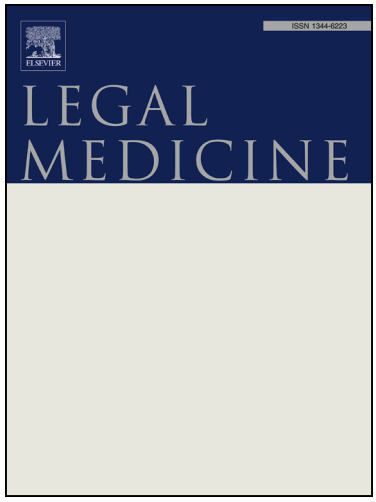
Journal Pre-Proof



Short Communication

Utility of the Ion S5™ and MiSeq FGx™ sequencing platforms to characterize challenging human remains

Kyleen Elwick, Magdalena M. Bus, Jonathan L. King, Joseph Chang, Sheree Hughes-Stamm, Bruce Budowle

PII: S1344-6223(19)30284-6
DOI: https://doi.org/10.1016/j.legalmed.2019.08.001
Reference: LEGMED 1623

To appear in: *Legal Medicine*

Received Date: 22 March 2019
Revised Date: 10 July 2019
Accepted Date: 3 August 2019

Please cite this article as: Elwick, K., Bus, M.M., King, J.L., Chang, J., Hughes-Stamm, S., Budowle, B., Utility of the Ion S5™ and MiSeq FGx™ sequencing platforms to characterize challenging human remains, *Legal Medicine* (2019), doi: https://doi.org/10.1016/j.legalmed.2019.08.001

Utility of the Ion S5™ and MiSeq FGx™ sequencing platforms to characterize challenging human remains

Kyleen Elwick [a, †], Magdalena M. Bus [b], Jonathan L. King [b], Joseph Chang [c], Sheree Hughes-Stamm[a, d], Bruce Budowle[b]

[a] Department of Forensic Science, Sam Houston State University, Huntsville, Texas, USA
[b] Center for Human Identification, University of North Texas Health Science Center, Fort Worth, Texas, USA
[c] Human Identification Division, Thermo Fisher Scientific, South San Francisco, California, USA
[d] School of Biomedical Sciences, University of Queensland, Brisbane, QLD, Australia

† Corresponding author at: Department of Forensic Science, Sam Houston State University, 1003 Bowers Blvd., Huntsville, TX 77340, USA.
E-mail: kee019@shsu.edu
ORCiD: 0000-0001-7174-710X

Abstract

Often in missing persons' and mass disaster cases the samples remaining for analysis are hard tissues such as bones, teeth, nails, and hair. These remains may have been exposed to harsh environmental conditions, which pose challenges for downstream genotyping. Short tandem repeat analysis (STR) via capillary electrophoresis (CE) is still the gold standard for DNA typing; however, a newer technology known as massively parallel sequencing (MPS) could improve upon our current techniques by typing different and more markers in a single analysis, and consequently improving the power of discrimination.

In this study, bone and tooth samples exposed to a variety of DNA insults (cremation, embalming, decomposition, thermal degradation, and fire) were assessed and sequenced using the Precision ID chemistry and a custom AmpliSeq™ STR and iiSNP panel on the Ion S5™ System, and the ForenSeq DNA Signature Prep Kit on the MiSeq FGx™ system, as well as the GlobalFiler™ PCR Amplification Kit on the 3500™ Genetic Analyzer.

The results demonstrated that using traditional CE-based genotyping performed as expected, producing a partial or full DNA profile for all samples, and that both sequencing chemistries and platforms were able to recover sufficient STR and SNP information from a majority of the same challenging samples. Run metrics including profile completeness and mean read depth produced good results with each system, considering the degree of damage of some samples. Most sample insults (except decomposed) produced similar numbers of alleles for both MPS systems. Comparable markers produced full concordance between the two platforms.

Keywords: Massively Parallel Sequencing, Ion S5™, MiSeq FGx™, missing persons, human remains, challenged remains

1 Introduction

Missing persons' cases, unidentified human remains, and mass disasters are problems encountered worldwide [1-3]. An overwhelming number of migrants and refugees have died or gone missing due to their efforts to cross borders or seas [4-6] or through human trafficking [7-9]. Routinely when identifying human remains in missing persons' cases, skeletal remains (bone, teeth) are the only samples available for DNA analysis [10-15]. However, some samples are more challenging to process than others due to their biological composition, environmental exposure (humidity, temperature, UV light, and microorganisms), DNA damage and/or degradation, the presence of inhibitors, and the possibility of contamination or comingled remains [10,13,16,17].

Currently, the technique of amplification of short tandem repeat (STR) loci combined with capillary electrophoresis (CE) is most commonly used to analyze such remains [16]. STRs are most frequently used because of their high discriminatory power. However, these severely compromised samples may not have suitable DNA fragment lengths to generate full CE-based STR profiles, decreasing the power of discrimination [18-20]. Therefore, other methods and genetic markers are being explored that may be more amenable to typing challenged samples. Single nucleotide polymorphisms (SNPs) using massively parallel sequencing (MPS) may be applicable for some degraded samples [21-25]. MPS demonstrates promising capabilities such as large sample multiplexing, improved mixture deconvolution, and the simultaneous analysis of different types of markers (e.g., identity-informative SNPs (iiSNPs), ancestry-informative SNPs (aiSNPs), STRs, and phenotypic-informative SNPs (piSNPs)) [21,26-32]. The use of multiple marker systems simultaneously (STRs and SNPs) can provide higher powers of discrimination and greater typing success with challenged samples than analyzing solely STRs. Furthermore, MPS can also detect sequence variation within the amplicons of these markers, many revealing SNPs within STR repeat regions and unreported microvariants [33-37], which were previously undetected using CE technology.

The goal of this study was to evaluate two MPS chemistries and platforms and compare their performance with traditional CE-based genotyping using challenged human remains that may be encountered in missing persons' cases. Bone and tooth samples were extracted using a total demineralization (TD) protocol [38]. The extracted DNA was quantified, STR-typed via CE, and then sequenced using both a custom AmpliSeq™ STR and iiSNP panel for degraded remains with Precision ID chemistry on the Ion S5™ system and the ForenSeq™ DNA Signature Prep Kit (using Primer Mix A) on the MiSeq FGx™. Performance between the two systems was determined by comparing read depth, heterozygote balance, and the total number of alleles or percentage of alleles. Percentage/number of alleles and the performance of the CODIS loci were compared between the three systems (two MPS systems and CE).

2 Materials and Methods

2.1 Sample Preparation

Bone (N = 19) (3x3 cm window cuts) and teeth (N = 5) samples from 14 cadavers were collected from the Applied Anatomical Research Center (AARC) at Sam Houston State University in Huntsville, Texas. These samples were subjected to a range of insults including cremation, embalming, decomposition, thermal degradation, and fire (Table 1). The remains were cremated in an oven at 900°C for 2.5 hours; embalmed remains were preserved with 30% glutaraldehyde for 880 days; teeth were thermally degraded in an oven at 232°C for 45 minutes; decomposed remains were surface exposed for 12-18 months; and burned remains were ignited with gasoline in a house (mock arson scene) and burned until they self-extinguished.

Bone sections were cleaned, chipped, and powdered as described in Zeng et al. [39]. Teeth were cleaned with a sterile toothbrush using 10% bleach, rinsing with DI H_2O, brushing with 70% ethanol, and rinsing again with DI H_2O. Teeth were individually wrapped in large task wipes, lightly crushed with a hammer, and powdered using a SPEX CertiPrep 6750 Freezer/Mill Cryogenic Grinder.

Three samples of each bone and tooth powders (300 mg) were extracted using a TD protocol [38]. Reference buccal swabs were collected before the cadavers were exposed to any insults (burning, decomposition, etc.). Reference swabs were then extracted using the AutoMate *Express*™ Forensic DNA Extraction System and PrepFiler *Express* (Thermo Fisher Scientific) according to the manufacturer's protocol [40]. Extracted DNA was quantified with Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA, USA) using a 7500 Real-Time PCR System (Thermo Fisher Scientific) according to manufacturer's instructions [41]. Quantification results and degradation indices for these samples are represented in Table 1.

2.2 CE-based STR Analysis

PCR amplification of STRs was performed using the GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific) on a ProFlex™ 96-well PCR System in accordance with the manufacturer's protocol [42]. DNA target input was 0.8 ng, whereas for low template samples (< 0.05 ng/µL) the full 15 µL of extract were amplified. Separation and detection were performed using a 3500™ Genetic Analyzer with POP-4™ polymer and a 36 cm capillary array (Thermo Fisher Scientific). Data were analyzed with GeneMapper™ ID-X v. 1.4 and an in-house excel workbook. An analytical threshold of 150 RFUs and a stochastic threshold of 600 RFUs were used to assign allele peaks. Average peak height (APH) was calculated by summing the peak heights at each locus of the sample replicates and dividing by the number of replicates. Average peak height ratios (APHR) were calculated by summing the peak height ratios at each locus for the sample replicates and dividing by the number of replicates. If allele or locus dropout occurred, the peak height ratio of that locus was given a value of zero. The standard deviation (SD) was calculated using the three replicates per sample.

2.3 Ion S5™ Sequencing

An automated library preparation method was chosen based on sample volume as up to 15 µL of DNA extract can be used with the Precision ID DL8 Kit (Thermo Fisher Scientific) on the Ion Chef™ System (Thermo Fisher Scientific), whereas manual library preparation is limited to 6 µL of extract. All low template samples (i.e., <0.16 ng/µL) were amplified and prepared using the DL8 kit and DNA samples greater than or equal to 0.16 ng were prepared manually using the Precision ID Library Kit (Thermo Fisher Scientific). A custom AmpliSeq™ STR and iiSNP primer panel (Thermo Fisher Scientific), including 32 STR markers, 1 Y-indel, 2 amelogenin sex markers, 41 iiSNPs, and 34 Y-SNPs, was used to amplify the extracted DNA. This panel consists of all STRs from the Precision ID

GlobalFiler NGS STR Panel v2 and 75 SNPs from the Precision ID Identity Panel (Thermo Fisher Scientific). All samples plus controls (N = 81) were sequenced in four runs. Two control samples (007 control DNA from Thermo Fisher Scientific) and two negative control samples (nuclease-free H_2O) were amplified with the manual library preparation. One control sample (007 control DNA) was amplified with each DL8 IonCode PCR plate. Libraries were quantified using the Ion Library TaqMan® Quantitation Assay (Thermo Fisher Scientific). Two pools of "high" quantity libraries were diluted to 50 pM, one pool of "mid-range" quantity libraries was combined neat at ~26 pM, and one pool of "low" quantity libraries was combined neat at ~12 pM (Supplemental Table 1). Templating and chip loading were performed using the Ion Chef™ System on a 530™ semiconductor chip, and sequencing was performed using the Ion S5™ Precision ID Chef and Sequencing Kit with the Ion S5™ System (Thermo Fisher Scientific). Data analyses were performed using Converge™ 2.0 (Thermo Fisher Scientific) and in-house excel workbooks. For STRs and SNPs, mean read depth was calculated by summing the total usable reads for the sample replicates and dividing by the number of replicates. STR and SNP heterozygote balance was calculated by averaging the heterozygote balance across sample replicates. Allele and locus dropout were treated as was done with CE-based dropout described previously. A minimum arbitrary detection threshold of 5X was used for both systems. A stochastic threshold for homozygote alleles was set at 10X.

2.4 MiSeq FGx™ Sequencing

Libraries were prepared using the ForenSeq™ DNA Signature Prep Kit (Verogen, Inc., San Diego, CA, USA) with Primer Mix A following the manufacturer's protocol [43]. Primer Mix A targets 27 autosomal STRs, 24 Y-STRs, and 7 X-STRs) and 94 iiSNPs. Samples with more than 0.2 ng of DNA were normalized to 0.2 ng and samples below 0.2 ng were used neat (0.01 ng – 0.1 ng) (5 µL maximum input). Normalized sample libraries including positive 2800M template control from the ForenSeq™ DNA Signature Prep Kit and a negative control (nuclease-free H_2O) were pooled in equal volumes according to the manufacturer's protocol [43] (Supplemental Table 2). Sequencing was performed on a MiSeq FGx™ (Illumina, San Diego, CA, USA) instrument using the MiSeq FGx™ Reagent Kit (Verogen) and the manufacturer's protocol [43]. Data analyses were performed using the ForenSeq Universal Analysis Software (Verogen), STRait Razor v2s [44], and in-house excel workbooks. The same data metrics calculated using the Ion S5 were also calculated in the same manner for the MiSeq.

3 Results and Discussion

3.1 Capillary Electrophoresis (CE)

3.1.1 Reportable Alleles

The number of reportable alleles was determined by the number of alleles present out of the total number of alleles expected. The expected number of alleles was determined by the total number of alleles in each panel. Full female profiles produced 44 alleles and full male profiles produced 46 alleles. All bone and teeth samples amplified with the GlobalFiler™ PCR Amplification Kit produced a STR profile to varying degrees of profile completeness. Reportable alleles ranged from 10 ± 3 to complete profiles across the samples (Supplemental Fig. 1). The thermally degraded teeth samples produced full profiles. The embalmed and cremated samples produced complete and near complete profiles, respectively. Mock arson burned samples produced profiles ranging from 27 ± 6 reportable alleles to full profiles. Decomposed skeletal remains produced the most degraded DNA profiles, ranging from 10 ± 3 to 32 ± 5.

3.1.2 Peak Height and Peak Height Ratios

APH across all samples ranged from 5154 ± 1952 relative fluorescence units (RFUs) to 210578 ± 8846 RFUs (Supplemental Fig. 2). Overall, the pattern observed across the sample types when considering the APH was consistent with the trend seen with STR profile completeness. The thermally

degraded samples produced the highest APHs ranging from 65684 ± 18897 RFUs to 210578 ± 8846 RFUs, while the decomposed remains produced the lowest APHs ranging from 5154 ± 1952 RFUs to 50751 ± 21380 RFUs.

APHRs showed a similar trend to both profile completeness and APH, decreasing from thermally degraded to decomposed samples. APHRs ranged from 8% ± 25% to 87% ± 11% across all samples, with just below half (46%) of the samples showing APHRs below 70% (Supplemental Fig. 3). The thermally degraded teeth ranged from 71% ± 19% to 87% ± 11% while the decomposed remains produced the least balanced profiles ranging from 8% ± 25% to 44% ± 42% APHRs.

3.1.3 Allelic Dropout

Allele dropout was determined by summing the number of alleles that dropped out at each locus across all samples. Allelic dropout was determined by comparison to a reference sample. Full profiles were obtained for CE samples. No allelic dropout was observed with the thermally degraded (five samples) and embalmed samples (one sample), the cremated sample (one sample) produced one dropout event at the DYS391 locus. The burned samples (eleven samples) showed 11% of allelic dropout, and the decomposed samples (six samples) resulted in the highest amount of allele dropout with 48% of alleles dropping out. As expected, the number of allelic dropout events increased as the size of the locus increased. Alleles at the loci D16S539, D7S820, FGA, CSF1PO, D18S51, D2S1338, TPOX, and SE33 experienced the most allelic dropout (Supplemental Fig. 4), which has been observed in several other studies [45-47]. Alleles at the SE33 locus dropped out most often with 43% of alleles dropping out overall. In contrast, only one instance of allelic dropout occurred at the Y INDEL and D22S1045 loci.

3.2 Massively Parallel Sequencing (MPS)

In this study, STR and SNP typing success was assessed via the number of reportable alleles, read depth, and heterozygote balance. Between the two platforms, all comparable results were concordant.

3.2.1 Reportable Alleles

The number of STR and SNP reportable alleles was calculated in the same manner as CE-based STRs. Full profiles were not available for all reference samples so a consensus method was used. For the Ion S5 system, full female STR profiles resulted in 64 alleles and SNP profiles produced 82 alleles. Full male STR profiles resulted in 67 alleles, and SNP profiles produced 116 alleles. For the MiSeq, full female STR profiles resulted in 70 alleles, full male STR profiles resulted in 88 alleles, and SNP profiles produced 188 alleles for both sexes.

All Ion S5 samples sequenced produced reportable alleles ranging from one allele to full profiles. However, using the MiSeq, two samples (one decomposed and one thermally degraded) produced no DNA profile. The decomposed sample that produced no profile with the MiSeq™ produced one allele with the Ion S5. In contrast, the thermally degraded sample that produced no profile with the MiSeq™ resulted in 98% of alleles using the Ion S5. Metrics are described for each sequencing run using the Ion S5 (Supplemental Table 1) and the MiSeq FGx (Supplemental Table 2).

Using the Ion S5, STR profiles ranged from 1 ± 1 allele to full profiles (Fig. 1a). All samples except decomposed remains produced >90% of alleles. Only three profiles produced below 50% of reportable alleles, all of which were decomposed skeletal remains. Reportable alleles for decomposed remains ranged from 1 ± 1 to 62 ± 3. For SNPs, the Ion S5 produced profiles ranging from 61 ± 55 alleles to full profiles (Fig. 1a). Similar to STRs, all samples except decomposed skeletal samples produced >90% of alleles. Decomposed remains resulted in profiles ranging from 47 ± 42 to 116 ± 0 alleles (Fig. 1a).

Using the MiSeq, STR profiles ranged from 0 alleles to complete profiles (Fig. 1b). Most thermally degraded, embalmed, and burned samples produced near complete or complete profiles. However, four samples (two burned and two thermally degraded) produced profiles <75%. Decomposed remains demonstrated the highest level of degradation (degradation index (DI) values from 1.6 to 18.5 and IPC ΔC_T values less than 1) with the number of reportable alleles ranging from 0 to 15 ± 14 alleles. No correlation was observed between the DI value and the number of alleles reported ($p > 0.05$). SNPs showed a similar pattern to STRs with reportable alleles ranging from 0 alleles to full profiles (Fig. 1b). Like STRs, embalmed and burned samples showed near complete or complete profiles. In contrast to STRs, all but one thermally degraded sample produced profiles ≤75%. The cremated sample produced similar results for SNPs and STRs showing ~70% of reportable alleles. The number of reportable alleles for decomposed skeletal samples ranged from 0 to 20 ± 16 alleles, demonstrating the most degradation as described above (Fig. 1b).

Overall, both sequencing platforms produced quality data for the types of challenged remains analyzed. On the Ion S5 the SNPs demonstrated higher profile completeness than that of STRs, producing ~10% more alleles than STRs overall (~93% ± 29% vs ~84% ± 16%), which was also observed in Xavier et al. using ancient bone samples [48] and is expected due to the smaller amplicon sizes. The severely compromised decomposed remains were especially difficult to analyze. However, with Precision ID DL8 library preparation 15 µL of low quantity sample were used increasing the DNA input amount compared to manual library preparation (6 µL). Therefore, the Precision ID DL8 library preparation on the Ion Chef was more flexible than manual library preparation when amplifying low template samples. In this study, there was a correlation observed between DNA input and percentage of reportable alleles ($p < 0.05$) (Supplemental Fig. 5a).

For most MiSeq samples, profile completeness between STRs and SNPs was comparable. In general, slightly more STR alleles were produced than SNPs (~66% ± 44% vs ~63% ± 44%). However, there were a few samples that demonstrated a >20% increase in STR profile completeness compared to SNPs. With MiSeq chemistry, only 5 µL of low template sample could be amplified resulting in a lower number of alleles being genotyped for compromised samples. Sample concentration would likely improve these results; DNA input versus profile completeness demonstrated a correlation in this study ($p < 0.05$) (Supplemental Fig. 5b).

CE-based STRs produced a greater percentage of total alleles than the Ion S5 for 2 out of 24 samples and for 8 out of 24 samples when using the MiSeq (Supplemental Fig. 6). For the less compromised remains (embalmed, cremated, and thermally degraded), all methods (CE and MPS) were comparable based on the common loci among the three systems. However, for the severely degraded remains (decomposed), the systems demonstrated variable results. In general, many of the burned remains were comparable for the three systems, but CE results were slightly lower. For the decomposed remains, CE results showed a lower profile completeness for most of the samples than those of the Ion S5, but for 2 samples, CE produced results when the other two systems did not (Supplemental Fig. 6). It is possible that these decomposed samples contained PCR inhibitors, affecting the MPS chemistries when maximum volume (15 µL) was amplified. Although the common loci of CE-based STRs are comparable to the MPS results, except for decomposed remains, MPS panels provided more information because they contain more markers (35 STR markers for the Ion S5 and 58 STRs for the MiSeq vs 24 STRs in GlobalFiler™, and many SNPs). Although, 2 decomposed samples produced almost no results for MPS, CE-based typing was able to recover >30% of alleles (Supplemental Fig. 6).

The success of typing 20 core CODIS loci was compared among the three platforms (CE, Ion S5, and MiSeq). Additionally, 12 out of 24 samples produced a full profile for all 3 platforms evaluated (Fig.

2). CE-generated STRs produced alleles ranging from 8 ± 3 to 40 (full profiles), with all but 3 samples yielding >50% of alleles. Samples sequenced using the Ion S5 system generated profiles ranging from 0 to 40 alleles, with all but three samples producing >50% of alleles. Two of the three samples producing <50% alleles generated no profile (both decomposed samples). CE-generated data produced 5 profiles more complete than those from the Ion S5 for the CODIS loci (Fig. 2). The 12 remaining samples that did not produce a full profile ranged from 0 to 40 alleles, with 4 out of 12 producing >50% of alleles using the MiSeq system. Similar to the Ion S5, only two samples failed to produce any alleles (one decomposed and one thermally degraded). However, 11 out of 12 CE-generated profiles showed more alleles for the CODIS loci. (Fig. 2). CE results demonstrated 2 out of 12 samples (Ion S5) and 7 out of 12 samples (MiSeq) with greater profile completeness than the MPS systems.

3.2.2 Read Depth

Using the Ion S5, mean read depth of STRs ranged from 19X ± 22X to 53648X ± 7873X, averaging ~17350X across all samples (Fig. 3a). The embalmed samples produced the highest mean read depth (53648X ± 7873X), thermally degraded and burned samples produced similar mean read depths, while the decomposed samples produced the lowest values (19X ± 22X to 7410X ± 6563X) (Fig 3a). For SNPs, mean read depth ranged from 2848X ± 2378X to 164801X ± 156816X, averaging ~74050X across all samples (Fig. 3a). Embalmed and cremated samples produced mean read depths >100000X, 162982X ± 23594X and 100850X ± 24823X, respectively, and decomposed samples produced the lowest mean read depths ranging from 2848X ± 2378X to 84588X ± 27600X (Fig. 3a).

For the MiSeq, mean read depth ranged from 0X to 202013X ± 23779X for STRs (Fig. 3b). Burned samples produced the highest read depth (3402X ± 2919X to 202013X to 23779X), with the decomposed remains ranging from 0X to 234X ± 298X with the lowest read depth (Fig. 3b). For SNPs, mean read depth ranged from 0X to 120061X ± 2690X (Fig. 3b). SNP mean read depth demonstrated a similar pattern to that of STRs with burned samples producing the highest read depth (2141X ± 1697X to 120061X ± 2690X), followed by the remaining sample insults and 0X to 109X ± 89X for decomposed samples (Fig. 3b).

In general, both platforms performed well and produced high sample read depth. Overall, Ion S5 SNPs produced higher read depth than STRs (74050X vs 17344X) for every sample. Both STRs and SNPs demonstrated proportional read depth across all samples types. The MiSeq STRs produced higher mean read depths than SNPs (70568X vs 31184X) for most samples. Only two samples produced higher mean read depth for SNPs than STRs, both decomposed remains. The MiSeq showed a large increase in mean read depth for burned samples and one thermally degraded sample for both STRs and SNPs. All other samples produced very low read depth compared to the burned samples, which may be due to run variability and/or a lower ability to type such samples.

Environmental insults to the DNA sample may damage the sample prior to extraction and affect the downstream success of certain types of samples. In this study, not all samples (mostly decomposed) were able to provide full MPS profiles on both MPS platforms due to damage or degradation of the DNA. In several studies [48-51], decomposed/ancient bone and tooth samples were analyzed using MPS. All studies demonstrated variable read depth and allele recovery for each sample, comparable to the results of this study.

3.2.3 Heterozygote Balance

Due to severe allelic imbalance (dropout) reads with less than 5X coverage were treated as dropout alleles resulting in a heterozygote balance of 0%. STRs produced using both platforms and SNPs produced on the MiSeq system resulted in multiple samples unable to calculate heterozygote balance.

Using the Ion S5 system STR heterozygote balance calculations could not be calculated for two samples (both decomposed) due to severe allelic or locus dropout. However, all SNP samples on the Ion S5 generated data that could be used for heterozygote balance calculations. For the MiSeq, five samples (one thermally degraded and four decomposed samples) demonstrated substantial allele dropout and heterozygote balance could not be calculated.

For Ion S5 STRs, heterozygote balance ranged from 0% to 81% ± 14%, averaging ~65% across all samples (Fig. 4a). However, most samples generated a heterozygote balance >70%. Out of 24 samples, 8 samples produced <70% heterozygote balance: two burned samples, and all decomposed samples. Thermally degraded teeth samples produced the highest heterozygote balance ranging from 76% ± 21% to 81% ± 14%. Embalmed and cremated samples produced comparable heterozygote balance of 77% ± 14% and 74% ± 19%, respectively. Burned samples resulted in heterozygote balance values ranging from 55% ± 25% to 77% ± 13-17%, and decomposed remains produced the least balanced profiles ranging from 0% to 67% ± 21% (Fig. 4a). With Ion S5 SNPs, heterozygote balance ranged from 32% ± 30% to 86% ± 2-4%, averaging ~79% across all samples (Fig. 4a). All but two samples (both decomposed) generated an average heterozygote balance >70%. Similar to STRs, thermally degraded samples produced the highest average heterozygote balance (83% ± 1% to 86% ± 4%), with the lowest heterozygote balances resulting from decomposed human remains, ranging from 32% ± 30% to 85% ± 1% (Fig. 4a).

Heterozygote balance for STRs on the MiSeq ranged from 0% to 82% ± 11% (Fig. 4b), averaging ~50% across all samples. Similar to profile completeness and read depth, burned samples generated the highest heterozygote balance ranging from 51% ± 40% to 82% ± 11%, while thermally degraded, embalmed, and cremated remains produced comparable heterozygote balances. Decomposed remains resulted in the lowest heterozygote balances ranging between 0% to 6% ± 22% (Fig. 4b). For SNPs, heterozygote balance ranged from 0% to 88% ± 10%, averaging ~51% across all samples (Fig. 4b). Analogous to STRs, SNPs demonstrated a similar trend with heterozygote balance (Fig. 4b).

In general, for the Ion S5, the majority of samples showed average heterozygote balances of >70%. Heterozygote balance averaged ~15% higher for SNPs than STRs. SNPs also demonstrated fewer samples with a heterozygote balance <70% compared with STRs (2 vs 8 samples). All samples except decomposed remains resulted in good heterozygote balance for SNPs. Overall, when sequenced on the MiSeq, just under half of the STR and SNP profiles demonstrated heterozygote balances <70%. For both STRs and SNPS, the burned samples demonstrated good heterozygous balance, while decomposed samples consistently demonstrated poor balance (<10%), and variability in all other samples.

Conclusions

Overall, MPS generated genetic data from challenged samples and provided more genetic data in 22 samples compared with the CE-based kit. Furthermore, a greater number of alleles will translate in greater power of discrimination. Although CE produced a usable DNA profile for identification purposes, based only on the 20 CODIS core loci for some more difficult samples, the greater number of loci included in MPS multiplexes allowed for more genetic information to be obtained from most samples barring the decomposed remains. Results suggest that MPS may recover more probative information from most samples, but CE-based methods were more robust for identifying skeletal samples. CE chemistry has been substantially developed over the past 25 years, while MPS kits for forensic applications have been around for less than five years. However, improvement in MPS panel design and chemistries could enhance performance.

References

1. da Silva LAF, Vilaça W, Azevedo D, Majella G, Silva IF, Silva BF (2009) Missing and unidentified persons database. Forensic Sci Int Genet Suppl Ser 2 (1):255-257.
2. Ritter N (2007) Missing persons and unidentified remains: the nation's silent mass disaster. NIJ Journal 256 (7)
3. Cattaneo C, Ritz-Timme S, Schutz HW, Collins M, Waite E, Boormann H et al. (2000) Unidentified cadavers and human remains in the EU: an unknown issue. Int J Leg Med 113 (3):2-3
4. Kovras I, Robins S (2016) Death as the border: Managing missing migrants and unidentified bodies at the EU's Mediterranean frontier. Polit Geogr 55:40-49.
5. Grant S, Solicitors HG (2015) Migrant deaths at sea: addressing the information deficit. Migration Policy Practice, V (1)
6. Last T, Spijkerboer T (2014) Tracking deaths in the Mediterranean.
7. Brayley H, Cockbain E, Laycock G (2011) The value of crime scripting: Deconstructing internal child sex trafficking. Policing: A Journal of Policy and Practice 5 (2):132-143
8. Miko FT Trafficking in persons: The US and international response. In, 2006. LIBRARY OF CONGRESS WASHINGTON DC CONGRESSIONAL RESEARCH SERVICE,
9. O'Neil K, Hamilton K, Papademetriou D (2005) Migration in the Americas. A paper prepared for the Policy Analysis and Research Programme of the Global Commission on International Migration Global Commission on International Migration, September
10. Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Eisenberg AJ, Budowle B et al. (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. Pathobiology 79 (5):228-238.
11. Loreille OM, Diegoli TM, Irwin JA, Coble MD, Parsons TJ (2007) High efficiency DNA extraction from bone by total demineralization. Forensic Sci Int Genet 1 (2):191-195.
12. Ferreira ST, Paula KA, Maia FA, Moraes AV (2013) A comparative study between two protocols for DNA extraction from bones. Forensic Sci Int Genet Suppl Ser 4 (1):e374-e375
13. Rothe J, Nagy M (2016) Comparison of two silica-based extraction methods for DNA isolation from bones. Leg Med (Tokyo) 22:36-41.
14. Karija Vlahovic M, Kubat M (2012) DNA extraction method from bones using Maxwell(R) 16. Leg Med (Tokyo) 14 (5):272-275.
15. Hartman D, Drummer O, Eckhoff C, Scheffer JW, Stringer P (2011) The contribution of DNA to the disaster victim identification (DVI) effort. Forensic Sci Int 205 (1-3):52-58.
16. Alonso A, Andelinovic S, Martin P, Sutlovic D, Erceg I, Huffine E et al. (2001) DNA typing from skeletal remains: evaluation of multiplex and megaplex STR systems on DNA isolated from bone and teeth samples. Croat Med J 42 (3):260-266
17. Mundorff AZ, Bartelink EJ, Mar-Cash E (2009) DNA preservation in skeletal elements from the World Trade Center disaster: recommendations for mass fatality management. J Forensic Sci 54 (4):739-745.
18. Freire-Aradas A, Fondevila M, Kriegel AK, Phillips C, Gill P, Prieto L et al. (2012) A new SNP assay for identification of highly degraded human DNA. Forensic Sci Int Genet 6 (3):341-349.
19. Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W (2006) A new "miniSTR-multiplex" displaying reduced amplicon lengths for the analysis of degraded DNA. Int J Leg Med 120 (2):115-120

20. Hughes-Stamm SR, Ashton KJ, van Daal A (2011) Assessment of DNA degradation and the genotyping success of highly degraded samples. Int J Legal Med 125 (3):341-348.
21. Seo SB, King JL, Warshauer DH, Davis CP, Ge J, Budowle B (2013) Single nucleotide polymorphism typing with massively parallel sequencing for human identification. Int J Legal Med 127 (6):1079-1086.
22. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456 (7218):53-59.
23. Wang J, Wang W, Li R, Li Y, Tian G, Goodman L et al. (2008) The diploid genome sequence of an Asian individual. Nature 456 (7218):60-65.
24. Daniel R, Santos C, Phillips C, Fondevila M, van Oorschot RA, Carracedo A et al. (2015) A SNaPshot of next generation sequencing for forensic SNP analysis. Forensic Sci Int Genet 14:50-60.
25. Kidd KK, Kidd JR, Speed WC, Fang R, Furtado MR, Hyland FC et al. (2012) Expanding data and resources for forensic use of SNPs in individual identification. Forensic Sci Int Genet 6 (5):646-652.
26. Guo F, 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 Genet 31:135-148
27. Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X et al. (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler NGS STR Panel and the Ion PGM System. Forensic Sci Int Genet 31:126-134.
28. Budowle B, van Daal A (2008) Forensically relevant SNP classes. Biotechniques 44 (5):603-608, 610.
29. Kayser M, Schneider PM (2009) DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. Forensic Sci Int Genet 3 (3):154-161.
30. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A et al. (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet 7 (1):98-115.
31. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci Int Genet 5 (3):170-180. doi:10.1016/j.fsigen.2010.02.004
32. Churchill JD, Schmedes SE, King JL, Budowle B (2016) Evaluation of the Illumina® beta version ForenSeq™ DNA signature prep kit for use in genetic profiling. Forensic Sci Int Genet 20:20-29
33. Pinto LM, de Oliveira CL, dos Santos LL, Tarazona-Santos E (2014) Molecular characterization and population genetics of non-CODIS microsatellites used for forensic applications in Brazilian populations. Forensic Sci Int-Gen 9:E16-E17.
34. Novroski NMM, Woerner AE, Budowle B (2018) Potential highly polymorphic short tandem repeat markers for enhanced forensic identity testing. Forensic Sci Int Genet 37:162-171.
35. Gettings KB, Aponte RA, Vallone PM, Butler JM (2015) STR allele sequence variation: Current knowledge and future issues. Forensic Sci Int Genet 18:118-130.
36. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA et al. (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. Forensic Sci Int Genet 21:15-21.

37. King JL, Churchill JD, Novroski NMM, Zeng X, Warshauer D, Seah LH et al. (2018) Increasing the discrimination power of ancestry-and identity-informative SNP loci within the ForenSeq™ DNA Signature Prep Kit. Forensic Sci Int Genet
38. Loreille OM, Parr RL, McGregor KA, Fitzpatrick CM, Lyon C, Yang DY et al. (2010) Integrated DNA and fingerprint analyses in the identification of 60-year-old mummified human remains discovered in an Alaskan glacier. J Forensic Sci 3:813-818.
39. Zeng X, Elwick K, Mayes C, Takahashi M, King JL, Gangitano D et al. (2018) Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success. Int J Leg Med.
40. Scientific TF (2018) AutoMate *Express*™ Instrument User Guide. F edn. Thermo Fisher Scientific,
41. Scientific TF (2017) Quantifiler™ HP and Trio DNA Quantification Kits: User Guide. F edn. Thermo Fisher Scientific,
42. Scientific TF (2015) GlobalFiler™ PCR Amplification Kit. Thermo Fisher Scientific, Thermo Fisher Scientific
43. Illumina (2015) ForenSeq DNA Signature Prep Reference Guide. vol 15049528, v01 edn., San Diego, CA
44. King JL, Wendt FR, Sun J, Budowle B (2017) STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems. Forensic Sci Int Genet 29:21-28.
45. Elwick K, Mayes C, Hughes-Stamm S (2018) Comparative Sensitivity and Inhibitor Tolerance of GlobalFiler® PCR Amplification and Investigator® 24plex QS Kits for Challenging Samples. Leg Med
46. Parys-Proszek A, Wrobel M, Marcinska M, Kupiec T (2018) Dual amplification strategy for improved efficiency of forensic DNA analysis using NGM Detect, NGM or Globalfiler kits. Forensic Sci Int Genet 35:46-49.
47. Tucker VC, Hopwood AJ, Sprecher CJ, McLaren RS, Rabbach DR, Ensenberger MG et al. (2011) Developmental validation of the PowerPlex® ESI 16 and PowerPlex® ESI 17 Systems: STR multiplexes for the new European standard. Forensic Sci Int Genet 5:436-448.
48. Xavier C, Parson W. (2017) Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit - MPS forensic application for the MiSeq FGx™ benchtop sequencer. Forensic Sci Int Genet 28: 188-194.
49. Ambers AD, Churchill JD, King JL, Stoljarova M, Gill-King H, Assidi M et al. (2016) More comprehensive forensic genetic marker analyses for accurate human remains identification using massively parallel DNA sequencing. BMC Genom 17(Suppl 9):750.
50. Strobl C, Eduardoff M, Buś MM, Allen M, Parson W. (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. Forensic Sci Int Genet 35: 21-25
51. Buś MM, Lembring M, Kjellström A, Strobl C, Zimmermann B, Parson W et al. (2019) Mitochondrial DNA analysis of a Viking Age mass grave in Sweden. Forensic Sci Int Genet.

Figure Captions

Fig. 1 Profile completeness (number of alleles) of STRs and SNPs sequenced using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ systems for 24 challenged human remains samples. Data presented as average + SD (N = 3). Dotted red line denotes the maximum number of male STRs and SNPs for each sequencing platform. Minimum detection threshold of 5X coverage

Fig. 2 Profile completeness (number of alleles) of the 20 core CODIS STRs using CE, the Ion S5™, and the MiSeq FGx™ systems for 24 challenging human remains samples. Data presented as average + SD (N = 3)

Fig. 3 Mean read depth of STRs and SNPs sequencing using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (triplicate). Minus error bars represent STR data and plus error bars represent SNP data. SNP mean read depth is determined by subtracting STR mean read depth

Fig. 4 Heterozygote balance of STRs and SNPs sequenced using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (N = 3). Minus error bars represent the STR data and plus error bars represent the SNP data

Figure Captions

Fig. 1 Profile completeness (number of alleles) of STRs and SNPs sequenced using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ systems for 24 challenged human remains samples. Data presented as average + SD (N = 3). Dotted red line denotes the maximum number of male STRs and SNPs for each sequencing platform. Minimum detection threshold of 5X coverage

Fig. 2 Profile completeness (number of alleles) of the 20 core CODIS STRs using CE, the Ion S5™, and the MiSeq FGx™ systems for 24 challenging human remains samples. Data presented as average + SD (N = 3)

Fig. 3 Mean read depth of STRs and SNPs sequencing using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (triplicate). Minus error bars represent STR data and plus error bars represent SNP data. SNP mean read depth is determined by subtracting STR mean read depth

Fig. 4 Heterozygote balance of STRs and SNPs sequenced using the **a.)** Ion S5™ and **b.)** MiSeq FGx™ for 24 challenging human remains samples. Data presented as average ± SD (N = 3). Minus error bars represent the STR data and plus error bars represent the SNP data

Highlights

- Two MPS chemistries and platforms were compared to traditional CE-based genotyping for use with challenged human remains.
- MPS generated more genetic data in 22 samples compared to CE.
- Using only the 20 CODIS loci, CE produced more alleles than MPS for some more difficult samples.
- MPS may recover more probative information because of the greater number of markers available.
- CE methods were more robust for identifying challenging human remains.

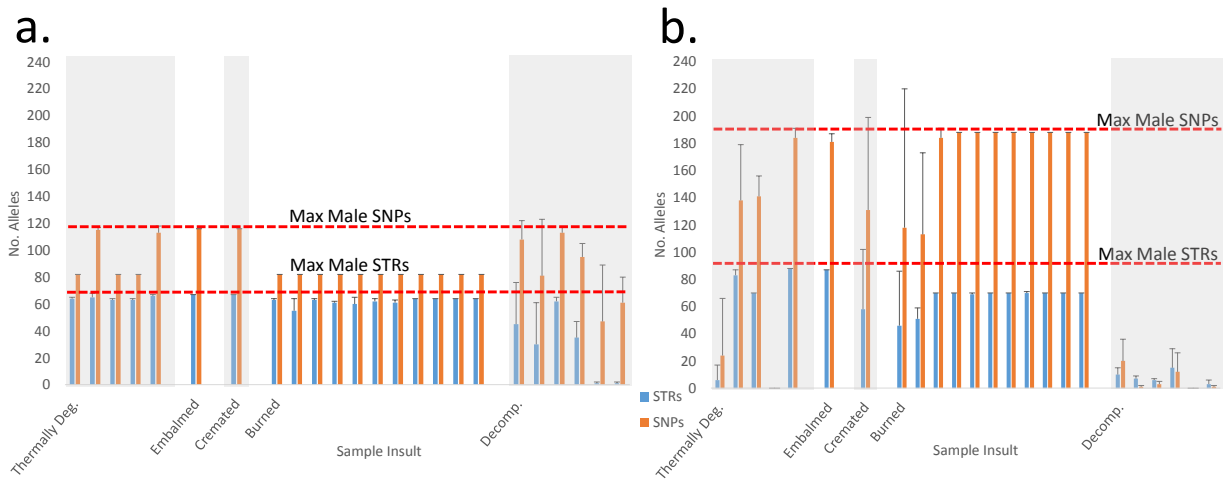


Fig. 1

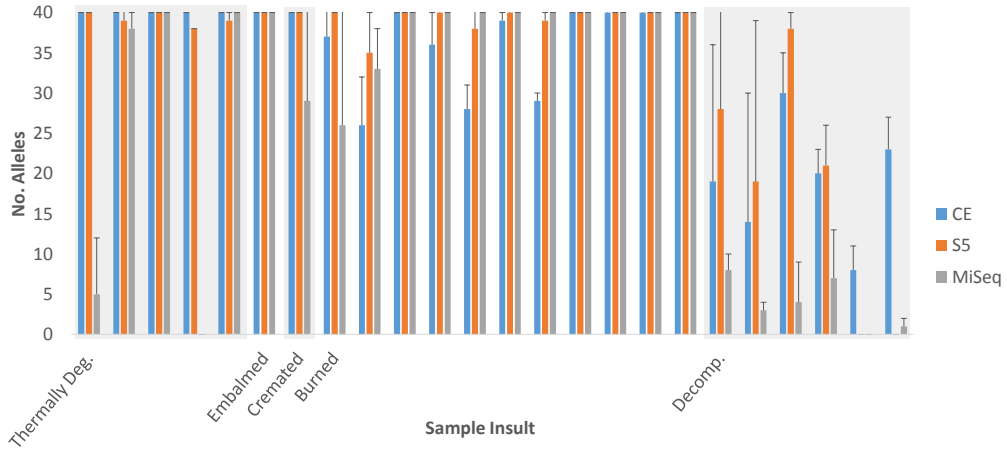


Fig. 2

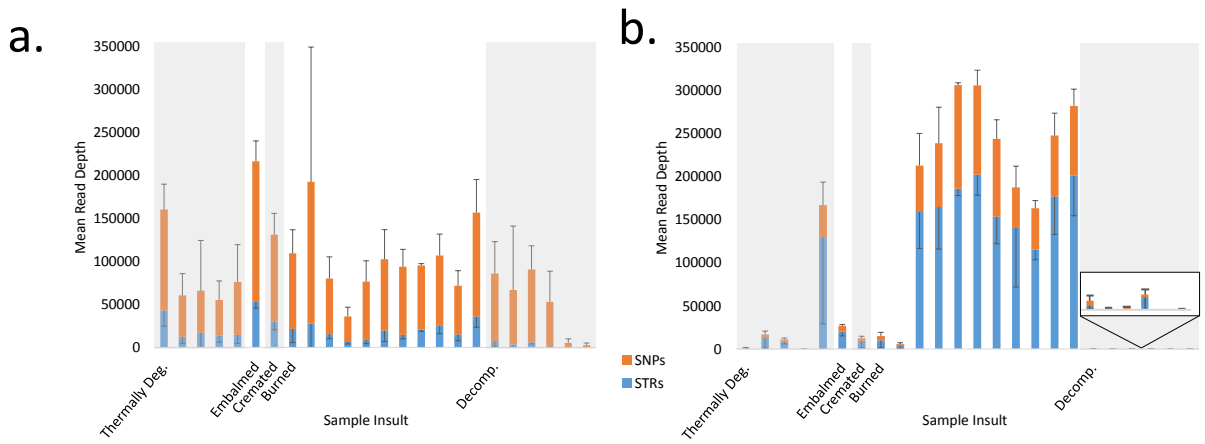


Fig. 3

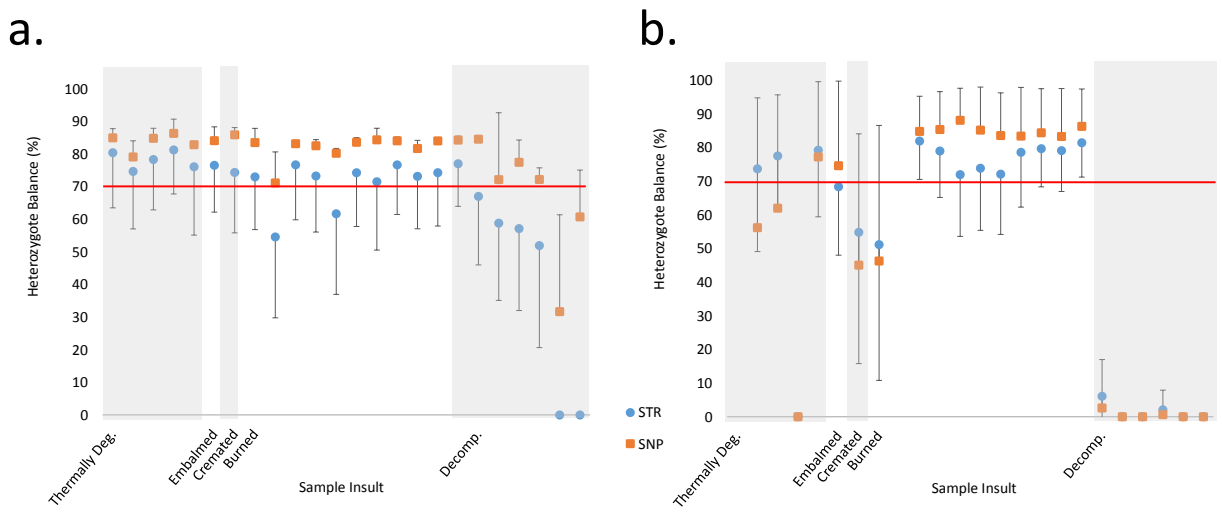


Fig. 4

Cadaver	Sex	Bone	Insult	Avg. Concentration (ng/µL)	Avg. Degradation Index (DI)
1	Female	Premolar	Thermally Degraded	3.89 ± 1.57	2.86 ± 0.98
2	Male	Molar	Thermally Degraded	22.66 ± 10.44	1.66 ± 0.51
3	Female	Premolar	Thermally Degraded	27.20 ± 4.26	1.87 ± 0.20
4	Female	Premolar	Thermally Degraded	4.07 ± 0.78	2.12 ± 0.10
5	Male	Molar	Thermally Degraded	46.17 ± 6.20	1.64 ± 0.14
6	Male	Femur	Embalmed	2.33 ± 0.47	1.11 ± 0.06
7	Male	Top Vertebral Arch	Cremated	3.17 ± 1.11	3.10 ± 0.47
8	Female	Femur	Burned	15.64 ± 21.20	3.94 ± 3.78
9	Female	Fibula	Burned	0.017 ± 0.004	3.06 ± 1.33
10	Female	Femur	Burned	7.21 ± 1.06	2.56 ± 0.25
		Humerus	Burned	6.35 ± 0.60	2.53 ± 0.15
		Tibia	Burned	32.72 ± 9.33	8.66 ± 0.88
11	Female	Femur	Burned	2.21 ± 0.36	2.70 ± 0.18
		Humerus	Burned	1.23 ± 0.23	3.79 ± 0.09
		Tibia	Burned	34.37 ± 5.56	2.80 ± 0.13
12	Female	Femur	Burned	4.34 ± 1.46	1.73 ± 0.05
		Humerus	Burned	3.07 ± 0.74	2.15 ± 0.37
		Tibia	Burned	1.63 ± 0.50	1.43 ± 0.11
13	Male	Femur	Decomposed	0.08 ± 0.03	2.19 ± 0.42
		Humerus	Decomposed	0.09 ± 0.02	1.79 ± 0.28
		Tibia	Decomposed	0.96 ± 0.38	16.24 ± 2.85
14	Male	Femur	Decomposed	0.025 ± 0.01	3.87 ± 0.52
		Humerus	Decomposed	0.015 ± 0.002	6.78 ± 4.12
		Tibia	Decomposed	0.019 ± 0.002	3.28 ± 0.32