana; Bašić, Željana; Parsons, Thomas J.; Holland, Mitchell M.

Source / Izvornik: **International Journal of Legal Medicine, 2019, 133, 1369 - 1380**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.1007/s00414-019-02110-4

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:227:769452

International Journal of Legal Medicine (2019) 133:1369–1380
https://doi.org/10.1007/s00414-019-02110-4

ORIGINAL ARTICLE

Impact of DNA degradation on massively parallel sequencing-based autosomal STR, iiSNP, and mitochondrial DNA typing systems

Elena I. Zavala[1] · Swetha Rajagopal[1,2] · George H. Perry[3] · Ivana Kruzic[4] · Željana Bašić[4] · Thomas J. Parsons[5] · Mitchell M. Holland[1]

Received: 1 February 2019 / Accepted: 19 June 2019 / Published online: 2 July 2019

Abstract

Biological samples, including skeletal remains exposed to environmental insults for extended periods of time, exhibit increasing levels of DNA damage and fragmentation. Human forensic identification methods typically use a combination of mitochondrial (mt) DNA sequencing and short tandem repeat (STR) analysis, which target segments of DNA ranging from 80 to 500 base pairs (bps). Larger templates are often unavailable as skeletal samples age and the associated DNA degrades. Single-nucleotide polymorphism (SNP) loci target shorter templates and may serve as a solution to the problem. Recently developed assays for STR and SNP analysis using a massively parallel sequencing approach, such as the ForenSeq kit (Verogen, San Diego, CA), offer a means for generating results from degraded samples as they target templates down to 60 to 170 bps. We performed a modeling study that demonstrates that SNPs can increase the significance of an identification when analyzing DNA down to an average size of 100 bps for input amounts between 0.375 and 1 ng of nuclear DNA. Observations from this study were then compared with human skeletal material results ($n = 14$, ninth to eighteenth centuries), which further demonstrated the utility of the ForenSeq kit for degraded samples. The robustness of the Promega PowerSeq™ Mito System was also tested with human skeletal remains ($n = 70$, ninth to eighteenth centuries), resulting in successful coverage of 99.29% of the mtDNA control region at 50× coverage or more. This was accompanied by modifications to a mainstream DNA extraction technique for skeletal remains that improved recovery of shorter templates.

Keywords Fragmented DNA · Massively parallel sequencing · SNPs · STRs

Introduction

The identification of skeletal remains plays a key role in missing persons, mass disaster, and military and human rights cases [1–5]. Working with these types of samples is difficult as they are often comprised of minute quantities of DNA less than 150 bps in size, while traditional forensic typing methods target the recovery, amplification, and analysis of longer pieces of DNA (typically 80 to 500 bps in size) [6]. With more degraded samples, the possibility of successfully amplifying the number of required targets for individualization decreases [1], particularly in cases of identification involving kinship matching. Mitochondrial (mt) DNA is present in higher amounts in the cell than nuclear (n) DNA (100–1000 copies compared with 2 copies per cell for nDNA, respectively) and therefore is more likely to persist over long periods of time. For this reason, mtDNA analysis is often used to support the identification process, which can associate remains to maternal lineages, and in cases involving the presence of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00414-019-02110-4) contains supplementary material, which is available to authorized users.

✉ Elena I. Zavala
elena_zavala@eva.mpg.de

[1] Department of Biochemistry and Molecular Biology, Forensic Science Program, Pennsylvania State University, State College, PA, USA

[2] Department of Forensic Science, John Jay College of Criminal Justice, New York, NY, USA

[3] Departments of Anthropology and Biology, Pennsylvania State University, State College, PA, USA

[4] University Department of Forensic Sciences, University of Split, Split, Croatia

[5] International Commission on Missing Persons, The Hague, Netherlands

heteroplasmy, can provide substantial weight for identification [7]. Nonetheless, mtDNA analysis is not as discriminating as the analysis of short tandem repeat (STR) and single-nucleotide polymorphism (SNP) loci. The challenges when working on human identification of older human skeletal material, however, goes beyond working with fragmented DNA, as analysts must also consider how to confirm endogenous results, minimize contamination concerns [8], and deconvolute sample results containing damage, contamination, or mixed profiles due, in part, to comingling of remains in mass graves [9].

Despite these challenges, there have been successful projects involving aged skeletal remains using STR analysis. An example is the successful identification of over 17,909 remains from conflicts during the break-up of the former Yugoslavia [4, 5]. With these successes, however, there are still samples where identifications have not been possible, as exemplified with a study on 150 Norwegian skeletal remains from World War II where 38% of samples failed to yield suitable DNA profiles [10]. The number of missing person cases continues to grow with current conflicts and natural disasters adding to the thousands of unidentified skeletal remains from conflicts associated such as the Korean War [11] and both World Wars [10, 12].

Advancements in human identification often come from both the fields of forensic science and ancient (a) DNA [13, 14]. Research in aDNA routinely leads to the development of cutting edge methods for the analysis of skeletal material, as illustrated through successful identification and genomic analysis of ancient hominins, including the discovery of a new hominin population group [15]. These discoveries are the result of methodological advancements relating to the successful recovery of ultra-short copies of endogenous DNA [16–19] library preparation for low copy number and single-stranded DNA [20], targeted capture of molecules of interest [21], massively parallel sequencing (MPS), and subsequent bioinformatics for genetic analysis [22]. These techniques, however, are optimized to answer biological questions aimed at identifying and comparing population groups, while forensic scientists look to identify individuals. Despite this difference, the common objective of recovering small amounts of highly degraded hominin DNA makes it clear that advancements and discoveries made in the field of aDNA have a high potential for positive impact in forensics.

The benefits of implementing MPS are clear, including the ability to target more loci of varying sizes at a single time to gain higher discriminatory information, both of which greatly increase the potential for the identification of biological samples [23–25]. While MPS is becoming the current standard in aDNA, applications in forensics are in the evaluation process for STR, SNP, and mtDNA analyses [7, 26–28]. The benefits when applied to mtDNA analysis have been demonstrated through advancements in the identification of human hairs [29], including the assessment of heteroplasmic drift [30], but have not been fully explored for aged human skeletal remains. The PowerSeq™ Mito System kit from Promega, which targets ten overlapping regions of the mtDNA control region (CR), was used for both of these studies, demonstrating the potential for its use with degraded samples. The ForenSeq kit from Verogen not only targets the 27 STR loci currently included in conventional STR typing kits but also decreases the size of the target amplicons and includes 94 informative identity SNPs and ancestral and phenotypically informative SNPs with even smaller target sizes [31–36]. SNPs not only have the benefit of having small target sizes but are easy to work with analytically and they tend not to be prone to artifacts. Due to these reasons, we anticipate that SNPs will be essential for increasing identification potential in forensic cases with highly degraded samples, including the identification of human skeletal remains.

During the implementation of MPS analysis of SNP loci in forensic case work, it will be important to understand the impact of DNA degradation on the interpretation of profile information. Stochastic sampling as a result of low copy number and/or degraded DNA is a known phenomenon, often resulting in differing parameters of analysis for these types of samples [37, 38]. Previous studies have generally been limited to scenarios where the nature (fragment size distribution, damage patterns, etc.) of the DNA was unknown [35, 36]. Here, we investigate the utility of both the ForenSeq system and the PowerSeq™ Mito kit to human skeletal remains for identification purposes. We first present a sensitivity test with degraded DNA of two different known size distributions and five different input amounts to evaluate both the sequencing results and analytical parameters of the ForenSeq system when working with highly degraded samples. In a second phase, we compared the observations from the sensitivity study with results from skeletal remains from the ninth to eighteenth centuries. Finally, the PowerSeq™ Mito kit was evaluated for mtDNA coverage and haplotype determination using seventy skeletal remains from the ninth to eighteenth centuries. This was accompanied by evaluation of modifications to a mainstream DNA extraction technique for skeletal remains with the goal of improving recovery of smaller DNA templates.

Material and methods

Sensitivity study

Biological samples from living donors were collected under the Penn State University internal review board (IRB) approved protocol STUDY00000970. Organic DNA extracts

from buccal swabs of three different donors were sheared to two different size distributions (average of 150 bp and 100 bp in size) and diluted to five different input amounts (1 ng, 250 pg, 125 pg, 25 pg, and 5 pg) for the ForenSeq system. In order to more closely mimic extracts from skeletal material, the experiment was repeated, this time mixing the sheared buccal extracts with demineralized animal bone lysate and purifying the lysate/buccal extract to generate a second extract. The ForenSeq results were evaluated for coverage (read depth), concordance, heterozygote balance, and dropout across input amounts for each size distribution of sheared DNA and type of extract (buccal versus buccal with demineralized animal bone). These results were also compared with the results of fourteen human skeletal extracts that were analyzed with the ForenSeq system.

Shearing sample preparation

The starting material with an average size of 150 bp (Fig. 1) was sheared from modern buccal DNA extracts using a Covaris S220; 130 uL sample (approximately 1 μg of total nDNA), 175 peak incident power (w), 10% duty factor, 200 cycles per burst, 430 s treatment time. The starting material with an average size of 100 bp (Fig. 1) was sheared from modern buccal DNA extracts using a Covaris M220; 130 uL sample (approximately 1 μg of total nDNA), 50 peak incident power (w), 20% duty factor, 200 cycles per burst, 2400 s treatment time. Size distributions of samples were confirmed and samples quantified using a Bioanalyzer 2100 with a high sensitivity chip (Agilent).

Sample quantification

Extracted samples were quantified for nDNA using the Quantifiler™ HP kit (cat. no. 4482911) from Thermo Fisher Scientific, which includes an 80-bp and 214-bp autosomal target. The short target was used for quantification, while the larger target was used to evaluate degradation (amount of smaller target compared with amount of larger target within a sample). Sheared samples were quantified using results from the Bioanalyzer 2100 with a high sensitivity chip and then diluted to the target input amounts for library preparation (1 ng, 375 pg, 125 pg, 25 pg, and 5 pg). Quantification of mtDNA copy number was determined using a custom quantification method [30], which includes a 69-bp and 283-bp target. The two targets allow for quantification of mtDNA copy number and an assessment of degradation, respectively.

Sensitivity study analysis

The results of the sensitivity study were evaluated for MPS coverage (read depth), concordance, heterozygote balance (ratio between the number of reads for each allele at a heterozygous loci), and dropout across input amounts for different sizes of sheared DNA (average of 150 bp and 100 bp). Variation of results between amplifications, donors, sequencing runs, and addition of demineralized animal bone to the sheared starting material was also evaluated. Dropout and concordance were evaluated by comparing the resulting genotypes with the known sequence and identifying full dropout (no observed alleles), false homozygote (result of stochastic sampling at a heterozygous locus), and consensus sequence. Analysis was completed with the recommended Universal Analysis Software (UAS) thresholds from Illumina [35]. A secondary analysis was completed by setting all thresholds to zero to evaluate if altered analytical parameters should be used when working with degraded samples. Results were compared with studies completed on non-sheared samples [35, 36].

The potential evidentiary weight of resulting genotypes was determined through the generation of random match probabilities [39] using the FBI CODIS allele frequencies [40] and the Forensic Resource Reference On Genetics (FROG) knowledge base compared with a P_x value (Eq. 1). Maximum random match probabilities were calculated using the alleles with the highest observed frequency at each locus for a given population group. The P_x value provides an identification threshold that specifies the most conservative random match probability required to identify a match as approaching source attribution for a given number of people (N) in a population to a specific degree of confidence (α). This was calculated using Eq. 1, with a 0.99 degree of confidence for a world population of 7 billion people, which provided an identification threshold of 1 in 733 billion, as a conceptual indication that the individual might be considered uniquely identified in the world population. In times when there is not a reference available for the person that is missing, kinship analysis is often used with references provided by parents or siblings. Statistics related to kinship analysis were not conducted for this study, but should be taken into account in future studies.

$$P_x = 1-(1-\alpha)^{1/N} \quad (1)$$

Human skeletal samples and controls

The human skeletal material used for this study originated from long bones and teeth from remains at three different sites in Croatia ranging from the ninth to eighteenth centuries (Velim-Velistak, Otok Vuletina Rupa-Grebcine, and Koljani). All skeletal material was selected and prepared at the University of Split and ground or reground with a Waring blender at the Pennsylvania State University. A new extraction protocol was identified (Supplemental Part 1 or SP1) and applied using 0.50 to 0.59 g of bone powder incubated overnight with rotation at 56 °C in 7.5-mL digestion buffer (0.5 M EDTA pH 8.0, 1% n-lauroylsarcosine, 10 mg

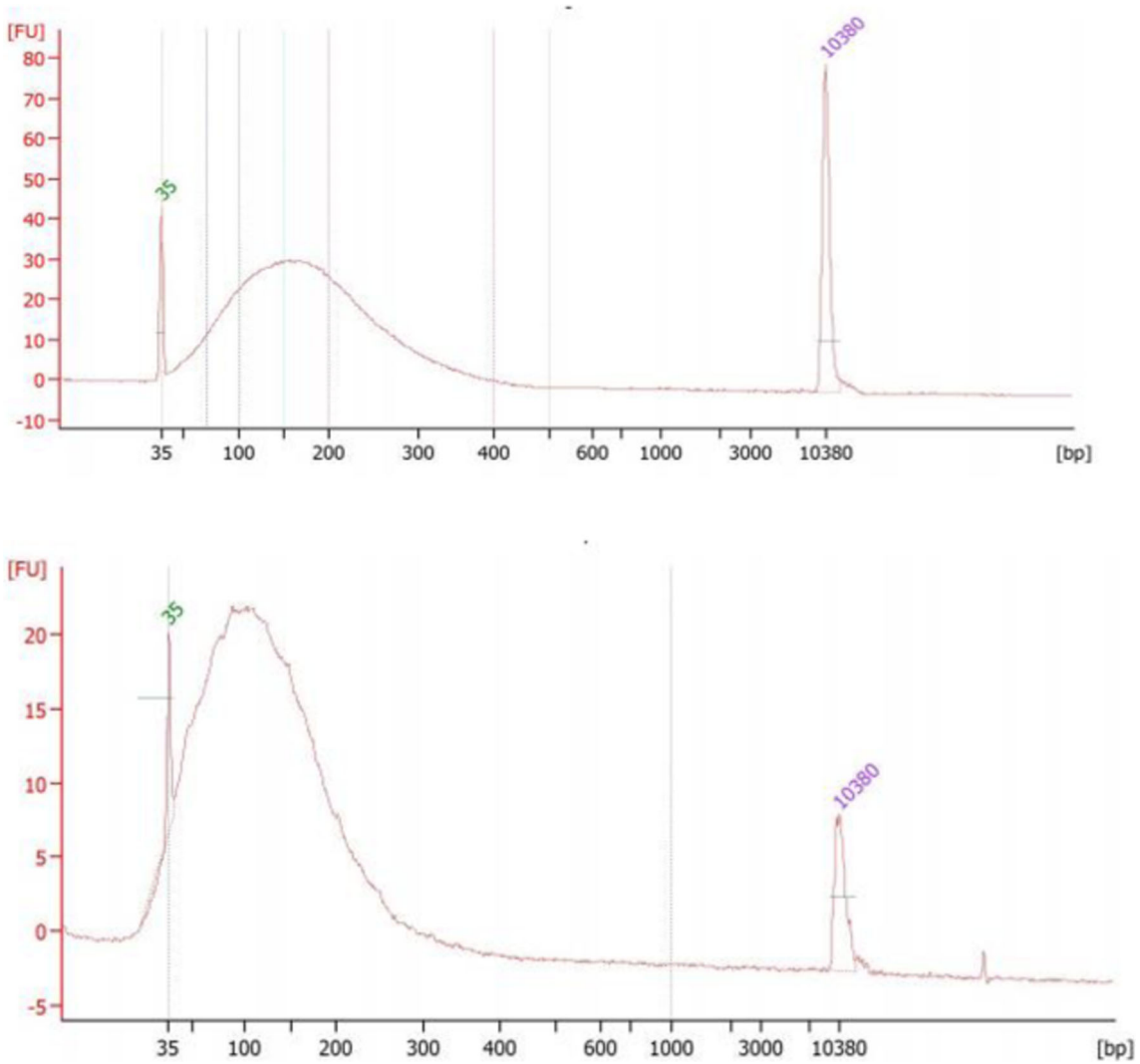


Fig. 1 Fragment length distribution of pristine DNA that was sheared using a Covaris S220 with an average fragment length of 150 bp (upper) and 100 bp (lower). Results are from a Bioanalyzer 2100 with a high sensitivity chip (Agilent) and indicate the separation of different sized DNA fragments in relation to size (bp). The highest point of the main peak is at approximately 150 bp (upper) and 100 bp (lower). The 35- and 10,380-bp peaks are internal sizing standards

proteinase K). The digestion buffer was preheated to 56 °C before addition to bone powder. Post-incubation, each tube was spun for five minutes at 2500 rpm and supernatant was removed and concentrated to approximately 300 μL using a 30-kDa Amicon Ultra-4 filter. Concentrate was then mixed with 5× Qiagen PB buffer and spun through a Qiagen MinElute column at 10.6 krpm (12.9 k × g) in an Eppendorf 5415D centrifuged for two minutes in three aliquots. Flow through was discarded after each spin. The column was then washed three times with 750 μL of Qiagen PE buffer with two minute spins at 10.6 krpm. A dry spin was completed at max speed (13.2 krpm or 16.1 k × g) for one min. The DNA was then eluted with 40 μL of Qiagen EB buffer after at least one minute of incubation at room temperature and spinning at 10.6 krpm (12.9 k × g) for one and a half minutes. Extracts were stored at − 20 °C.

All contamination controls and modern samples used for the sensitivity study were taken from buccal swabs in accordance with Penn State University internal review board (IRB) approved protocol Study 00000970. A cutting from each swab was placed in a 1.7-mL tube with 400-μL stain extraction buffer (2% SDS; 10 mM EDTA; 100 mM NaCl; 7.6 mM Tris-HCl, pH 8.0) and 10 μL Proteinase K (20 mg/mL) and incubated for at least 1 h at 56 °C. The cotton cutting was then removed, 400 μL PCIA was added to the tube and the layers were mixed through inversion. The tube was then centrifuged at 7000 rpm (8.5 k × g) for 10 min. The upper aqueous layer was removed and placed in a new tube where it was mixed with 40 μL of 3 M sodium acetate, pH 5.3, and then 1.0 mL ice-cold 100% ethanol. The tube was then incubated at 4 °C for at least 12 h. After incubation, samples were spun for 10 min at 11 krpm (13.4 k × g) and the supernatant was

discarded. The pellet was washed with 1.0 mL of 70% ethanol and spun for 5 min at 11 krpm. Supernatant was discarded and the pellet was left to dry at room temperature in a hood overnight. The DNA pellet was resuspended in 40 μL of low TE buffer (1 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) and incubated at 56 °C for 30 min to reconstitute the DNA extract.

DNA library preparation, sequencing, and analysis

The protocol for the PowerSeq™ Mito System kit was followed for sequencing the mtDNA CR of the seventy human skeletal samples. For samples that quantified below the recommended input amount, the maximum input volume of 10 μL was used. Analysis was performed using GeneMarker® HTS software by SoftGenetics, LLC version 1.0.1.1191 [41]. Haplotypes were determined and coverage evaluated using a minimum coverage of 50×. Haplogroups were determined using the EMPOP mtDNA database, v3/R11. The ForenSeq kit with Primer Set B from Illumina, which targets 27 STRs, 94 identity SNPs, and phenotypic and ancestry SNPs, was used for the degradation sensitivity assessment. For human skeletal samples that were quantified below the recommended input value (0.2 ng/uL), the maximum input volume of 5 μL was used. Results were analyzed using the ForenSeq UAS. The UAS software provides an overview of the sequencing run quality metrics as well as allele and genotype calling for the targeted loci in the ForenSeq kit based on preset analytical and interpretation thresholds [35]. The user is able to review the results through histograms and pie charts of each locus that represent the read coverage for each observed allele at that locus. For routine analysis, the automatic threshold settings were used (1.5% for analytical and 4.5% for interpretation with a minimum of 650 reads, Table 1), while for the threshold evaluation, the individual thresholds were set to zero.

Contamination containment

When working with aDNA from human skeletal remains, contamination containment is an important consideration, requiring confirmation of endogenous results [8]. Buccal swabs were collected and typed from all members of both the Pennsylvania State University and University of Split laboratories who were known to be in areas where the materials were handled, and the experimental DNA results compared with their STR and mtDNA profiles. Before any laboratory work was performed, all surfaces and tools were cleaned with 10% bleach and 70% ethanol and subjected to a UV crosslinker for at least 45 min. The outer superficial layer of all bones was removed with a grinding stone, and the bone further cleaned with bleach, ethanol, and water in order to remove any surface contamination before pulverizing the bone sample. Reagent blanks and negative controls were carried through each experiment, and all extractions were performed in a room maintained for low copy number samples with pre- and post-PCR rooms separated (see Supplemental Part 3, or SP3, for reagent blank and negative control results).

Table 1 Overview of the standard analysis settings used with the Illumina UAS software

Threshold	Setting	Measurement	Purpose
Analytical threshold	1.5%	Read coverage of allele divided by total read coverage at locus (minimum locus read coverage of 650 reads)	Alleles with coverage below this amount are not analyzed
Interpretation threshold	4.5%	Read coverage of allele divided by total read coverage at locus (minimum locus read coverage of 650 reads)	For a locus to be identified as homozygous, the observed allele must have coverage above this threshold
Intralocus balance (heterozygous balance)	50%	Read coverage of minimum intensity typed allele divided by read coverage of maximum intensity typed allele	Identify possible allelic imbalance at a locus and correctly call heterozygous loci
Stutter	Varies per locus (0–25%)	Read coverage of allele in question divided by read coverage of parent allele	Differentiate observed stutter from true alleles

Results

Sensitivity study

A highly significant identification RMP was achieved when using either the STRs or identity SNPs from the ForenSeq kit for samples down to 125 pgs of input DNA with an average size of 150 bps (Table 2). While identification potential was lost at 375 pgs of input DNA with an average size of 100 bps for STRs and SNPs in certain population groups, Hispanic American/Mexican and European American/Caucasian population groups maintained a significant minimum RMP down to 375 pgs. This was determined using maximum RMPs for Caucasian, African American, and Southwest and Southeast Hispanic population groups for the average observed consensus loci for each input amount and size of DNA for STRs and SNPs from Fig. 2 and comparing them with a P_x value of 1 in 733 billion. There was no significant difference observed between the performance of starting material associated with sheared DNA and samples of sheared DNA mixed with

Table 2 Corresponding maximum RMPs for observed consensus loci (Fig. 2) for highest three input amounts. In parenthesis is the population group that had the highest maximum RMP for each number and target loci. *SWH*, Southwest Hispanic; *EA/C*, European American/Caucasian; *HA/M*, Hispanic American/Mexican; *AA*, African American; *C*, Caucasian

Input amount	Target	Sheared DNA 150 bp ($n = 16$)		Sheared DNA 100 bp ($n = 18$)	
		Observed consensus loci	Maximum RMP	Observed consensus loci	Maximum RMP
1 ng	STRs	25	< 1 in 5.5 × 10^21 (SWH)	14	1 in 7.6 × 10^11 (SWH)
	SNPs	90	1 in 6.9 × 10^36 (EA/C)	66	1 in 6.2 × 10^22 (AA)
375 pg	STRs	23	< 1 in 5.5 × 10^21 (SWH)	8	1 in 9.0 × 10^5 (SWH)
	SNPs	82	1 in 3.4 × 10^34 (HA/M)	48	1 in 1.0 × 10^14 (AA)
125 pg	STRs	19	1 in 2.7 × 10^18 (SWH)	4	1 in 220 (C)

demineralized animal bone (*t* test, autosomal STR average *P* values 0.217, SNP average *P* value 0.552, with Table 1 of Supplemental Part 2, or SP2, providing values for individual comparisons). The variance observed between sequencing runs, amplifications, and donors was also not significant, and was therefore combined for an $n = 18$ for each input amount.

Aligning the observations from the sheared data to the sequenced human skeletal material revealed that the sheared DNA with an average size of 100 bp performed more closely to actual observations (Tables 2 and 3) and that combining identity SNPs with STRs increases the ability to make a highly significant identification in these types of cases. These human skeletal extracts were also run on a capillary electrophoresis (CE) instrument (Applied Biosystems 3130xl Genetic Analyzer) with the Fusion 6C kit (Promega PowerPlex®), and no significant difference was observed in terms of number of alleles between the MPS and CE systems (*t* test, $n = 14$, *P* value = 0.07237). Across the fourteen samples, alleles were observed above a stochastic threshold at sixteen different loci on the CE system. Of these loci, all but four were of comparable size or smaller in relation to the ForenSeq targets of the same loci.

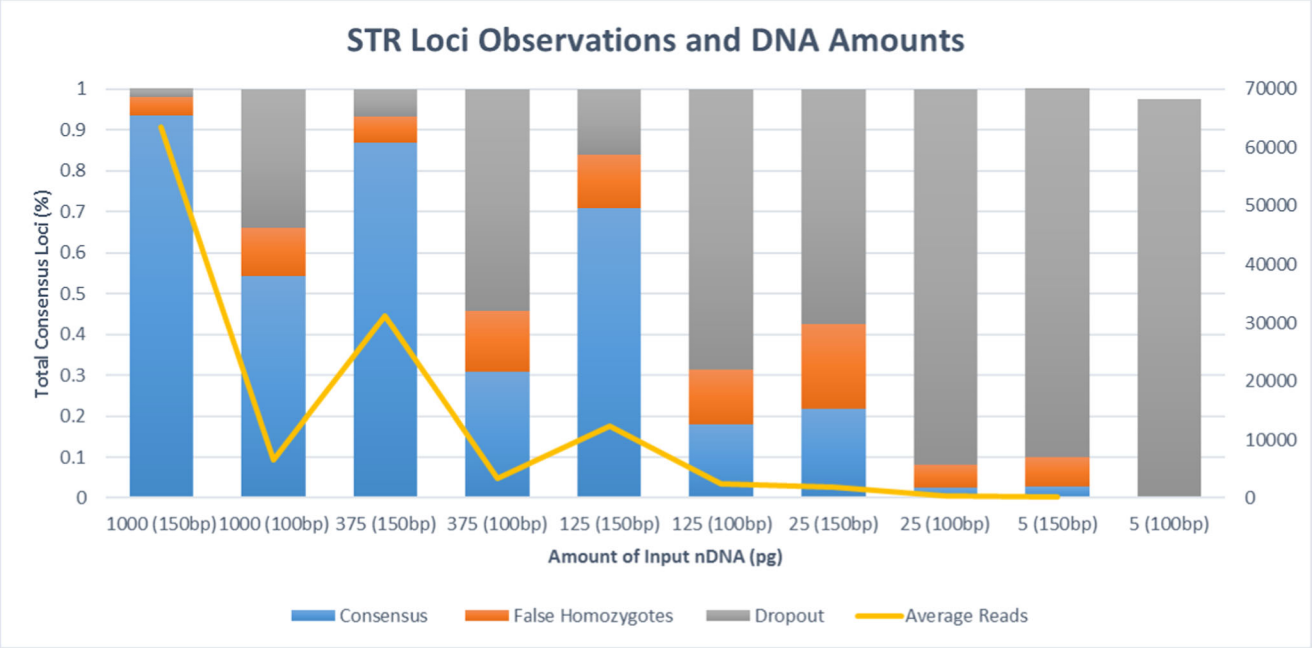
Threshold evaluation

The UAS system automatically sets thresholds for analysis of typical forensic samples based on coverage (read depth), but defaults to set parameters for samples with coverage below 650 reads. With older, degraded samples, it is expected that lower coverage will be observed. This evaluation was performed to investigate the impact this may have on the standardized UAS threshold. The average observed coverage for all input amounts and shear sizes, except for the 150 bp average samples with input amounts of 1 ng and 350 pg, was below 650 reads (Fig. 3). This means the default analysis settings would be used in most cases for analysis with the UAS software. The difference in read coverage between samples of an average size of 150 bps compared with samples of an average size of 100 bps was less pronounced for SNPs than for STRs (Fig. 3). The average heterozygote balance for STRs and SNPs of the 150 bp sized fragmented DNA (65.1% ± 22% for STRs and 74.12% ± 18% for SNPs) was above the threshold of 50% across input amounts, and the range (5 to 100% for STRs and 15 to 100% for SNPs) demonstrated clear stochastic sampling. Similar observations were found for the average heterozygote balance for STRs and SNPs for the 100 bp sized fragmented DNA (59% ± 24% for STRs and 70% ± 20% for SNPs, Supplemental Part 4 or SP4) and range (7 to 100% for STRs and 13 to 100% for SNPs). Elevated stutter levels were also observed for lower input amounts (SP4).

mtDNA control region sequencing

The extraction of DNA from 77 human skeletal remains from three different sites in Croatia resulted in an average recovery of 395 copies of mtDNA/μL (Fig. 4). Inhibition was not observed when assessed for a selection of the extracts (SP3). A subset of seventy of the extracts were sequenced, fifteen in replicate, for a total of 85 samples. Despite the low levels of recovered mtDNA, on average, 99.29% of the targeted CR was successfully sequenced at 50× coverage or more across all samples.

No evidence of contamination from scientists at the University of Split or Penn State University laboratories was observed that interfered with haplotype analysis. Quantification results demonstrated DNA degradation in all samples. All samples that were amplified in duplicate resulted in the same haplotypes. Of the 70 samples sequenced, 30 resulted in single source haplotypes while 40 showed background signal suggesting potential mixtures of two to four persons. In an effort to avoid overinterpreting these data, the mixtures were only evaluated for potential contamination. Of the forty samples that did not appear to be single source, thirteen were also sequenced on the ForenSeq system which



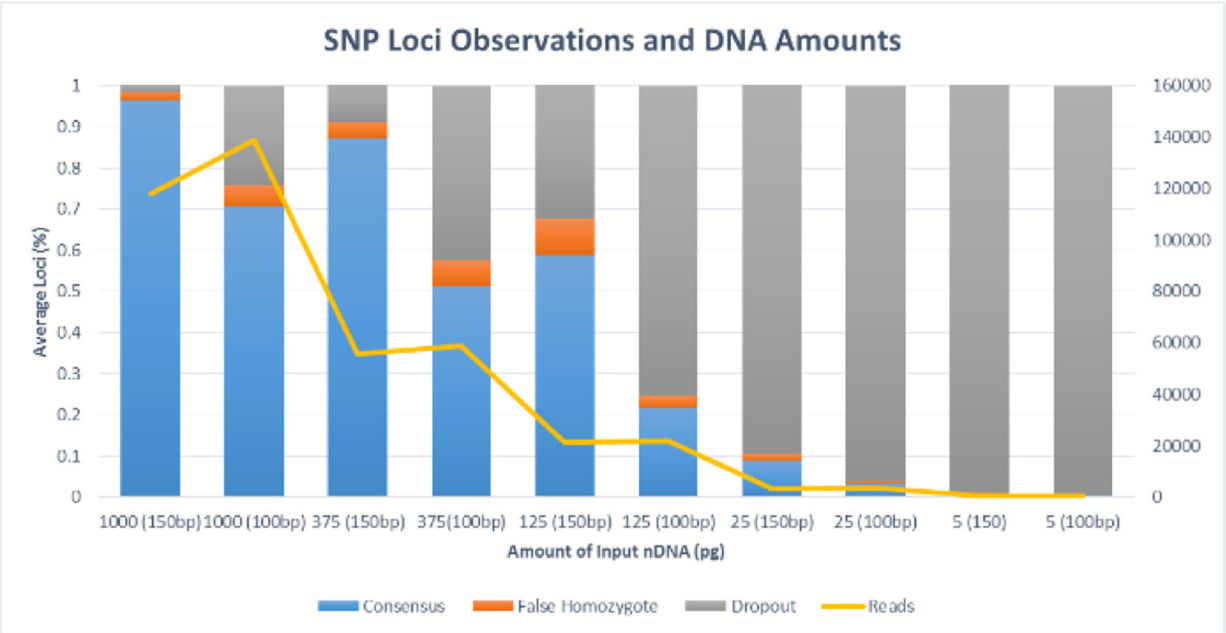


Fig. 2 Comparison between observed consensus, false homozygous, and dropout alleles at STR (upper) and SNP (lower) loci for fragmented DNA with an average size of 150 bp and 100 bp as compared with known sequences for 27 STR loci. The left *y*-axis shows the percent of the 27 loci that fall into each category while the right *y*-axis shows read depth; $n = 16$ for 150 bp and $n = 18$ for 100 bp

showed no evidence of a mixture. The haplogroups observed in the 30 single source samples reflect those commonly observed in medieval Europe (Table 4, [42–45]).

Discussion

In this study, we investigated the utility of both the ForenSeq system and the PowerSeq™ Mito kit for the identification of human skeletal remains. The sensitivity study demonstrated that for samples with an average fragment length of at least 150 bp, an identification can be confirmed down to an input of 125 pg and that combining the results of both SNP and STR typing is required for more degraded samples. The Promega PowerSeq™ Mito kit demonstrated its robustness with degraded samples with the successful 50× coverage across 99.29% of the mtCR for the 70 skeletal samples tested.

Table 3 The number of STR and SNP loci with alleles above interpretation (stochastic) and analytical threshold for fourteen human bone extracts and their respective input amounts for sequencing

Sample no.	nDNA input amount (pg)	Observed STR loci	Observed SNP loci
V-V 22	1 ng	0	0
V-V 14	1 ng	7	4
K 54.2	300	0	0
V-V 40	300	13	36
K 54.2	200	0	0
V-V 14	100	14	24
V-V 38	56.5	3	7
V-V 38	50	4	4
K 43.2	22	3	3
K 43.2	15.5	1	1
K 50	13	0	0
K 28.2	–	5	16
K 28.2	–	1	7
K 28.2	–	2	8

Previous sensitivity studies using DNA primer mix B from the ForenSeq kit resulted in 100% concordance down to approximately 125 pg of input genomic DNA for autosomal STRs and identity SNPs, and above 90% concordance below 62.5 pg of input [35, 36]. This sharply differs from observations in the current study with degraded DNA, which show below 90% concordance with input amounts less than 1 ng for samples averaging 150 bp in size and below 90% concordance for more fragmented samples at all tested input amounts, affirming the need for in-depth studies of these types of samples. The differences in these observations most likely are due to the fragmented nature of DNA used for our test. The potential for SNPs to provide supplementary information in highly degraded samples has been observed in other studies [36], but not consistently with normal input amounts of 1 ng [35]. The results of this study clearly demonstrate the ability of SNPs to increase identification potential for highly degraded DNA as they remained below the 50% line of dropout until 125 pg and 375 pg of input respectively for 150 bp and 100 bp nDNA. This is emphasized further when comparing the random match probabilities in Table 2 for both STRs and SNPs.

As expected when working with degraded and low copy number samples, stochastic effects were observed throughout the study. The observed impact of these effects was greater on STR loci than on SNP loci. With decreased input amounts, allelic dropout increased resulting in false homozygote calls for both STRs and SNPs (Fig. 2). Elevated stutter for the STR loci was also observed, but 81% of the observed stutter at each loci across input amounts stayed below the stutter thresholds (SP4 Fig. 4) and the average heterozygote balance stayed above 50% for input amounts of 125 pg and greater across most loci (SP4 Fig. 3). This heterozygous imbalance was less pronounced for SNPs than STRs, with averages above 60% for input amounts of 25 pg and higher (SP4 Fig. 2). These observations not only emphasize the need for duplicate amplification, as already recommended in other studies [38], but also demonstrate that SNPs not only provide supplementary information for when STR alleles cannot be typed but also provide more reliable results for cases with highly degraded samples.

The difference in observations from previous studies that focused on more pristine samples [35, 36] and between the different types of input DNA used in this study (two different sized sheared DNA and human skeletal remains) point out the importance of understanding the nature of the DNA being analyzed. The sensitivity study in this paper utilizes shearing with a Covaris, which results in a normal fragmentation curve (Fig. 1), while previous studies [46] show an exponential size decay. This means that actual samples would have a greater proportion of smaller DNA fragments than the artificially fragmented samples used in this study. This could also explain the difference between the observations of the sensitivity test and results with actual skeletal material. Previous tests on DNA fragmentation patterns have been performed in ancient DNA labs, focusing on much older samples than are typically seen in forensic casework. MPS opens the door to increase our understanding of how to work with highly degraded samples, and in turn, the ability to identify people from more and more challenging samples. Based on the results of this study, for STR loci, it would be recommended to use loci observed to be heterozygous for input amounts less than 375 pg with highly fragmented samples in order to avoid false homozygote genotyping. Overall, our results suggest that supplementing or switching from STRs to SNPs for human identification cases with samples comprised of low copy number and highly fragmented DNA (average size of 150 bps or less) would both increase identification potential and decrease the chance of reporting erroneous findings.

Successful MPS analysis of the mtDNA CR at 50× coverage and an average of 99.29% of the CR demonstrate the robustness of the PowerSeq™ Mito kit for cases with low quantity and highly degraded mtDNA. Recommendations for confirming the endogenous nature of the results for aDNA and forensics [8] were followed to the extent possible. Based on phylogenetic sense (Table 1), negative and contamination controls (SP3), quantification amounts (Fig. 4), and reproducibility through sequencing (SP3), the mtDNA results for samples with a single major profile were identified as endogenous. To avoid over interpretation, we decided not to deconvolute the observed mixture profiles in the remaining samples. The ability to confirm the endogenous nature of results based on molecular behavior (i.e., damage patterns and fragment size distribution) is limited for forensic

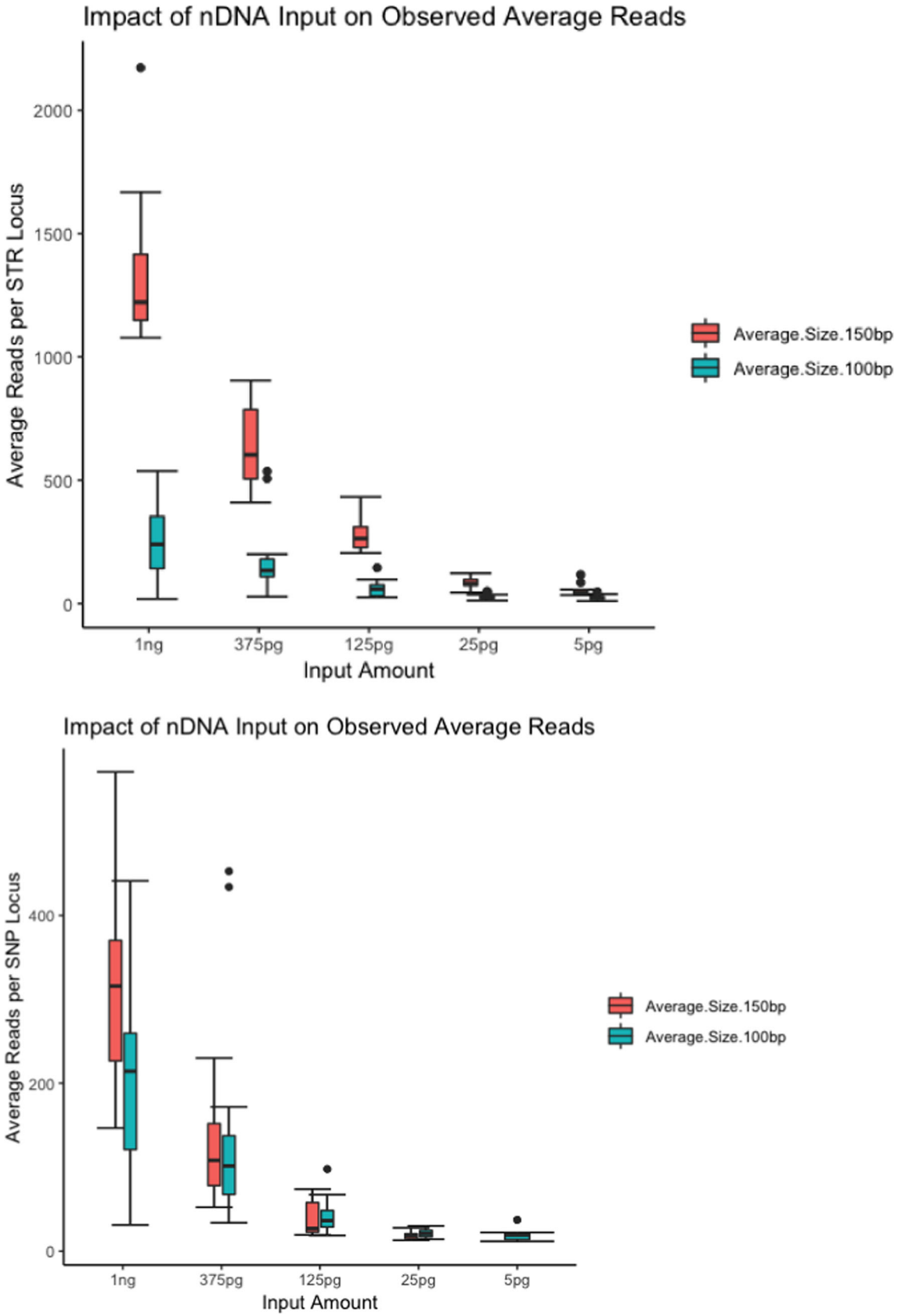


Fig. 3 Median total reads observed per autosomal STR (upper) and SNP (lower) locus across five different input amounts for DNA of an average size of 150 bp ($n = 12$) and 100 bp ($n = 18$)

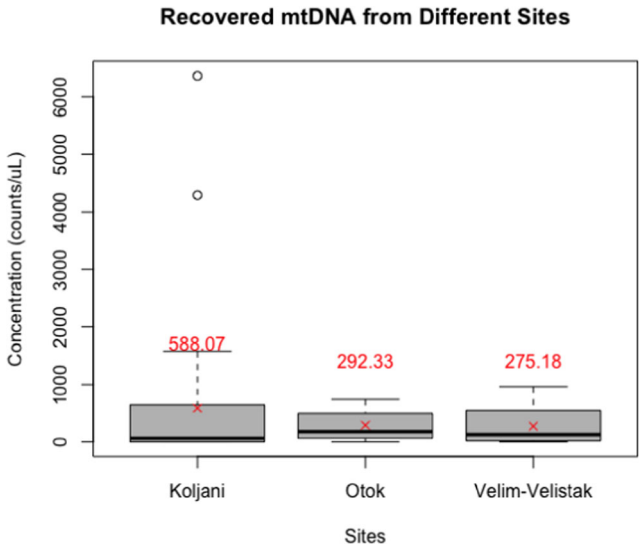


Fig. 4 The distribution of recovered mtDNA concentrations (counts/μL) from three different archeological sites in Croatia using a 60-bp target for quantification by qPCR. The black line represents the median, while the red "x" the respective mean. The mean for each site is also noted above each site in red

science due to the use of PCR-based sequencing and rare access to shotgun sequencing in crime laboratories. Current studies are beginning to characterize expected damage patterns for more recent samples [47, 48], but more extensive studies need to be performed before these characterizations are implemented widely. MPS has already been used to characterize these damage problems [48, 49] and has been suggested as a way to help in deconvoluting mixtures [25]. For future studies, taking smaller samples of powder from various locations on a single bone [1] may help to differentiate between contamination, damage, and endogenous DNA.

Human identification cases related to mass disasters or grave sites can involve thousands of unidentified remains, while missing person cases may only involve one individual. In order to efficiently handle this range in a number of samples, developed methods must be scalable and cost-effective at both ends of the sample spectrum. This includes not only the cost of reagents but also the hands-on time required by technicians and analysts. With the implementation of robots and automation [50] into these pipelines, there is the potential to improve work efficiency.

Extensive studies and validations have been completed for multiple MPS kits and typical forensic samples [31–36], demonstrating the strength of the application to forensic casework. This study also demonstrates the ability for the forensic field to more confidently and accurately work with more degraded samples by including SNPs. For those cases with highly degraded DNA, increased studies on the characterization of DNA recovered from more recent human skeletal material will continue to increase our ability to correctly analyze the results from these types of samples and provide better guidelines for analysis. This characterization will only be possible through the implementation of MPS in order to evaluate more completely the fragment size distributions and damage patterns, as has been done for aDNA samples.

Table 4 Observed haplogroup frequencies from the thirty samples that produced a single donor mtDNA profile in comparison with previously observed modern [43] and medieval [45] haplogroup frequencies. The haplogroup M is not typically associated with European samples. *V-V*, Velim-Velistak; *O*, Otok Vuletina Rupa

Haplogroup	Site			Modern Croatia ($n = 488$)	Medieval Croatia ($n = 7$)
	Koljani ($n = 15$)	V-V ($n = 14$)	O ($n = 1$)		
F	0.00	0.00	0.00	0.20	0.00
H	73.33	21.43	100.00	45.29	62.00
HV	6.67	0.00	0.00	4.07	4.00
I	0.00	7.14	0.00	2.61	0.00
J	0.00	0.00	0.00	9.83	15.00
K	6.67	0.00	0.00	4.30	3.00
L	0.00	7.14	0.00	0.20	0.00
N	0.00	0.00	0.00	0.82	0.00
R	0.00	14.29	0.00	0.00	0.00
T	0.00	7.14	0.00	5.98	2.00
U	6.67	35.71	0.00	18.85	13.00
V	0.00	0.00	0.00	3.89	1.00
W	0.00	0.00	0.00	1.84	0.00
X	6.67	0.00	0.00	1.84	0.00
M	0	7.143	0	0	0

Conclusions

The combined observations for sequencing results with both the PowerSeq™ Mito kit and Illumina's ForenSeq kit for highly fragmented DNA demonstrate the success of using MPS with highly degraded samples for STR, iiSNP, and mtDNA analyses. The findings using the PowerSeq™ Mito kit resulted in at least 50× coverage across 99.29% of the mtCR for 70 samples from the ninth to eighteenth centuries. The assessment of the ForenSeq system successfully demonstrated its potential to improve identification capacity when working with degraded DNA. In cases where the average fragment size of recovered nDNA is 150 bp, identification can be confirmed for total DNA inputs from 125 pg to 1 ng. For more highly fragmented DNA, at an average size of 100 bp, both SNP and STR typing are required for a significant match for total DNA inputs from 375 pg to 1 ng.

Acknowledgments The authors would like to thank the following people for their input and guidance throughout the project: Jennifer McElhoe and Charity Holland (Holland Research laboratory at the Pennsylvania State University), Molly Rathbun (The Pennsylvania State University), Sylvain Amory, Stefan Prost, Rene Huel (ICMP), and Ana Bilic (ICMP). All sheared samples were prepared with the help of the Penn State Genomics Core Facility – University Park, PA.

Funding information This project was partially funded thanks to the Carol DeForest Grant from the Northeastern Association of Forensic Science, Illumina, and the National Institute of Justice (2015-DN-BX-K025).

Compliance with ethical standards

Informed consent Informed consent was obtained from all individual participants included in the study. Skeletal material was provided with permission from the University of Split.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The Ethical committee from the Medical School (University of Split) approved the research on skeletal remains (approval no. 45-1106 from 6 March 2006).

References

1. Holland MM, Cave CA, Holland CA, Bille TW (2003) Development of a quality, high throughput DNA analysis procedure for skeletal samples to assist with the identification of victims from the World Trade Center attacks. Croat Med J 44:264–272
2. Forensic Genetics Policy Initiative (2017) Establishing best practices for forensic DNA databases. http://dnapolicyinitiative.org/wp-content/uploads/2017/08/BestPractice-Report-plus-cover-final.pdf. Accessed September 2017
3. International Committee of the Red Cross (2009) Missing People, DNA analysis and identification of human remains: a guide to best practice in armed conflicts and other situations of armed violence. https://www.icrc.org/en/doc/assets/files/other/icrc_002_4010.pdf. Accessed September 2017
4. Parsons TJ, Huel RLM (2017) DNA and missing persons identification: practice, progress and perspectives. In: Handbook of forensic genetics: biodiversity and heredity in civil and criminal investigation, vol 2. World Scientific, Singapore, pp 337–373
5. Parsons TJ, Huel RML, Bajunovic Z, Rizvic A (2019) Large scale DNA identification: the ICMP experience. Forensic Sci Int Genet. https://doi.org/10.1016/j.fsigen.2018.11.08
6. Butler JM, Hill CR (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. Forensic Sci Rev 24(1): 15–26
7. Parson W, Gusmao L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, Parsons TJ (2014) DNA Commission of the International Society for Forensic Genetics: revised and extended guidelines for mitochondrial DNA typing. Forensic Sci Int Genet. https://doi.org/10.1016/j.fsigen.2014.07.010
8. Poinar HN (2003) The top 10 list: criteria of authenticity for DNA from ancient and forensic samples. Int Congr Ser 1239:575–579
9. Pilli E, Modi A, Serpico C, Achilli A, Lancioni H, Lippi B, Bertoldi F, Gelichi S, Lari M, Caramelli D (2013) Monitoring DNA contamination in handled vs directly excavated ancient human skeletal remains. PLoS One. https://doi.org/10.1371/journal.pone.0052524
10. Morild I et al (2015) Identification of missing Norwegian World War II soldiers, in Karelia Russia. J Forensic Sci 60(4):1104–1110. https://doi.org/10.1111/1556-4029.12767
11. Oh YN, Park J, Hong S (2017) Genetic analysis of old skeletal remains from Korean War victims using PowerPlex® Fusion 6C and MiniSTR system for human identification. Forensic Sci Int Gen Suppl Ser. https://doi.org/10.1016/j.fsigss.2017.09.068
12. Gaudio D, Betto A, Vanin S, De Guio A, Galasso A, Cattaneo C (2015) Excavation and study of skeletal remains from a World War I mass grave. Int J Osteoarchaeol 25(5):585–592. https://doi.org/10.1002/oa.2333
13. Eduardoff M, Xavier C, Strobl C, Casas-Vargas A, Parson W (2017) Optimized mtDNA control region primer extension capture analysis for forensically relevant samples and highly compromised mtDNA of different age and origin. Genes 8(10):237. https://doi.org/10.3390/genes8100237
14. Amory S, Huel R, Bilić A, Loreille O, Parsons TJ (2012) Automatable full demineralization DNA extraction procedure from degraded skeletal remains. Forensic Sci Int – Gen. 6(3):398–406. https://doi.org/10.1016/j.fsigen.2011.08.004
15. Krause J et al (2010) The complete mitochondrial DNA genome of an unknown hominin from Southern Siberia. Nature 464:894–897
16. Glocke I, Meyer M (2017) Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. Genome Res 27(7):1230–1237. https://doi.org/10.1101/gr.219675.116
17. Korlević P et al (2015) Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques 59(2):87–93. https://doi.org/10.2144/000114320
18. Yang DY, Waye JS, Dudar JC, Saunders SR (1998) Technical note: improved DNA extraction from ancient bone using silica-based spin columns. Am J Phys Anthropol 105(4):539–543. https://doi.org/10.1002/(SICI)1096-8644(199804)105
19. Dabney J, Knapp M, Glocke I, Gansauge M-T, Weihmann A, Nickel B, Valdiosera C, Garcia N, Paabo S, Arsuaga J-L, Meyer M (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc Natl Acad Sci 110(39):15758-15763
20. Gansauge M et al (2017) Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. Nucleic Acids Res 45(10). https://doi.org/10.1093/nar/gkx033

21. Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res 22(5):939–946. https://doi.org/10.1101/gr.128124.111
22. Prüfer K et al (2017) A high-coverage Neandertal genome from Vindija Cave in Croatia. Science 358(6363):655–658. https://doi.org/10.1126/science.aao1887
23. Carboni I, Fattorini P, Previderè C, Sorçaburu S, Cilieri, Iozzi S, Nutini AL, Contini E, Pescucci C, Torricelli F, Ricci U (2015) Evaluation of the reliability of the data generated by next generation sequencing from artificially degraded DNA samples. Forensic Sci Int Gen Suppl Ser. https://doi.org/10.1016/j.fsigss.2015.09.034
24. Yang Y, Xie B, Yan J (2014) Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics 12:190–197. https://doi.org/10.1016/j.gpb.2014.09.001
25. Holland M, McQuillan M, O'Hanlon K (2011) Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy. Croat Med J. 52(3):299–313. https://doi.org/10.3325/cmi.2011.52.299
26. Phillips C et al (2016) D5S2500 is an ambiguously characterized STR: identification and description of forensic microsatellites in the genomics age. Forensic Sci Int – Gen. 23:19–24. https://doi.org/10.1016/j.fsigen.2016.03.002
27. Eduardoff M et al (2015) Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM™. Forensic Sci Int – Gen. 17:110–121. https://doi.org/10.1016/j.fsigen.2015.04.007
28. Gettings KB, Kiesler KM, Vallone PM (2015) Performance of a next generation sequencing SNP assay on degraded DNA. Forensic Sci Int Genet 19:1–9. https://doi.org/10.1016/j.fsigen.2015.04.010
29. Parson W et al (2015) Massively parallel sequencing of complete mitochondrial genomes from hair shaft samples. Forensic Sci Int – Gen. 15(8–15). https://doi.org/10.1016/j.fsigen.2014.11.009
30. Gallimore JM, McElhoe JA, Holland MM (2018) Assessing heteroplasmic variant drift in the mtDNA control region of human hairs using an MPS approach. Forensic Sci Int – Gen. 32:7–17. https://doi.org/10.1016/j.fsigen.2017.09.013
31. Churchill J, Schmedes S, King J, Budowle B (2015) Evaluation of the Illumina® Beta Version ForenSeq™ DNA Signature Prep Kit for use in genetic profiling. Forensic Sci Int 20:20–29. https://doi.org/10.1016/j.fsigen.2015.09.009
32. Just T, Moreno L, Smerick J, Irwin J (2017) Performance and concordance of the ForenSeq™ system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. Forensic Sci Int – Gen. 28:1–9. https://doi.org/10.1016/j.fsig.2017.01.001
33. Silvia AL, Shugarts N, Smith J (2017) A preliminary assessment of the ForenSeq™ FGx System: next generation sequencing of an STR and SNP multiplex. Int J Legal Med 131(1):73–86. https://doi.org/10.1007/s00414-016-1457-6
34. Fei G, Yu J, Zhang L, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA signature prep kit on the MiSeq FGx™ forensic genomics system. Forensic Sci Int – Gen. 31:135–148. https://doi.org/10.1016/j.fsigen.2017.09.003
35. Jäger AC et al (2017) Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories. Forensic Sci Int – Gen. 28:52–70. https://doi.org/10.1016/j.fsigen.2017.01.011
36. Parson W, Xavier C (2017) Evaluation of the Illumina ForenSeq DNA Signature Prep Kit – MPS forensic application for the MiSeq FGx benchtop sequencer. Forensic Sci Int – Gen. 28:188–194. https://doi.org/10.1016/j.fsigen.2017.02.018
37. Whitaker JP, Cotton EA, Gill P (2001) A comparison the characteristics of profiles produced with the AMPFISTR SGM PLUS™ multiplex system for both standard and low copy number (LCN) STR DNA analysis. Forensic Sci Int 123(2–3):215–223
38. Benschop CC, van der Beek CP, Meiland HC, van Gorp AG, Westen AA, Sljen T (2011) Low template STR typing: effect of replicate number and consensus method on genotyping reliability and DNA database search results. Forensic Sci Int – Gen. 5(4):316–328. https://doi.org/10.1016/j.fsigen.2010.06.006
39. Budowle B, Chakraborty R, Carmody G, Monson KL (2000) Source attribution of a forensic DNA profile. Forensic Sci Commun 2(3)
40. Moretti T, Moreno L, Smerick J, Pignone M, Hizon R, Buckleton J, Bright J, Onorato A (2016) Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. Forensic Sci Int – Gen. 25:175–18.1. https://doi.org/10.1016/j.fsigen.2016.07.022
41. Holland MM, Pack ED, McElhoe JA (2017) Evaluation of GeneMarker® HTS for improved alignment of mtDNA MPS data, haplotype determination, and heteroplasmy assessment. Forensic Sci Int – Gen. 28:90–98. https://doi.org/10.1016/j.fsigen.2017.01.016
42. Csákyová V, Szécsényi-Nagy A, Bauerová M (2016) Maternal genetic composition of a medieval population from a Hungarian-Slavic contact zone in Central Europe. PLoS One. https://doi.org/10.1371/journal.pone.0151206
43. Šarac J et al (2014) Maternal genetic heritage of Southeastern Europe reveals a new Croatian isolate and a novel, local subbranching in the x2 haplogroup. Ann Hum Genet 78(3):178–194. https://doi.org/10.1111/ahg.12056
44. Csősz A et al (2016) Maternal genetic ancestry and legacy of 10th century AD Hungarians. Sci Rep 6:33446. https://doi.org/10.1038/srep33446
45. Bašić Z, Fox A et al (2015) Cultural inter-population differences do not reflect biological distances: an example of interdisciplinary analysis of population from Eastern Adriatic coast. Croat Med J 56(3):230–238. https://doi.org/10.3325/cmj.2015.56.230
46. Kistler L et al (2017) A new model for ancient DNA decay based on paleogenomic meta-analysis. Nucleic Acids Res 45(11):6310–6320. https://doi.org/10.1093/nar/gkx361
47. Arbeithuber B, Makova K, Tiemann-Boege I (2016) Artefactual mutations resulting from DNA lesions limit detection levels in ultrasensitive sequencing applications. DNA Res 23(6):547–559. https://doi.org/10.1093/dnares/dsw038
48. Rathbun MM et al (2017) Considering DNA damage when interpreting mtDNA heteroplasmy in deep sequencing data. Forensic Sci Int – Gen. 26:1–11. https://doi.org/10.1016/j.fsigen.2016.09.008
49. Gorden EM, Sturk-Andreaggi K, Marshal C (2018) Repair of DNA damage caused by cytosine deamination in mitochondrial DNA of forensic case samples. Forensic Sci Int – Gen. 34:257–264. https://doi.org/10.1016/j.fsigen.2018.02.015
50. Holland MM, Wendt F Evaluation of the RapidHIT™ 200 (2015) an automated human identification system for STR analysis of single source samples. Forensic Sci Int Genet 14:76–85. https://doi.org/10.1016/j.fsigen.2014.08.010