

# PUBLISHED VERSION

Jennifer E L Templeton, Paul M Brotherton, Bastien Llamas, Julien Soubrier, Wolfgang Haak, Alan Cooper and Jeremy J Austin

**DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification**

Investigative Genetics, 2013; 4(26):1-13

© 2013 Templeton et al.; licensee BioMed Central Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Originally published at:

<http://doi.org/10.1186/2041-2223-4-26>

## PERMISSIONS

<http://creativecommons.org/licenses/by/2.0/>



Attribution 2.0 Generic (CC BY 2.0)

This is a human-readable summary of (and not a substitute for) the [license](#).

[Disclaimer](#)



### You are free to:

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material

for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

### Under the following terms:



**Attribution** — You must give **appropriate credit**, provide a link to the license, and **indicate if changes were made**. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**No additional restrictions** — You may not apply legal terms or **technological measures** that legally restrict others from doing anything the license permits.

<http://hdl.handle.net/2440/83202>

**METHODOLOGY**

**Open Access**

# DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification

Jennifer E L Templeton<sup>1,2</sup>, Paul M Brotherton<sup>1,3</sup>, Bastien Llamas<sup>1</sup>, Julien Soubrier<sup>1</sup>, Wolfgang Haak<sup>1</sup>, Alan Cooper<sup>1</sup> and Jeremy J Austin<sup>1,4\*</sup>

## Abstract

**Background:** Mitochondrial DNA (mtDNA) typing can be a useful aid for identifying people from compromised samples when nuclear DNA is too damaged, degraded or below detection thresholds for routine short tandem repeat (STR)-based analysis. Standard mtDNA typing, focused on PCR amplicon sequencing of the control region (HVS I and HVS II), is limited by the resolving power of this short sequence, which misses up to 70% of the variation present in the mtDNA genome.

**Methods:** We used in-solution hybridisation-based DNA capture (using DNA capture probes prepared from modern human mtDNA) to recover mtDNA from post-mortem human remains in which the majority of DNA is both highly fragmented (<100 base pairs in length) and chemically damaged. The method 'immortalises' the finite quantities of DNA in valuable extracts as DNA libraries, which is followed by the targeted enrichment of endogenous mtDNA sequences and characterisation by next-generation sequencing (NGS).

**Results:** We sequenced whole mitochondrial genomes for human identification from samples where standard nuclear STR typing produced only partial profiles or demonstrably failed and/or where standard mtDNA hypervariable region sequences lacked resolving power. Multiple rounds of enrichment can substantially improve coverage and sequencing depth of mtDNA genomes from highly degraded samples. The application of this method has led to the reliable mitochondrial sequencing of human skeletal remains from unidentified World War Two (WWII) casualties approximately 70 years old and from archaeological remains (up to 2,500 years old).

**Conclusions:** This approach has potential applications in forensic science, historical human identification cases, archived medical samples, kinship analysis and population studies. In particular the methodology can be applied to any case, involving human or non-human species, where whole mitochondrial genome sequences are required to provide the highest level of maternal lineage discrimination. Multiple rounds of in-solution hybridisation-based DNA capture can retrieve whole mitochondrial genome sequences from even the most challenging samples.

**Keywords:** Mitochondrial DNA, Degraded DNA, Ancient DNA, DNA hybridisation, DNA enrichment, Forensic science, Next-generation sequencing

\* Correspondence: [jeremy.austin@adelaide.edu.au](mailto:jeremy.austin@adelaide.edu.au)

<sup>1</sup>Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia

<sup>4</sup>Sciences Department, Museum Victoria, Carlton Gardens, Melbourne, Vic 3001, Australia

Full list of author information is available at the end of the article

## Background

Nuclear DNA short tandem repeat (STR) profiling is currently the preferred method for human identification in forensic practice [1]. However, analysis of low copy number (LCN) and highly damaged or degraded DNA from trace sources or poorly preserved human remains is challenging due to stochastic effects and can often fail completely [2]. Typical complications observed in the analysis of trace amounts of DNA include issues with contamination, amplification failure (allele and locus dropout) [3], preferential amplification of shorter amplicons [4] and artefacts (enzymatic stutter, allele drop-in and off-ladder peaks). Complete amplification failure can be due to PCR inhibition or the fragmentation of all DNA templates below target amplicon sizes, which generally range from 100 to 400 base pairs (bp) [5]. Another complication – ‘jumping PCR’ – can generate non-authentic chimeric amplicons from discrete DNA template molecules, particularly when DNA fragmentation levels are high [6,7]. Additionally, chemical DNA modification due to miscoding lesions can terminate amplification reactions by halting DNA polymerase extension [7]. A combination of all these factors can lead to a poor or misleading DNA profile, or no profile at all in extreme cases.

The development and optimisation of nuclear SNP (single nucleotide polymorphism) typing protocols, shorter amplicon commercial STR kits (mini-STRs) [8], optimisation of PCR conditions, capillary electrophoresis and statistical interpretation techniques [9] have improved standard profiling methods [10-12]. However, in spite of these developments, the STR profiling of degraded, low-template DNA often has limited success. Furthermore, a large number of nuclear SNPs are required (50 to 80 loci) to obtain a similar level of discrimination as a full nuclear 16-loci STR profile [13]. In these cases, genetic identification from degraded samples may succeed through the analysis of mitochondrial DNA (mtDNA).

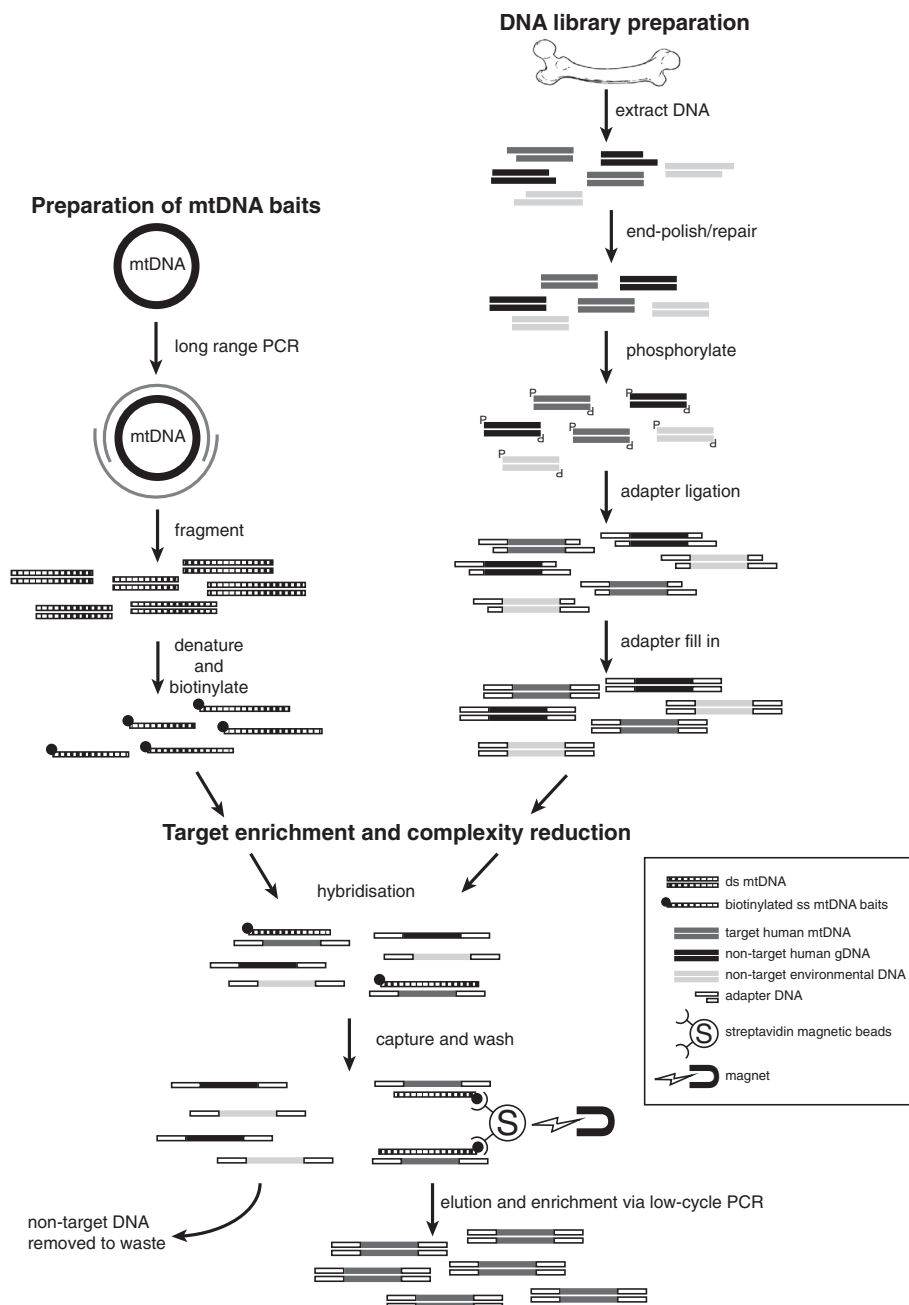
Mitochondrial DNA has several features that can make it a useful marker for human identification. As there can be thousands of copies of the mitochondrial genome in many cells (compared to only two copies for autosomal nuclear DNA), mtDNA typing is well suited to biological specimens where DNA fragmentation has occurred or the total DNA copy numbers are naturally low or have been severely reduced through post-mortem damage and degradation [14]. Suitable materials include bones, teeth, hair shafts, faeces and other biological materials. The lack of recombination events in the mtDNA genome and strict uniparental inheritance, in contrast to the nuclear genome, can allow maternal relatives separated by several generations to serve as reference samples. This latter feature is particularly beneficial in missing-person identification, where suitable ante-mortem or family reference samples may be unavailable.

Standard PCR-based sequencing approaches for mitochondrial hypervariable regions I and II (HVS I and II) typically amplify 2 to 12 overlapping fragments of approximately 150 bp to 600 bp in length [15-18] but are labour intensive, consume significant amounts of valuable DNA extract and can be template-length dependent and costly. Repeated singleplex PCR amplifications also bring an increased risk of contamination with exogenous human DNA due to the multiple lab steps required. Multiplex PCR amplification could in theory provide a solution for medium-sized PCR target fragments but still require hundreds of overlapping amplicons [1] in cases where whole mitochondrial genome sequences are needed for high-resolution identification.

Another disadvantage of typing just the mitochondrial HVS I/II is that short sequences from this single locus are far less powerful for identification purposes than a full multi-locus STR profile [19]. This can become a significant problem when many individuals in a population share a common haplogroup, such as the >40% of Western Europeans who belong to mitochondrial haplogroup H, or when distantly related individuals share a maternal ancestry that may not be known [20]. Recent studies sequencing whole mitochondrial genomes have shown that >70% of the mtDNA variation can be located outside HVS I/II for some haplogroups [21], so that full mitochondrial genome sequences provide far greater resolving power for human identification [22,23].

Ancient DNA studies of human archaeological samples routinely generate complete mitochondrial genomes via DNA hybridisation-based enrichment of mtDNA target sequences [21,24-26], and the creation of barcoded/indexed DNA libraries, followed by next-generation sequencing (NGS). Multiple samples can be processed in parallel in a high-throughput fashion [25], greatly reducing processing contamination risks, labour and costs compared to traditional Sanger sequencing approaches. These kinds of DNA capture strategies generally rely on the hybridisation of target DNA sequences to probes that are either immobilised on a surface (such as a microarray) or in solution [27,28]. Despite the significant potential of these new approaches, they have not been applied or examined in a forensic context for human identification.

The aim of this study was to sequence whole mitochondrial genomes from a range of human skeletal samples (in this case ranging from 10 to 2,500 years old) at an affordable cost using standard laboratory equipment and home-made DNA-capture probes for use in hybridisation-based target enrichment (Figure 1). Our previous application of this method [21] used three rounds of in-solution capture-based enrichment so we also aimed to explore the efficacy of using one or two rounds of enrichment to reduce costs and improve workflows. We deliberately used samples that had previously failed or had the potential to



**Figure 1 mtDNA hybridisation enrichment protocol.** mtDNA baits were prepared using  $2 \times 8$  kb mtDNA long-range PCR products (spanning 16,569 base pairs), generated using a DNA extract from a present-day sample (of a known haplotype) and the Roche LR Expand PCR kit. The products were fragmented by physical shearing to create 200 bp to 600 bp fragments prior to end labelling with biotin. The DNA library was prepared as follows. Damaged DNA leaves 5' and 3' overhangs. T4 DNA polymerase was used to polish the DNA by creating blunt ends and T4 PNK phosphorylated 5' ends, which is required for adaptor ligation. T4 ligase attached universal hybridisation adaptors (Uni-hyb A and Uni-hyb B) to the phosphorylated ends. Klenow polymerase filled in the short-arm adaptor ligation to create double-stranded adaptors (through the use of deoxyribonucleotide triphosphates - dNTPs). Adaptor complementary primers and Taq polymerase amplified the entire library to immortalise the sample. Single-stranded probe DNA was mixed with single-stranded library DNA and left to hybridise overnight (in the presence of blocking oligos). Biotinylated probe and bound library DNA were fixed to streptavidin beads on a magnetic rack, and non-specific or weakly bound library DNA was washed away through a series of three stringency washes (by increasing temperature and decreasing salt concentration) from the library-probe-streptavidin interaction. The single-stranded library DNA was converted to double stranded DNA and eluted from the probe-streptavidin interaction using the Bst strand-displacing enzyme (in the presence of dNTPs). Bst recognises nicks in the template and displaces library DNA into solution. Probe DNA remained bound to the magnet. Eluted library DNA was enriched through low cycle PCR, using adaptor complementary primers. Library DNA was then prepared for next-generation sequencing. mtDNA, mitochondrial DNA; PCR, polymerase chain reaction.

fail nuclear STR typing (Table 1). STR profiling was predominantly performed to assess the likelihood of obtaining full STR profiles from degraded samples and not to identify the samples. To identify the samples would require reference profiles for comparison and replicate testing of the samples by LCN analyses. The capture-probe method is designed to focus on the recovery of human mtDNA fragments <100 bp in length (with the vast majority in the 20 to 70 bp range) (Figure 2), from samples that yield highly damaged and fragmented DNA templates available only with low copy number. We anticipate the method will be useful for samples that cannot be typed successfully using standard STR kits and for detecting key or private SNPs within whole mitochondrial genome sequences (that would otherwise remain undetected with traditional mtDNA HVS I/II sequencing) for human identification.

## Methods

### Samples

Bone and tooth samples were selected representing a range of ages, preservation conditions and contexts (Table 1). Three samples were from missing-person cases, two of which were from Australian servicemen killed in World War II. Two samples were recovered from archaeological contexts.

### Degraded DNA work

To avoid the potential for contamination of samples with contemporary human DNA or previously amplified PCR products, all steps preceding DNA library amplification were carried out in a dedicated ancient DNA laboratory geographically separated (by approximately 1.5 km) from post-PCR and other molecular biology laboratories at the University of Adelaide.

Strict decontamination procedures were followed [29]. There were ultraviolet lights in every room. There was positive air pressure and the one-way airflow was filtered using high efficiency particulate air filters. There were

separate workrooms each containing dead-air glove boxes. Equipment and work areas were cleaned with sodium hypochlorite and isopropanol before and after use. Personnel protective clothing included a full-body suit, face mask, face shield, boots and triple-gloves. There was a strict one-way movement of personnel (from shower to freshly laundered clothes to ancient DNA laboratory to post-PCR laboratory).

Non-template controls and extraction blanks were included in each experiment to monitor potential contamination from exogenous human DNA sources and cross-contamination from other samples. The complete mitochondrial genomes of all staff involved in the handling of the samples (JT, WH, BL and PB) were sequenced to monitor potential contamination (Additional file 1: Table S1). The mitochondrial genome of an anonymised present-day sample used to generate mtDNA capture probes was also fully sequenced (haplotype: J1c8a) to monitor contamination (Additional file 1: Table S1).

### DNA extraction, quantification and STR profiling

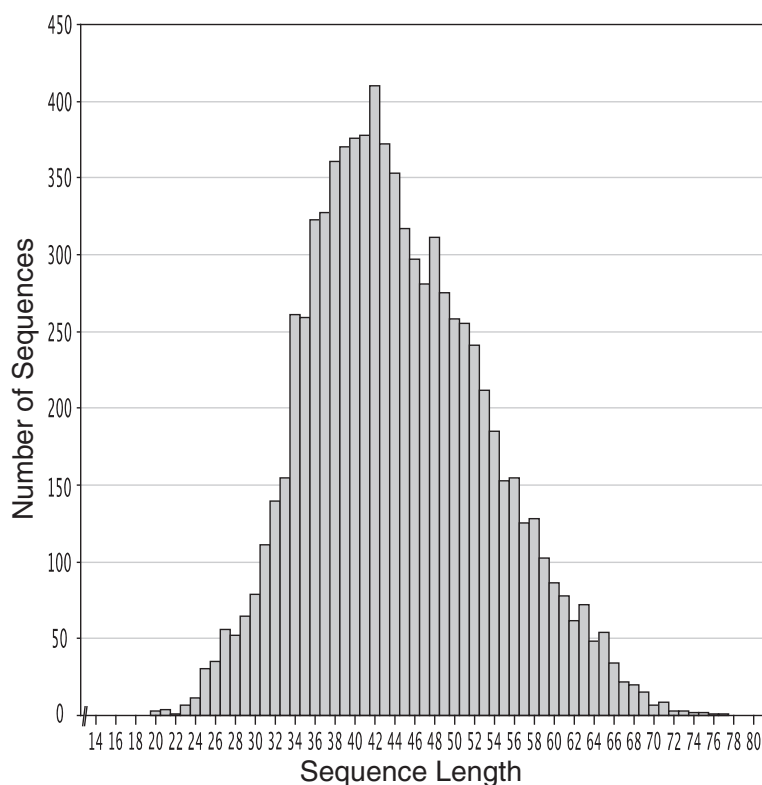
DNA was extracted from five samples using an optimised method as previously described [21]; see Additional file 1.

For sample 9210A, a small quantity of DNA extract was available so only mitochondrial testing was performed. The four additional samples (4464B, 10730A, 8727C and 11995A) were subjected to both nuclear STR and mitochondrial sequence analyses. Nuclear DNA and mtDNA were quantified in all extracts using quantitative PCR (qPCR) with SYBR<sup>®</sup> green chemistry and previously published 67 bp nuclear [30] and 77 bp mitochondrial [31] PCR targets (Additional file 1: Table S5). The total 10 µL qPCR reaction mix consisted of 1× Brilliant II SYBR<sup>®</sup> green master mix (Agilent Technologies, USA), 0.15 µM forward primer (Additional file 1: Table S5), 0.15 µM reverse primer (Additional file 1: Table S5), 400 ng/µL rabbit serum albumin and 1 µL DNA extract. Samples were run in triplicate, and negative (no template) and positive controls (male genomic control DNA, Applied

**Table 1 Samples used in this study**

ACAD ID	Description	Age	Excavated	Locality	Environmental conditions of site	Preservation status on collection
8727C	Cranium fragment	Approximately 10 years	Unknown	South-east Queensland, Australia	Surface deposit	Well preserved
11995A	Long bone fragment	Approximately 70 years	2011	Papua New Guinea	Tropical lowland battlefield, wet	Poor
9210A	Tooth	Approximately 70 years	2006	Christmas Island, Australia	Tropical lowland, burial, wet	Well preserved
4464B	Long bone fragment	Approximately 2500 years (Iron Age)	2007	Latsch, South Tyrol, Italy	Temperate, burial	Well preserved
10730A	Long bone fragment	Approximately 600 to 1,000 years (Ychsma culture)	2002	Huaca Pucllana, Lima, Peru	Temperate, high altitude, burial	Well preserved

ACAD, Australian Centre for Ancient DNA.



**Figure 2 Distribution of the length of human mtDNA reads from sample ID 8727C.** In total, 277,470 reads were trimmed to remove primers and the resulting 218,102 mapped (to the Reconstructed Sapiens Reference Sequence) reads were filtered to remove duplicates as a result of clonal reads. The resulting 7,756 unique mapped reads (2.8% of the total reads) gave 99.5% mitochondrial genome sequence (16,487 base pairs) coverage. The mean fold coverage was 20x (SD  $\pm$  9.7), mean unique read length was 43.2 bp (SD  $\pm$  9.3) and all fragments captured were <80 bp.

Biosystems, USA) were included in all runs. Extraction blank controls were also quantified. Cycling was performed using a Corbett 6000 Rotor-Gene real-time PCR thermocycler and consisted of an initial 5 min denaturation step at 95°C, followed by 45 cycles of 95°C for 10 s, 59°C for 20 s and 72°C for 15 s. Results were analysed using the Rotor-Gene 6000 Series Software 1.7. The DNA concentration was determined using the comparative cycle threshold method where unknown samples are compared to a standard curve. The standard curve for the nuclear target was created using male genomic control DNA (Applied Biosystems, USA). The standard curve for mitochondrial DNA was created using a PCR product (Additional file 1: Table S5).

STR typing was performed using AmpFLSTR ProfilerPlus™ (Applied Biosystems, USA). The final 12.5  $\mu$ L reaction volume consisted of 4.6  $\mu$ L ProfilerPlus™ reaction mix, 2.5  $\mu$ L of ProfilerPlus™ primer mix, 0.4  $\mu$ L AmpliTaq Gold™ and 5  $\mu$ L of DNA extract. Cycling was performed on a 9700 GeneAmp thermal cycler and consisted of an initial denaturation at 95°C for 10 min followed by 34 cycles of 94°C for 1 min, 59°C for 1 min and 72°C for 1 min, then a final extension at 60°C for 45 min. PCR products were analysed on a 3130xl Genetic Analyser in a 17.3  $\mu$ L final

volume that consisted of 2  $\mu$ L of PCR product, 15  $\mu$ L HiDi™ formamide and 0.3  $\mu$ L ROX-500 size standard (Applied Biosystems, USA). Results were analysed using Genemapper ID software (v3.2.1). Alleles were interpreted based on peak heights reaching a set threshold value of 25 relative fluorescence units (RFU) above a clean baseline. A wild-card designation was used, with peak heights <150 RFU, to account for potential allele dropout (for example, '11, F' instead of '11,11'). A profile was considered full when all alleles were detected above the threshold RFU. A profile was defined as partial when peaks were detected above the threshold RFU and when at least one locus was successfully called. A profile in forensic terms can be described as partial when at least one locus has been called (even if this is not an informative profile).

#### Mitochondrial DNA capture and enrichment

Biotinylated DNA-capture probes of a known haplotype and DNA libraries were generated as described in Additional file 1 and as previously described [21]. Whilst Brotherton *et al.* [21] used three rounds of enrichment for all samples, we explored the effects of using one or two rounds of enrichment on the number of unique reads, coverage and sequencing depth. Hybridisation was carried out in a final

volume of 30  $\mu$ L consisting of 100 ng of probe and 400 ng of library DNA. The thermal profile used was: denature DNA for 5 min at 95°C, followed by 14 to 18 hours incubation at 50°C to allow the DNA-capture probes to hybridise to fragments of DNA with closely matched sequences from complementary human DNA regions. The two library primers (Additional file 1: Table S4) were included as part of the hybridisation mix at 0.67  $\mu$ M to 1.0  $\mu$ M, to act as 'blocking' oligonucleotides. The blocking oligonucleotides are complementary to the library adaptors and have a dual role during the hybridisation reaction: (i) to minimise unwanted hybridisation between the adaptor-tagged flanking regions of otherwise unrelated single-stranded library DNA molecules and (ii) to enable strand displacement of probe DNA from library DNA as explained below.

Following overnight hybridisation at 50°C, 50  $\mu$ L of magnetic streptavidin beads in solution (Invitrogen) were added to 30  $\mu$ L of hybrid DNA and the beads were immobilised on a magnetic rack. The clear supernatant was discarded. The bead complex (DNA-capture probe/library DNA) immobilised to the magnet was subjected to successively increased-stringency washes, to remove progressively non- or weakly-hybridised single-stranded library DNA molecules, using decreased salt and increased temperature: 2 $\times$  saline sodium citrate (SSC)/0.1% sodium dodecyl sulphate (SDS) at 37°C for 1 min; 2 $\times$  SSC/0.1% SDS at 42°C for 10 min; 1 $\times$  SSC/0.1% SDS at 43°C for 10 min; 0.5 $\times$  SSC/0.1% SDS at 44°C for 10 min; 0.5 $\times$  SSC/0.1% SDS at 45°C for 10 min.

The strand-displacing Bst DNA polymerase enzyme (large fragment, New England Biolabs) was used to release library DNA from the DNA-capture probe (immobilised to beads on the magnet). Reactions were performed at 35  $\mu$ L final volume comprising 1 $\times$  Thermopol buffer (New England Biolabs), 200  $\mu$ M of each dNTP (to convert single-stranded library DNA to dsDNA), and 100  $\mu$ g/mL of Bovine Serum Albumin (New England Biolabs). The reaction was pre-heated to 60°C and 2  $\mu$ L of Bst enzyme was added last to each reaction. Tubes were incubated at 60°C for 5 min with regular agitation. The reaction tube was then applied to the magnetic rack at 60°C and 35  $\mu$ L of supernatant was transferred to a fresh PCR tube. This tube was immediately incubated at 80°C for 20 min to inactivate the enzyme.

The heat-inactivated supernatant was split between eight PCR re-amplification reactions (total combined volume 140  $\mu$ L), designed so that upon the addition of the sub-portion of Bst buffer, the final composition of the reactions would be 1 $\times$  AmpliTaq Gold buffer II, 2.5 mM MgCl<sub>2</sub>, 250  $\mu$ M of each dNTP, 1.0 U AmpliTaq Gold (Applied Biosystems), and 0.5  $\mu$ M of PCR primers UniHyb-PCR-A and UniHyb-PCR-B (Additional file 1: Table S4). Thermocycling was at 94°C for 11 min, followed by 12 cycles of

30 s at 95°C, 30 s at 60°C and 1 min (+2 seconds per cycle) at 72°C, followed by a final 10 min at 72°C. Amplification reactions were pooled and library amplicons purified using MinElute spin columns (Qiagen) and eluted into 15  $\mu$ L as per the manufacturer's instructions. These comprised the 'first enrichment' DNA libraries and amplification products were sized and quantified via gel electrophoresis against size markers (HyperLadder™ V, Bionline) and a Nanodrop 2000 (Thermo Scientific).

For cases where a second round of enrichment took place, the overnight hybridisation and wash steps were repeated to produce 'second enrichment' DNA libraries highly enriched for mtDNA sequences.

#### Ion Torrent PGM sequencing

Enriched library DNA was prepared for Ion Torrent sequencing by re-amplification using Ion Torrent barcoded primers (Additional file 1: Table S6). Eight 24  $\mu$ L reaction volumes per sample were re-amplified using 1  $\mu$ L of purified library DNA as the template. Final reactions conditions comprised of 1 $\times$  AmpliTaq Gold buffer II, 2.5 mM MgCl<sub>2</sub>, 2.5 U AmpliTaq Gold (Applied Biosystems), 250  $\mu$ M of each dNTP (Invitrogen), and 0.5  $\mu$ M of each PCR primer. The thermocycling profile was 94°C for 12 min, followed by 12 cycles of 30 s at 95°C, 30 s at 60°C and 45 s at 72°C, followed by a final 10 min at 72°C. The eight amplified samples per reaction were pooled and purified using MinElute spin columns (Qiagen) and eluted into 15  $\mu$ L as per the manufacturer's instructions. The DNA was sized and quantified via gel electrophoresis against size markers (HyperLadder™ V, Bionline) and a Nanodrop 2000 (Thermo Scientific). Library DNA was size-selected above 120 bp and further purified to remove adaptor dimer, using Qiagen's gel extraction purification kit following the manufacturer's instructions.

Prior to sequencing, the fragment size distribution and DNA concentration of individual libraries were measured using a Bioanalyzer 2100 (Agilent Technologies) following the manufacturer's instructions. The quantified indexed library DNA was pooled to an equimolar concentration prior to the One Touch. The pooled library DNA was adjusted to a final concentration of 10 to 15 pM prior to amplification (by emulsion PCR) and enriched for positive ion sphere particles (ISPs) using the Ion Torrent One Touch System II (Life Technologies) and the Ion One Touch 200 template kit v2 DL (Life Technologies), following the manufacturer's instructions.

Templated ISPs were sequenced on a 316 micro-chip (up to 100 Mb of data expected) using the Ion Torrent Personal Genome Machine (PGM; Life Technologies) and the Ion PGM 200 sequencing kit v2 chemistry (Life Technologies) for 130 cycles (520 flows). After sequencing, the individual sequence reads were filtered within the PGM software to remove low-quality and polyclonal sequences.

Sequences matching the PGM 3' adaptor were also automatically trimmed prior to bioinformatics analysis.

### Bioinformatic sequence analysis

Ion Torrent PGM data from the mtDNA capture was processed using a customisable analytical pipeline. The scripts `fastx_barcode_splitter.pl` and `fastx_trimmer` (from the FASTX toolkit [32]) were used to demultiplex the reads by barcode, using a strict zero mismatch threshold. `Cutadapt v1.1` [33] was then used to trim adapters using a maximum error rate of 0.33 ( $-e 0.3333$ ), and to remove short ( $-m 25$ ), long ( $-M 110$ ) and low-quality sequences ( $-q 20$ ), for a total of five passes ( $-n 5$ ). The filtered reads were checked with `FastQC` [34] before being mapped against the Reconstructed Sapiens Reference Sequence (RSRS) [35] using `TMAP v3.2.1` [36] with the following options: `-g 3 -M 3 -n 7 -v stage1 -stage-keep-all map1 -seed-length 12 -seed-max-diff 4 stage2 map2 -z-best 5 map3 -max-seed-hits 10`. The program TMAP has been optimised to align Ion Torrent PGM reads against a reference genome [37]. Mapped reads with mapping quality below Phred 30 and read duplicates were removed using `Samtools v0.1.18` [38] and the `MarkDuplicates` tool of `Picard Tools v1.79` [39]. The GC content of mapped reads was analysed using the `CollectGcBiasMetrics` tool of `Picard Tools v1.79`. Misincorporation patterns were assessed using `mapDamage v0.3.6` [40]. The resulting sequence assembly was visualised using `Biomatters Geneious Pro v5.6.2` software [41] and mitochondrial haplotypes were defined for each individual according to `phyloree.org` [4].

### Confirming SNP calls by hypervariable region I sequencing

HVS I was amplified using a minimum of four short overlapping primer pairs, as previously described [42,43]. Minisequencing of 22 coding region SNPs (`GenoCoRe22`) using a multiplex and `SNaPshot` based approach was carried out, as previously described [42,43].

### Results

Quantitative PCR on four of the five samples with sufficient DNA extract volume indicate a 14,000 to 300,000-fold difference in the amount of recovered nuclear DNA:mtDNA, highlighting the greater potential for mtDNA typing in degraded remains. The total nuclear DNA in all four samples was very low ( $<2$  pg/ $\mu$ L). Subsequently, all four samples only produced partial STR profiles using low copy number techniques (34 cycles of PCR and reduced reaction volumes with higher concentrations of *Taq* DNA polymerase) (Table 2). Locus dropout was observed in each degraded sample analysed for nuclear STR typing. Only the positive control DNA produced a full STR-DNA profile, which matched the reference profile at all ten loci examined (Table 2). All negative controls were blank.

In contrast, a higher concentration of mitochondrial DNA was detected in all four samples using qPCR (Table 2). DNA library preparation, mtDNA enrichment and NGS were completed for all five samples. After one round of hybridisation and enrichment we obtained 96% to 97% of the mitochondrial genome at an average 15 to 18-fold coverage from two well-preserved samples but only 62% of the mitochondrial genome at an average 1-fold coverage on a poorly preserved sample (Table 3, Figure 3). However, after two rounds of hybridisation and enrichment we obtained 98% to 100% of the mitochondrial genome at an average 1646-fold coverage from all five samples, irrespective of morphological preservation of the sample (Table 3, Figure 3). Complete or near complete mitogenomes were recovered from samples with as few as 350 copies/ $\mu$ L of the 77 bp mtDNA fragment.

Two rounds of enrichment substantially improved the number of unique reads that mapped to the mtDNA genome (from 2- to 11-fold) (Figure 4, Additional file 1: Figure S1) and the average redundancy per site of the genome (from 1 to 18 $\times$  to 16 to 46 $\times$ ) but did not alter the mean fragment length of mtDNA recovered (42 bp after one round and 45 bp after two rounds) (Table 3). The second round of enrichment proved to be particularly important for the less well-preserved sample 11995A, for which it provided an 11-fold increase in the total number of unique reads, which also substantially improved the coverage from 62% to a near complete mitochondrial genome (98%) and therefore allowed an unambiguous haplotype designation (Table 3).

Coverage was uneven across all five samples after one and two rounds of enrichment (Additional file 1: Figure S1). This variation in coverage has been reported previously for modern and ancient human and Neanderthal mtDNA genomes [5,24,44-46] and is positively correlated with GC content. This may be due to loss (denaturation) of short AT-rich sequences before or during the library preparation [5,46].

Damage patterns in all samples followed expectations for degraded DNA, with a larger than usual amount of deaminated cytosine residues accumulated towards the ends of the molecules. In addition, we could observe a high frequency of indels, a well-identified homopolymer sequencing error characteristic of PGM technology [47-50]. However, indels were randomly distributed and did not affect the final consensus sequences, as each called position was covered with enough depth to prevent false-positive base calls.

Stringent quality filtering during analysis removed a large proportion of the total reads for each sample. Post-filtering provided on average, across all samples, a very small proportion of unique mapped reads vs total reads (0.04% to 2.4%) (Figure 5). However, the pattern of mapped reads had an adequate level of coverage for each sample to allow



**Table 2 Mitochondrial and nuclear quantitative PCR (qPCR) and STR typing result**

Sample ID	Mitochondrial DNA			Nuclear DNA		
	qPCR	qPCR	qPCR	STR typing (Profiler Plus kit, 10 loci) 34 cycles		
	77 bp target (copies/ $\mu$ L)	67 bp target (copies/ $\mu$ L)	67 bp target (ng/ $\mu$ L)	Number of loci	Percentage genotyping success	None, partial or full profile obtained
4464B	15,727	<1 copy (0.54)	0.001794	8	80%	Partial
10730A	62,592	0	0	Amel only	10%	Partial
8727C	350	<1 copy (0.025)	0.000082	3	30%	Partial
11995A	2,715	<1 copy (0.0091)	0.00003	2	20%	Partial
9210A	n/a	n/a	n/a	n/a	n/a	n/a
Negative control (H2O)	0	<1 copy (0.034)	0.000111	0	0%	None
Female positive control	7,896,895	543	1.79213	10	100%	Full

n/a, non-applicable (The sample was excluded from testing due to the low quantity of DNA extract. Quantitative PCR data is presented as a geometric average using triplicate values.).

detection of variants in the mitochondrial genome. Traditional HVS I sequencing and coding region SNP mini-sequencing of all five samples produced identical results to those obtained by whole mtDNA genome sequencing.

### Discussion

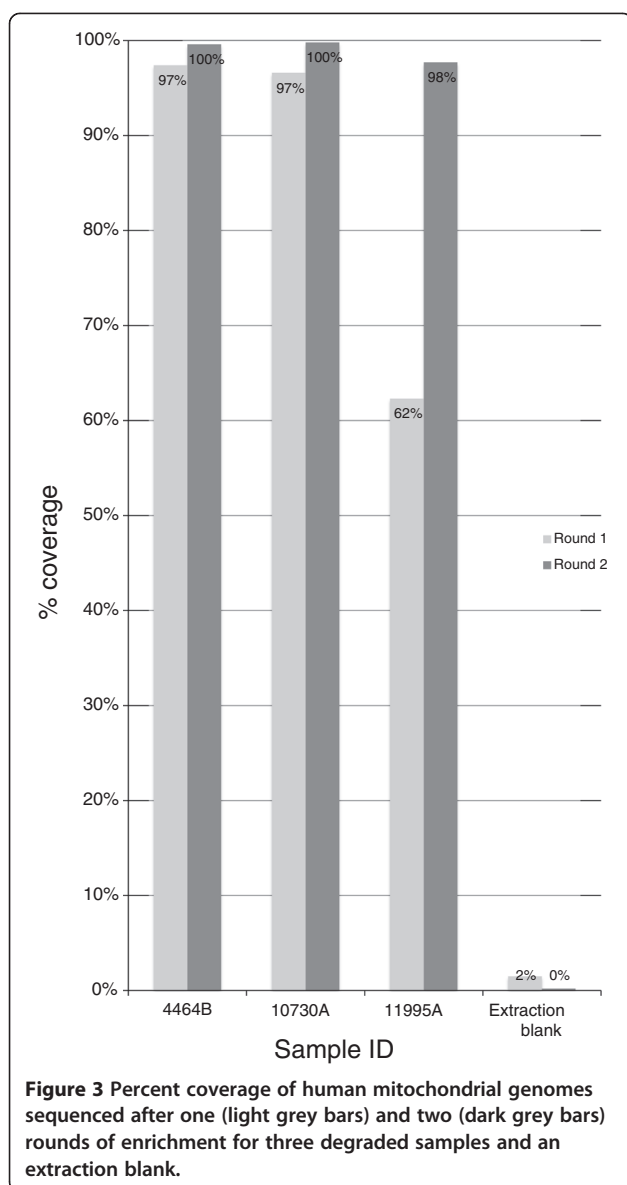
Low amounts of DNA combined with high levels of damage and fragmentation make STR typing of degraded samples challenging. DNA capture coupled with next-generation sequencing can retrieve whole mitochondrial genome sequences from degraded samples when nuclear DNA is below detection levels. Despite high levels of DNA decay in skeletal remains, whole mtDNA genome sequencing is possible due to the copy-number advantage and reduced rate of fragmentation of mtDNA (compared with nuclear DNA) combined with the ability to capture and sequence DNA fragments in the 20 bp to 70 bp range. Quantitative PCR can be used to determine the amount of DNA available from extracted materials and will indicate the likelihood of obtaining a nuclear STR-DNA profile

from a degraded sample. This is of particular importance in cases where total nuclear DNA quantity is <100 pg, which reduces the likelihood of obtaining a full nuclear STR DNA profile even when applying LCN techniques. In contrast, near complete mitochondrial genome sequences can be obtained with a single round of enrichment from samples with >10,000 77 bp mtDNA copies/ $\mu$ L and with two rounds of enrichment from samples with <3,000 77 bp mtDNA copies/ $\mu$ L. Our work builds on previous in-solution capture-based enrichment methodologies [21,24-26] and demonstrates the importance of using multiple rounds of enrichment to improve mtDNA genome recovery from samples with low amounts of endogenous DNA. Repeating the enrichment process on samples with very low amounts of mtDNA can more than double the number of unique reads and average coverage, and substantially improve the overall coverage of the mtDNA genome (Additional file 1: Figure S1). The methodology has the ability to capture DNA templates that are damaged and fragmented (<100 bp in length) (Figure 2) and that are generally difficult to recover using traditional methods

**Table 3 Ion Torrent PGM whole mitochondrial DNA sequencing data, after one and two rounds of enrichment**

ACAD ID	Rounds capture	Percentage coverage	Average length ( $\pm$ SD)	Number of bases covered	Unique reads	Mean X coverage ( $\pm$ SD)	Haplotype
4464B	1	97.4%	42 ( $\pm$ 11.0)	16,136	7,069	18 ( $\pm$ 12.3)	HV0e
4464B (ancient)	2	99.60%	45 ( $\pm$ 11.7)	16,507	14,837	40 ( $\pm$ 19.8)	HV0e
10730A	1	96.6%	43 ( $\pm$ 11.3)	16,004	5,760	15 ( $\pm$ 10.7)	B2b
10730A (ancient)	2	99.80%	47 ( $\pm$ 12.2)	16,540	16,138	46 ( $\pm$ 20.2)	B2b
11995A	1	62.3%	40 ( $\pm$ 10.2)	10,321	553	1 ( $\pm$ 1.6)	Low coverage
11995A (forensic)	2	97.70%	45 ( $\pm$ 12.9)	16,188	6,012	16 ( $\pm$ 11.2)	H1a
Extraction blank	1	1.5%	41 ( $\pm$ 18.4)	248	6	0	n/a
Extraction blank (EBC11049)	2	0.20%	0	28	1	0	n/a
8727C (forensic)	2	99.50%	43 ( $\pm$ 9.3)	16,487	7,756	20 ( $\pm$ 9.7)	U5a2a1fNEW
9210A (forensic)	2	100%	65 ( $\pm$ 20.4)	16,569	7,865	31 ( $\pm$ 14.1)	J1c12

ACAD, Australian Centre for Ancient DNA; n/a, non-applicable.



of PCR-based amplification and sequencing [51]. This is of particular importance in cases where DNA has been exposed to prolonged heat, moisture, ultraviolet light and microbial attack, which generally results in template fragmentation (in extreme cases there can be no surviving endogenous DNA templates >100 bp) [52].

Two common concerns with mitochondrial DNA testing can be eliminated or reduced using this whole mtDNA genome sequencing approach. Traditional HVS I/II sequencing requires 2 to 12 separate PCR amplifications and up to 24 separate DNA sequencing reactions. This multi-tube, multi-step approach introduces the potential for sample mix-up during laboratory processing and increases the risk of introducing contaminating DNA. Our whole mtDNA genome approach eliminates this risk, massively reducing opportunities for sample mix-up, while the barcoded

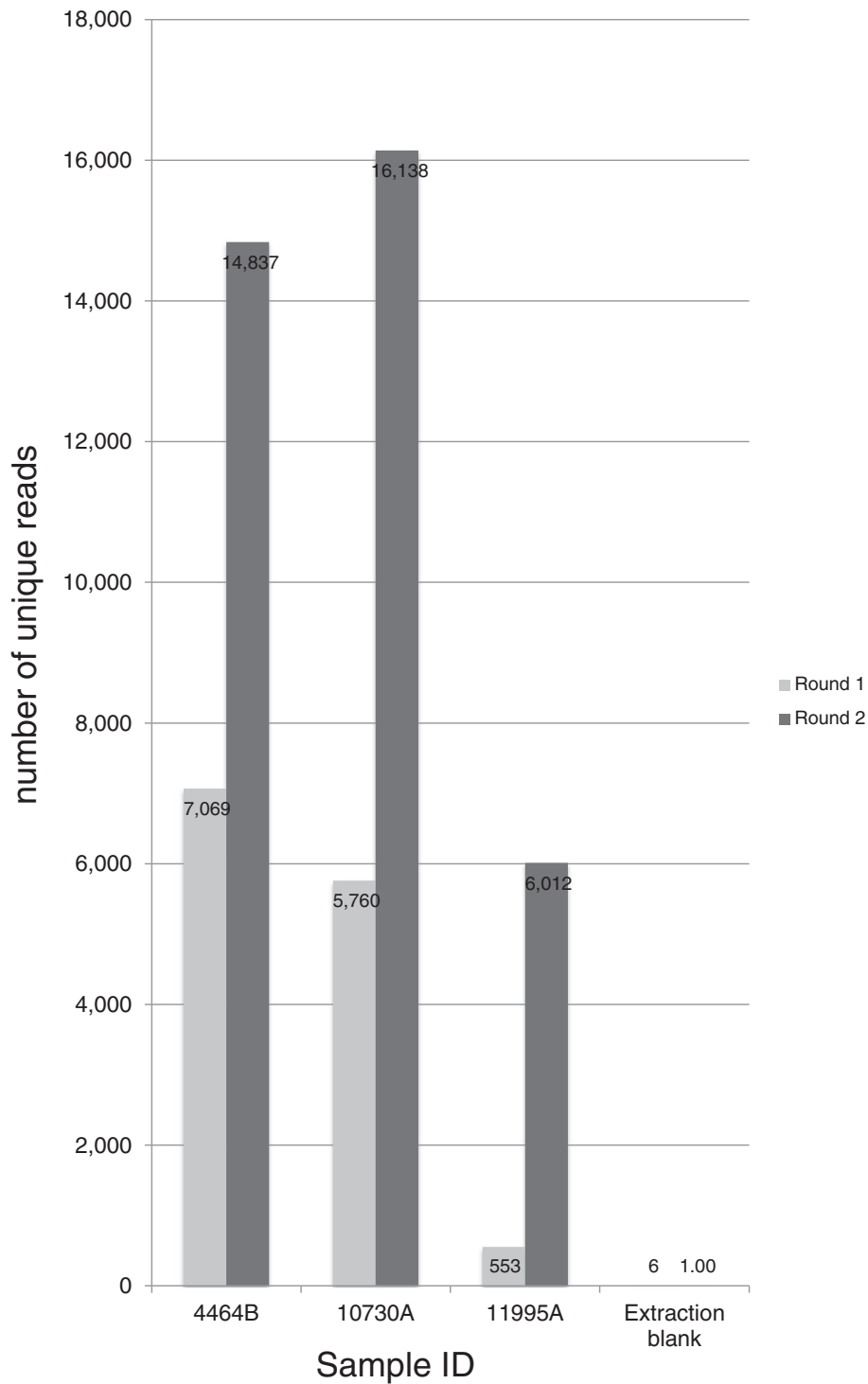
adapters ligated to the DNA provide an additional means to eliminate (or identify and screen out) contamination introduced in later steps. In addition, barcoding allows many samples to be pooled for high-throughput screening efforts and can reduce the cost of sequencing.

DNA capture and related approaches have been shown to give preferential enrichment of short endogenous DNA templates over longer exogenous contaminant DNA in a sample [7]. This is particularly important where small quantities of endogenous DNA in a sample have become saturated by larger quantities of exogenous contamination (human and microbial), consequently leading to poor PCR amplification, mistyping of target loci via artefacts or even complete PCR amplification failure [1].

Traditional forensic and archaeogenetic studies using mtDNA have relied on HVS I/II sequencing. However, this relatively short sequence has limited resolving power and can fail to discriminate between distinct maternal lineages [20]. Outside the control region, coding region SNPs provide additional resolution and discriminatory power [20,53]. To date, this additional information has been obtained via case [20,54], region [55,56], continental [57] or haplogroup [58] specific SNP multiplex assays. In contrast, our whole mtDNA genome sequencing approach is a universal solution for obtaining high-resolution mtDNA data, which can discriminate between closely related maternal lineages. However, although our methodology provides a mechanism to generate whole mtDNA genome sequences from difficult and degraded samples, there is a clear need for the parallel development of high-quality mitochondrial genome databases [20,53].

Complete mitochondrial genomes sequences can aid human identification efforts by placing an individual into specific haplotypes based on private SNPs. This high-resolution discrimination can be used to include or exclude closely related maternal lineages [21], especially in populations with high frequencies of particular haplotypes. By resorting to whole mtDNA sequencing, we were able to gain additional haplogroup and haplotype resolution relative to traditional HVS I/II sequencing. This information has already proved critical in a comparison with maternal relatives in a case where the HVS I/II sequence alone could not exclude a maternal relationship. Our approach could assist large-scale identification efforts when more comprehensive mtDNA reference databases become available to the forensic community.

Validation studies have confirmed that mtDNA typing is a reliable means of forensic identification [59]. However, a worldwide effort will be required with labs collaborating and producing large databases, estimating the frequency of particular mtDNA haplotypes and improving the statistical basis of the databases. In the meantime, techniques used to sequence whole mtDNA in archaeological and population studies will continue to advance at a rapid pace.

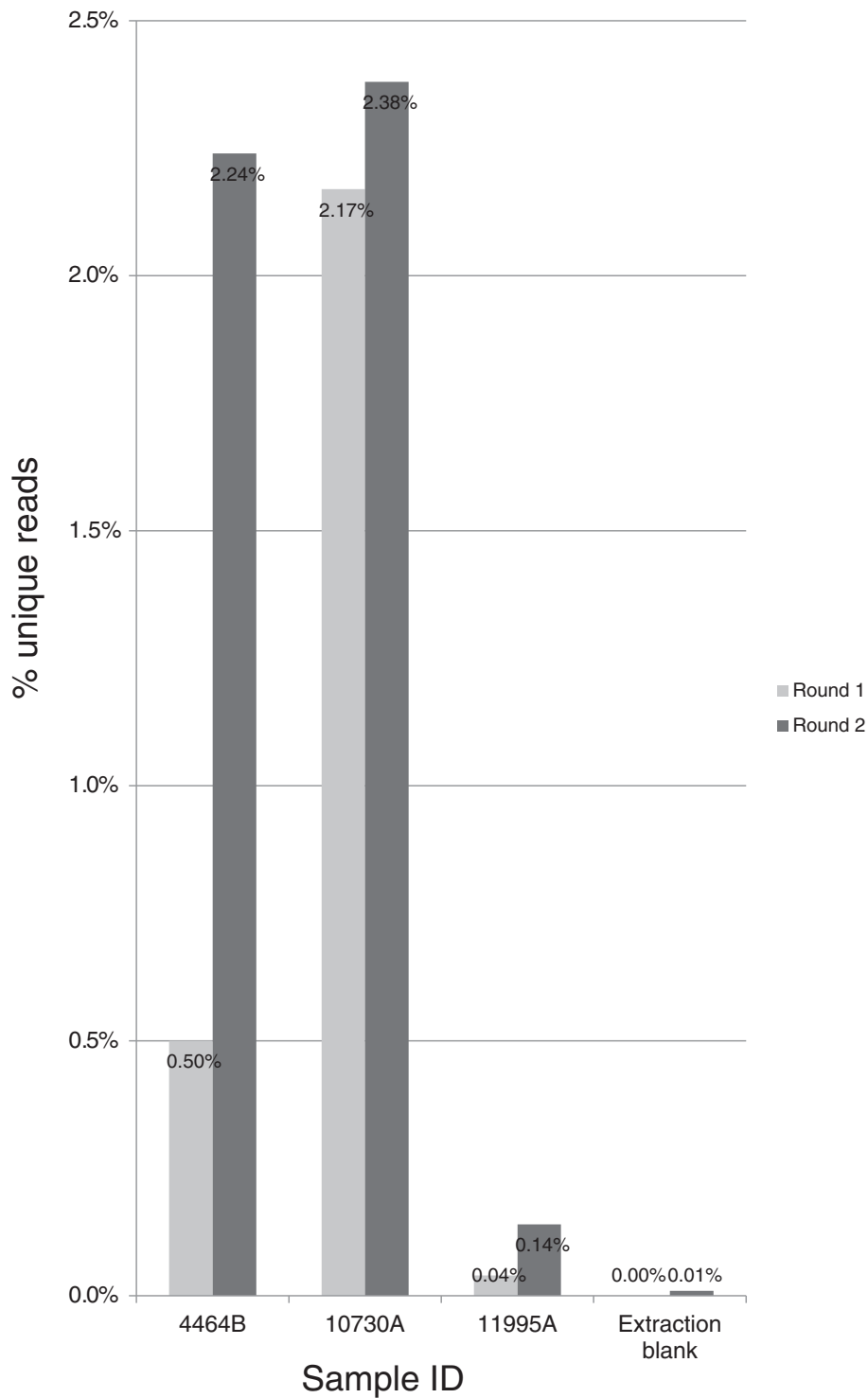


**Figure 4** Number of unique sequencing reads mapped to the Reconstructed Sapiens Reference Sequence mitochondrial genome after one (light grey bars) and two (dark grey bars) rounds of enrichment for three degraded samples and an extraction blank.

### Conclusions

In-solution capture-based whole mitochondrial genome sequencing immortalises the limited and important contents of the DNA extract in the form of a DNA library, and

is followed by targeted enrichment of mtDNA sequences. The application of these methods using hybridisation enrichment and NGS has led to the reliable genotyping of human remains for which standard nuclear PCR



**Figure 5** Percentage of unique sequencing reads that map to the Reconstructed Sapiens Reference Sequence mitochondrial genome after one (light grey bars) and two (dark grey bars) rounds of enrichment for three degraded samples and an extraction blank.

protocols had been unsuccessful. This result indicates that the technique can be applied to obtain whole mitochondrial genomes even from particularly challenging samples. Additionally, as NGS platforms become more

affordable and widely available and with the advent of DNA library, barcoding (to monitor contamination and allow multiple samples to be processed), new methods for mtDNA analysis should be considered.

## Additional file

**Additional file 1: Detailed description of the methods used to extract DNA from bone samples, prepare capture-bait library and prepare libraries from degraded DNA.** Primer sequences are shown in Tables S2, S3, S4, S5, and S6. Mitochondrial genome haplotypes for laboratory staff and degraded bone samples are shown in Tables S1 and S8, respectively. Details of Ion Torrent sample barcoding and sequencing runs are shown in Table S7. Mapping of individual sequence reads to the reference mitochondrial genome for all five degraded samples are shown in Figure S1.

### Abbreviations

ACAD: Australian Centre for Ancient DNA; bp: base pair; HVS: Hypervariable region; ISP: Ion sphere particle; kb: kilobase; LCN: Low copy number; mtDNA: mitochondrial DNA; NGS: Next-generation sequencing; PCR: Polymerase chain reaction; PGM: Personal Genome Machine; qPCR: quantitative PCR; RFU: Relative fluorescence units; SSC: Saline sodium citrate; SDS: Sodium dodecyl sulphate; SNP: Single nucleotide polymorphism; STR: Short tandem repeat.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

JELT co-developed the protocol, processed samples, performed next-generation sequencing, co-analysed data and wrote the manuscript. PB designed and developed the DNA extraction, library preparation and targeted enrichment protocol, co-developed the underlying research concept and assisted manuscript preparation. BL processed sample extraction and co-developed the analytical pipeline for data analysis with JS. BL and JS performed data analysis and assisted with manuscript preparation. WH contributed to the experimental design, provided archaeological samples, processed sample extractions and library preparations and assisted with manuscript preparation. AC co-developed the underlying concept, contributed to the experimental design and assisted with manuscript preparation. JA provided forensic samples, processed sample extractions, contributed to the experimental design and assisted with manuscript preparation. All authors read and approved the final manuscript.

### Acknowledgments

We thank the Queensland Police, Australian Navy and the Unrecovered War Casualties – Army (Australian Defence Force) for permission to use samples. Members of the Australian Centre for Ancient DNA, in particular Dr Denise Higgins, provided helpful and constructive comments on the research and manuscript. We thank Dr Hubert Steiner, Dr Isabel Flores Espinoza and Mrs Maria Ines Barreto for providing archaeological samples. Additional thanks to Professor Jeremy Timmis, Rosalie Kenyon, Dr Mark Van der Hoek, Dr Paul Gooding, Dr Steve Richards and Dr Hugh Cross for technical support. This research was funded by an Australian Research Council Linkage project to AC (LP0822622), an ARC Future Fellowship to JA (FT100100108) and an ARC Discovery Project to WH and BL (DP1095782).

### Author details

<sup>1</sup>Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia. <sup>2</sup>School of Biological Sciences, Flinders University, Bedford Park, Adelaide, South Australia 5001, Australia. <sup>3</sup>Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK. <sup>4</sup>Sciences Department, Museum Victoria, Carlton Gardens, Melbourne, Vic 3001, Australia.

Received: 19 July 2013 Accepted: 18 October 2013

Published: 2 December 2013

### References

1. Butler JM: *Advanced Topics in Forensic DNA Typing: Methodology*. London: Academic; 2011.

- Nandineni MR, Vedanayagam JP: Selective enrichment of human DNA from non-human DNAs for DNA typing of decomposed skeletal remains. *Forensic Sci Int-Gen Supplement Series* 2009, **2**:520–521.
- Alonso A, Martín P, Albarrán C, García P, Primorac D, García O, de Fernández Simón L, García-Hirschfeld J, Sancho M, Fernández-Piqueras J: Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies. *Croat Med J* 2003, **44**:273–280.
- van Oven M, Kayser M: Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* 2009, **30**:E386–E394.
- Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Laluzza-Fox C, Rudan P, Brajković D, Kučan Z, Gušić I, Schmitz R, Doronichev VB, Golovanova LV, Rasilla M, Fortea J, Rosas A, Pääbo S: Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science* 2009, **325**:318–321.
- Anderung C, Persson P, Bouwman A, Elburg R, Götherström A: Fishing for ancient DNA. *Forensic Sci Int-Gen* 2008, **2**:104–107.
- Brotherton P, Sanchez JJ, Cooper A, Endicott P: Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX. *Nucleic Acids Res.* 2010, **38**:e7.
- Parsons TJ, Huel R, Davoren J, Katzmarzyk C, Miloš A, Selmanović A, Smajlović L, Coble MD, Rizvić A: Application of novel 'mini-amplicon' STR multiplexes to high volume casework on degraded skeletal remains. *Forensic Sci Int-Gen* 2007, **1**:175–179.
- Bright J-A, Taylor D, Curran JM, Buckleton JS: Developing allelic and stutter peak height models for a continuous method of DNA interpretation. *Forensic Sci Int-Gen* 2013, **7**:296–304.
- Hill CR, Kline MC, Coble MD, Butler JM: Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J Forensic Sci* 2008, **53**:73–80.
- Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin K-J: Simple and highly effective DNA extraction methods from old skeletal remains using silica columns. *Forensic Sci Int-Gen* 2010, **4**:275–280.
- Fregeau CJ, De Moors A: Competition for DNA binding sites using Promega DNA IQ (TM) paramagnetic beads. *Forensic Sci Int-Gen* 2012, **6**:511–522.
- Gill P: An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *Int J Legal Med* 2001, **114**:204–210.
- Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL: Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci Int* 1999, **103**:23–35.
- Allen M, Engström AS, Meyers S, Handt O, Saldeen T, Von Haeseler A, Pääbo S, Gyllenstein U: Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities. *J Forensic Sci* 1998, **43**:453–464.
- Boles TC, Snow CC, Stover E: Forensic DNA testing on skeletal remains from mass graves: a pilot project in Guatemala. *J Forensic Sci* 1995, **40**:349–355.
- Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K: Identification of the remains of the Romanov family by DNA analysis. *Nat Genet* 1994, **6**:130–135.
- Ginther C, Issel-Tarver L, King M-C: Identifying individuals by sequencing mitochondrial DNA from teeth. *Nat Genet* 1992, **2**:135–138.
- Butler JM, Levin BC: Forensic applications of mitochondrial DNA. *Trends Biotechnol* 1998, **16**:158–162.
- Just RS, Loreille OM, Molto JE, Merriwether AD, Woodward SR, Matheson C, Creed J, McGrath SE, Sturk-Andreaggi K, Coble MD, Irwin JA, Ruffman A, Parr RL: Titanic's unknown child: the critical role of the mitochondrial DNA coding region in a re-identification effort. *Forensic Sci Int-Gen* 2011, **5**:231–235.
- Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissian C, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek MB, Koriach J, Luong K, Ho SYW, Quintana-Murci L, Behar D, Meller H, Alt KW, Cooper A, the Genographic Consortium: Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun* 2013, **4**:1764.
- Melton T, Clifford S, Kayser M, Nasidze I, Batzer M, Stoneking M: Diversity and heterogeneity in mitochondrial DNA of North American populations. *J Forensic Sci* 2001, **46**:46–52.
- Grubwieser P, Mühlmann R, Parson W: New sensitive amplification primers for the STR locus D2S1338 for degraded casework DNA. *Int J Legal Med* 2003, **117**:185–188.
- Maricic T, Whitten M, Pääbo S: Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 2010, **5**:e14004.
- Knapp M, Stiller M, Meyer M: Generating barcoded libraries for multiplex high-throughput sequencing. In *Edited by Shapiro B, Hofreiter M*. Edited by Ancient DNA. Humana Press: New York; 2012:155–170.

26. Fu Q, Mittnik A, Johnson PLF, Bos K, Lari M, Bollongino R, Sun C, Giemisch L, Schmitz R, Burger J, Ronchitelli AM, Martini F, Cremonesi RG, Svoboda J, Bauer P, Caramelli D, Castellano S, Reich D, Pääbo S, Krause J: **A revised timescale for human evolution based on ancient mitochondrial genomes.** *Curr Biol* 2013, **23**:553–559.
27. Hodges E, Xuan Z, Baliya V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR: **Genome-wide *in situ* exon capture for selective resequencing.** *Nat Genet* 2007, **39**:1522–1527.
28. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES, Nussbaum C: **Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing.** *Nat Biotechnol* 2009, **27**:182–189.
29. Cooper A, Poinar HN: **Ancient DNA: do it right or not at all.** *Science* 2000, **289**:1139–1139.
30. Swango KL, Timken MD, Chong MD, Buoncristiani MR: **A quantitative PCR assay for the assessment of DNA degradation in forensic samples.** *Forensic Sci Int* 2006, **158**:14–26.
31. Adler CJ, Haak W, Donlon D, Cooper A, the Genographic Consortium: **Survival and recovery of DNA from ancient teeth and bones.** *J Archaeol Sci* 2011, **38**:956–964.
32. **FASTX-Toolkit.** [http://hannonlab.csh.edu/fastx\\_toolkit/index.html](http://hannonlab.csh.edu/fastx_toolkit/index.html).
33. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *EMBnet Journal* 2011, **1**:1–10.
34. Andrews S: **FastQC. A quality control tool for high throughput sequence data.** <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>.
35. Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli E-L, Silva NM, Kivisild T, Torroni A, Villems R: **A 'Copernican' reassessment of the human mitochondrial DNA tree from its root.** *Am J Hum Gen* 2012, **90**:675–684.
36. **TMAP v3.2.1.** <https://github.com/nh13/TMAP>.
37. Merriman B, Torrent I, Rothberg JM: **Progress in Ion Torrent semiconductor chip based sequencing.** *Electrophoresis* 2012, **33**:3397–3417.
38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The sequence alignment/map format and SAMtools.** *Bioinformatics* 2009, **25**:2078–2079.
39. **Picard Tools.** <http://picard.sourceforge.net>.
40. Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L: **mapDamage: testing for damage patterns in ancient DNA sequences.** *Bioinformatics* 2011, **27**:2153–2155.
41. **Geneious v6.0 created by Biomatters.** <http://www.geneious.com/>.
42. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, Burger J: **Ancient DNA from the first European farmers in 7500-year-old Neolithic sites.** *Science* 2005, **310**:1016–1018.
43. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, der Sarkissian C, Brandt G, Schwarz C, Nicklisch N, Dresely V, Fritsch B, Balanovska E, Villems R, Meller H, Alt KW, Cooper A, the Genographic Consortium: **Ancient DNA from European early Neolithic farmers reveals their near eastern affinities.** *PLoS Biol.* 2010, **8**:e1000536.
44. Sánchez-Quinto F, Schroeder H, Ramirez O, Ávila-Arcos MC, Pybus M, Olalde I, Velazquez A, Marcos MEP, Encinas JMV, Bertranpetit J, Orlando L, Gilbert MTP, Lalueza-Fox C: **Genomic affinities of two 7,000-year-old Iberian hunter-gatherers.** *Curr Biol* 2012, **22**:1494–1499.
45. Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, Pääbo S: **A complete mtDNA genome of an early modern human from Kostenki, Russia.** *Curr Biol* 2010, **20**:231–236.
46. Green RE, Malaspina A-S, Krause J, Briggs AW, Johnson PL, Uhler C, Meyer M, Good JM, Maricic T, Stenzel U, Prüfer K, Siebauer M, Burbano H, Ronan M, Rothberg JM, Egholm M, Rudan P, Brajković, Kučan Z, Gušić I, Wikström M, Laakkonen L, Kelso J, Slatkin M, Pääbo S: **A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing.** *Cell* 2008, **134**:416–426.
47. Yeo ZX, Chan M, Yap YS, Ang P, Rozen S, Lee ASG: **Improving indel detection specificity of the Ion Torrent PGM benchtop sequencer.** *PLoS One* 2012, **7**:e45798.
48. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y: **A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.** *BMC genomics* 2012, **13**:341.
49. Golan D, Medvedev P: **Using state machines to model the Ion Torrent sequencing process and to improve read error rates.** *Bioinformatics* 2013, **29**:i344–i351.
50. Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW: **Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data.** *PLoS Comput Biol* 2013, **9**:e1003031.
51. Koehnemann S, Pfeiffer H: **Application of mtDNA SNP analysis in forensic casework.** *Forensic Sci Int-Gen* 2011, **5**:216–221.
52. Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, Campos PF, Samaniego JA, Gilbert MTP, Willerslev E: **The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils.** *Proc Roy Soc B: Biol Sci* 2012, **279**:4724–4733.
53. Irwin JA, Parson W, Coble MD, Just RS: **mtGenome reference population databases and the future of forensic mtDNA analysis.** *Forensic Sci Int-Gen* 2011, **5**:222–225.
54. Just RS, Loney MD, Barritt SM, Los CW, Smith BC, Holland TD, Parsons TJ: **The use of mitochondrial DNA single nucleotide polymorphisms to assist in the resolution of three challenging forensic cases.** *J Forensic Sci* 2009, **54**:887–891.
55. Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ: **Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians.** *Int J Legal Med* 2004, **118**:137–146.
56. Ballantyne KN, van Oven M, Ralf A, Stoneking M, Mitchell RJ, van Oorschot RAH, Kayser M: **MtDNA SNP multiplexes for efficient inference of matrilineal genetic ancestry within Oceania.** *Forensic Sci Int-Gen* 2012, **6**:425–436.
57. van Oven M, Vermeulen M, Kayser M: **Multiplex genotyping system for efficient inference of matrilineal genetic ancestry with continental resolution.** *Invest Genet* 2011, **2**:6–6.
58. Quintans B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A: **Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing.** *Forensic Sci Int* 2004, **140**:251–257.
59. Wilson MR, DiZinno JA, Polansky D, Replogle J, Budowle B: **Validation of mitochondrial DNA sequencing for forensic casework analysis.** *Int J Legal Med* 1995, **108**:68–74.

doi:10.1186/2041-2223-4-26

**Cite this article as:** Templeton *et al.*: DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investigative Genetics* 2013 **4**:26.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

