



Internal validation and improvement of mitochondrial genome sequencing using the Precision ID mtDNA Whole Genome Panel

Christian Faccinnetto¹ · Daniele Sabbatini² · Patrizia Serventi¹ · Martina Rigato³ · Cecilia Salvoro³ · Gianluca Casamassima¹ · Gianluca Margiotta¹ · Sara De Fanti^{4,5} · Stefania Sarno⁴ · Nicola Staiti¹ · Donata Luiselli⁶ · Alberto Marino¹ · Giovanni Vazza³

Received: 12 April 2021 / Accepted: 13 August 2021 / Published online: 7 September 2021
© The Author(s) 2021

Abstract

With the recent advances in next-generation sequencing (NGS), mitochondrial whole-genome sequencing has begun to be applied to the field of the forensic biology as an alternative to the traditional Sanger-type sequencing (STS). However, experimental workflows, commercial solutions, and output data analysis must be strictly validated before being implemented into the forensic laboratory. In this study, we performed an internal validation for an NGS-based typing of the entire mitochondrial genome using the Precision ID mtDNA Whole Genome Panel (Thermo Fisher Scientific) on the Ion S5 sequencer (Thermo Fisher Scientific). Concordance, repeatability, reproducibility, sensitivity, and heteroplasmy detection analyses were assessed using the 2800 M and 9947A standard control DNA as well as typical casework specimens, and results were compared with conventional Sanger sequencing and another NGS sequencer in a different laboratory. We discuss the strengths and limitations of this approach, highlighting some issues regarding noise thresholds and heteroplasmy detection, and suggesting solutions to mitigate these effects and improve overall data interpretation. Results confirmed that the Precision ID Whole mtDNA Genome Panel is highly reproducible and sensitive, yielding useful full mitochondrial DNA sequences also from challenging DNA specimens, thus providing further support for its use in forensic practice.

Keywords Whole mitochondrial genome · Massive parallel sequencing · NGS · Internal validation · Ion Torrent · Forensic genetics

Christian Faccinnetto and Daniele Sabbatini contributed equally to this work

✉ Christian Faccinnetto
christian.faccinnetto@carabinieri.it

✉ Giovanni Vazza
giovanni.vazza@unipd.it

¹ Reparto Carabinieri Investigazioni Scientifiche Di Parma, Sezione Biologia, Parma, Italy

² Department of Neurosciences DNS, University of Padova, Padova, Italy

³ Department of Biology, University of Padova, Padova, Italy

⁴ Department of Biological Geological and Environmental Sciences, University of Bologna, Bologna, Italy

⁵ Interdepartmental Centre Alma Mater Research Institute On Global Challenges and Climate Change (Alma Climate), University of Bologna, Bologna, Italy

⁶ Department of Cultural Heritage, University of Bologna, Ravenna, Italy

Introduction

Portions of the nuclear genome, such as autosomal short tandem repeat (STR) loci, are more often utilized during forensic investigations, as their typing results are much more informative than mitochondrial DNA (mtDNA). However, STR analysis on degraded samples remains highly controversial, mainly due to the difficulty of reliably interpreting results. Nonetheless, the lack of recombination, the abundance in cells, the matrilineal inheritance, and the apparently lower sensitivity to degradation make the mtDNA molecule highly suitable for forensic genetic investigations, especially for the analyses of bones, hairs shafts, teeth, and highly degraded biological samples [1].

So far, mtDNA analysis has been mainly based on the Sanger-type sequencing (STS) of the hypervariable segments (HVR-I, HVR-II) of the control region (CR), consisting of about 1100 bp with high polymorphic information content [2]. More recently, however, the development

of robust sequencing protocols, the growth of high-quality databases, and the publication of guidelines for typing, annotating, and interpreting results from the International Society for Forensic Genetics allowed the scientific community to validate and rationalize the analysis of mtDNA in forensic casework [3, 4].

The recent advances in next-generation sequencing (NGS) technologies have generated an increasing interest in the use of mtDNA in forensic sciences, because of the potential capability of NGS to capture variants along the entire mitochondrial genome and to detect heteroplasmy at very low levels [5–7].

Among the recently developed NGS platforms, Illumina and Ion Torrent are the most used in forensic laboratories currently. Both platforms have been coupled with several typing kits and software extensions for sequencing and analysing the control region and the entire mitochondrial genome, reporting in general good performances in terms of reproducibility, specificity, and sensitivity when tested for casework implementation [8–12].

Despite these results, still further data and experience need to be accumulated to optimize technical and analytical protocols, and fully understand how these methods behave toward some biological and non-biological mtDNA features, as for example low-level heteroplasmy, mixture/contamination, and primer-binding site polymorphisms.

In this work, we describe the internal validation study for the NGS-based typing of the mtDNA genome, conducted using the Precision ID mtDNA Whole Genome Panel v.2.2 (Thermo Fisher Scientific, Waltham, MA, USA) on the Ion S5 system (Thermo Fisher Scientific), in accordance with the Validation Guidelines for Forensic DNA Analysis Methods of the Scientific Working Group on DNA Analysis Methods (SWGDM) [13] and the European Network of Forensic Science Institute (ENFSI, Recommended Minimum Criteria for the Validation of Various Aspects of the DNA Profiling Process) [14].

Materials and methods

Sample description

The selected samples, analysed at the Reparto Carabinieri Investigazioni Scientifiche di Parma—RIS (Italy), were divided into a series of experiments in order to assess concordance, repeatability, reproducibility, sensitivity, and casework performance of the Precision ID mtDNA Whole Genome Panel (Table S1).

Concordance was determined in 6 replicates of each of the forensic standard control DNA samples 9947A (Thermo Fisher Scientific) and 2800 M (Promega; Madison, WI, USA), analysed using 0.1 ng of genomic DNA (gDNA)

according to manufacturer's recommendations. In addition, 8 saliva samples (DB4523, DB4538, DB4553, DB4579, DB4582, DB4588, DB4595, and DB4597), collected from anonymous volunteers after informed consent, were analysed using 0.1 ng of input gDNA and compared with their mtDNA control regions (CR) obtained at RIS by conventional Sanger-type sequencing (STS) (see Supplemental material for STS protocol).

An inter-laboratory concordance study was also performed by comparing NGS data produced with the Ion S5 system (Thermo Fisher Scientific) at RIS with those produced with an Ion PGM™ System (Thermo Fisher Scientific) at the Laboratory of Molecular Anthropology (Department of Biological, Geological and Environmental Sciences—BiGeA, University of Bologna, Italy). Two single-source samples from buccal swabs (BO08 and BO09), obtained from voluntary adults following written informed consent, were analysed at RIS, using manufacturer's recommended 0.1 ng input gDNA, and at BiGeA, following the protocol described in De Fanti and collaborators [15].

The precision and accuracy of the Precision ID mtDNA Whole Genome Panel were assessed with two experiments. First, a repeatability study was performed using 0.1 ng of input gDNA of forensic standard control 9947A, analysed in 6 replicates by the same scientist on the same Ion S5 System. Second, a reproducibility study was performed using 0.1 ng of input gDNA of control samples 9947A and 2800 M; each examined in triplicate by two different scientists on the same Ion S5 System.

For sensitivity studies, 13 serial dilutions were manually prepared in UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific) using the 9947A control DNA, and examined in triplicate at the final gDNA input of 0.1 ng (X.1), 20 pg (X.2), 10 pg (X.3), 5 pg (X.4), 2.5 pg (X.5), 1.2 pg (X.6), 0.6 pg (X.7), 0.3 pg (X.8), 0.15 pg (X.9), 0.075 pg (X.10), 0.0375 pg (X.11), 0.01875 pg (X.12), and 0.009375 pg (X.13).

The casework study sample set consisted of 6 challenging DNA specimens (6 hairs shafts) analysed using ≤ 0.1 ng input gDNA, and their putative reference samples (1 post-mortem blood and 1 buccal swab) at the final gDNA input of 0.1 ng.

DNA extraction and quantification

Since mtDNA typing is extremely sensitive to contamination, in both the RIS and BiGeA laboratories, all pre-amplification analyses were performed in a dedicated mtDNA laboratory, physically and logistically isolated from post-PCR facilities, following rigorous laboratory measures commonly used for ancient DNA analysis [16].

Saliva samples were collected with the Oragene DNA (OG-500) kit (DNA Genotek; Ottawa, Ontario, Canada) and

extracted at BiGeA with the prepIT-L2P kit (DNA Genotek). Human buccal cell samples for the inter-laboratory concordance study were extracted at BiGeA with the QIAamp® DNA Mini Kit (Qiagen GmbH; Hilden, Germany) following manufacturer's protocol.

Casework hair shaft samples were previously cleaned with different washing steps (10% Sodium hypochlorite, UltraPure™ DNase/RNase-Free Distilled Water, and 70% Ethanol), and DNA was extracted at the RIS laboratory using the Tissue and Hair Extraction Kit combined with the DNA IQ™ System (Promega) following manufacturer's protocols for mitochondrial DNA isolation. Finally, the DNA of the putative reference samples of the casework specimens were isolated at the RIS using the EZ1® DNA Investigator® Kit (Qiagen) with BioRobot EZ1 system (Qiagen) [17].

The amount of gDNA was determined in duplicate using the Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific) on an Applied Biosystems 7500 Fast Real-Time PCR instrument and the HID Real-Time PCR Software v 1.2 (Thermo Fisher Scientific). DNA samples were then diluted in UltraPure™ DNase/RNase-Free Distilled Water to achieve 0.1 ng of final DNA, except for the sensitivity and casework study samples, where different concentrations were prepared as described above.

Library and template preparation

Library preparation was performed using the Precision ID mtDNA Whole Genome Panel and the AmpliSeq™ Precision ID Library Kit 2.0 (Thermo Fisher Scientific) according to manufacturer's user guide (revision B.0, 2019) for a "two-in-one" method. This panel, specifically developed for forensic applications, consists in a 2-pool AmpliSeq multiplex assay of 81 primer pairs (average amplicon length of 162 bp with amplicon overlap of 11 bp), and a large number of additional degenerate primers (~280) to ensure full amplicon coverage with extremely degraded specimens [10, 18, 19].

Both extraction negative and library negative controls were included and processed identically to positive samples, to monitor the absence of contamination throughout the laboratory process. Furthermore, the 9947A forensic standard control DNA was used as a positive amplification control and a library preparation control during the experimental steps.

After primer digestion and adapter ligation, all the libraries were quantified using the Ion Library TaqMan™ Quantification Kit (Thermo Fisher Scientific) following manufacturer's instructions and normalized to 30 pM in order to ensure an equal representation of each library in the pool. Samples with less than the desired 30 pM library concentration were used undiluted for library pooling. The barcoded libraries were combined (5 µl of each), and the resulting pool was quantified in triplicate to verify that the

previous library normalization step was performed correctly. Fully automated template preparation, enrichment of template-containing beads, and chip loading were performed by the Ion Chef™ System (Thermo Fisher Scientific) using the Ion S5™ Precision ID Chef Kit (Thermo Fisher Scientific), according to manufacturer's recommendations (ThermoFisher Scientific Application Guide: Precision ID mtDNA Panels with the HID Ion S5™/HID Ion GeneStudio™ S5 System, Revision B.0; January 19, 2019).

Sequencing and data analysis

All samples analysed in this study were processed in five sequencing runs on an Ion S5™ System using the Ion S5™ Precision ID Sequencing Kit and the Ion 520™ sequencing chips (Thermo Fisher Scientific) following manufacturer's instructions. Primary sequencing data were obtained using the Torrent Suite™ Software (TSS) v.5.10 (Thermo Fisher Scientific) and aligned to the revised Cambridge reference sequence (rCRS + 80) (NCBI reference NC_012920) [20] plus a repetition of 80 nucleotides after position 16,569 (as provided by Thermo Fisher Scientific) with default alignment options. Secondary sequencing analyses were performed with the HID Genotyper v.2.1 plug-in with Converge Software v.2.1 (both Thermo Fisher Scientific). Converge uses "mito variant caller" (MVC) [21], an optimized Smith-Waterman alignment algorithm [22] that integrates PhyloTree mtDNA phylogeny (<http://www.phyloree.org>) [23] and EMPOP (<http://www.empop.org>) [24] information into the scoring function. The MVC parameters were set as default: a minimum total read coverage per position of 20 reads, a minimum variant coverage of 20 reads to call, a coverage threshold to mark a region of 20 reads, and a minimum coverage percent compared to the median of the amplicon of 5.0. Additional MVC parameters included the following thresholds: 96.0 for confirming variant calls, 10.0 for point heteroplasmy (PHP), 20.0 for insertion, and 30.0 for deletion. The results concerning variant, coverage, and quality score were generated automatically by Converge Software v.2.1 in tabular format and linear-circular plots. Read alignments from the mvc.BAM (binary alignment map) and mvc.BAI (binary alignment index) files were manually inspected by two different scientists to confirm variants and anomalies, using both Integrative Genomes Viewer v.2.4.16 (IGV) [25] and the mitoIGV tool available in Converge. Mitochondrial haplogroups provided by Converge were confirmed using Haplogrep2 [26], a web server based on PhyloTree, Built 17. Data handling was performed with custom Python scripts, statistical analyses were done in R (version 4.0.2), and graphics were generated using the ggplot2 R package (version 2.2.1).

Results and discussion

Analysis of amplification-negatives

As previously reported, it is well known that NGS methods tend to produce not negligible levels of background noise in sequencing results. This can be due both to technical artefacts during library preparation and target amplification (i.e. cytosine deamination, guanine oxidation) and to exogenous contamination of reagents, consumables or the laboratory environment [27–29]. To assess the potential impact of background noise, we analysed results from 9 amplification negatives sequenced throughout the study. Negative controls were distributed across the 5 runs as follows: one per run for runs 1, 2, and 3; four in run 4; and two in run 5.

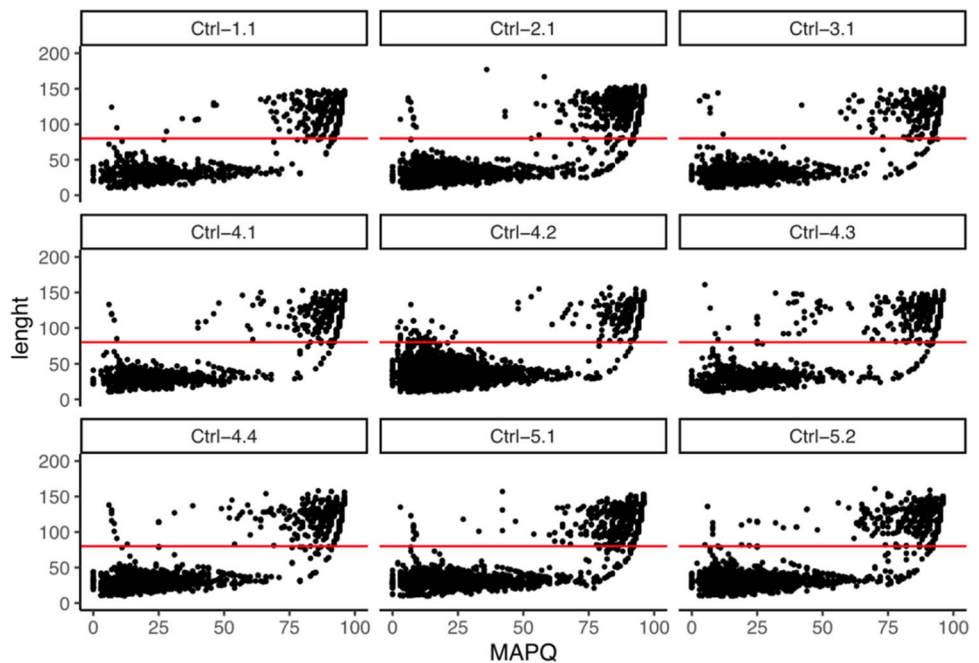
Table 1 Short reads in amplification negatives

Sample	Total mapped reads	Reads < 80 bp	% of reads < 80 bp
Ctrl-1.1	2972	1762	59.2
Ctrl-2.1	7244	3958	54.6
Ctrl-3.1	3049	2094	69.0
Ctrl-4.1	2829	1807	63.8
Ctrl-4.2	6301	5494	87.1
Ctrl-4.3	2963	1852	62.0
Ctrl-4.4	6386	4982	78.0
Ctrl-5.1	8764	6334	72.2
Ctrl-5.2	6949	5247	75.5

None of the 9 negative controls provided a complete mtDNA sequence, nor usable data for comparison; however, all of them exhibited few amplicons covered by aligned reads with a local maximum coverage ranging from 77 to 657. A deeper investigation of these reads clearly showed that most of them (mean of 69.05%, 95% CI 47.15–90.95%) were noticeably short reads (Table 1) with poor mapping quality (MAPQ) scores, as reported in Fig. 1. Conversely, long reads with high-quality alignment were a small proportion of the signal in all negatives.

Despite the origin of these reads being unclear, they can result at least in part from non-target micro-amplicons, created by the forward primer of one targeted amplicon and the reverse primer of a second, overlapping amplicon, or from very low levels of contaminating DNA. Regardless, these reads seem shorter than all the amplicons of the Precision ID mtDNA Whole Genome Panel whose expected size ranges from 125 to 174 bp. To reduce the noise associated with these short reads, we performed a cluster analysis of reads based on their length and MAPQ. The analysis identified the coordinates of the centroids of 2 clusters, of short low-quality and long high-quality reads, with an average of midpoints of 74.16 bp for length (95% CI 70.62–79.7) and of 57.6 for MAPQ (95% CI 55.95–59.27). We thus defined a length cut-off threshold of 80 bp and removed sequences < 80 bp by setting this filter in Torrent Suite™ Software before alignment of reads and variant calling. We also evaluated the possibility to include a MAPQ threshold for the filtering, which would likely provide more accurate results, but unfortunately, the MAPQ score was not easily manageable in Torrent Suite™ Software. A similar approach, based on

Fig. 1 Mapped reads in amplification negatives. The scatter-plots show the distribution of read length vs MAPQ in each amplification negative. The red line indicates the length cut-off value of 80 bp



read length only, has been effectively adopted by Michael D. Brandhagen and colleagues for the validation of another NGS-based mitochondrial DNA analysis, with the PowerSeq™ CRM Nested System (Promega) on the Illumina MiSeq platform [12]. Further, our read filtering step successfully removed the noise due to short sequences, providing an average read depth per position in negative controls of 9.4 reads (standard deviation of 5.9).

Interestingly, the removed short reads were not randomly distributed across the mtDNA but mainly occurred in the mt_125, mt_139, mt_164, mt_130, and mt_133 amplicons (also considering different runs), thus suggesting that they may be technical artefacts due to the formation of micro-amplicon sequences (Table S2). In contrast, longer reads left after filtering displayed a random distribution in negative controls, between and within different runs.

Concordance study

Concordance was assessed by (i) comparing standard control DNAs with their known reference mtDNA profiles (6 replicates of control DNAs 9947A and 2800 M) [12, 30, 31], (ii) comparing sequences from 8 saliva samples to their previously obtained STS sequences of the CRs mtDNA (DB4523, DB4538, DB4553, DB4579, DB4582, DB4588, DB4595, and DB4597), and (iii) comparing whole mtGenomes sequenced in two laboratories with different Ion Torrent-based sequencing platforms (BO08 and BO09 samples) (Table S1).

The analysis of all the samples followed the same procedure as above including the short-read filtering (< 80 bp) to reduce noise. Although in these samples short reads were on average a small fraction of all mapped reads (2.70%, 95% CI 1.88–3.5%), they showed a non-homogeneous distribution similar to that seen in negative controls, and they could thus impact variant calling and heteroplasmy interpretation.

Across the 22 samples, the median value of the total number of mapped reads per sample was 202,312 (interquartile range [IQR], 183,269–241,234) with an average of 233,399. The median of the average base coverage per sample was 1383 reads ([IQR] 1246–1650), with an average of 96.87% uniformity of base coverage. The average amplicon coverage per sample had a median value across the samples of 1254 reads ([IQR] 1135–1495) with an average of 1448 reads (Table S3).

The entire mtDNA sequence was obtained for all the samples, except for two saliva specimens (DB4523 and DB4588), each presenting gaps on their mtDNA sequence due to low coverage amplicons. In sample DB4523, a fragment of 33 bp (nucleotide position (np) 13,216–13,248, margins included) and one of 69 bp (np 13,248–13,316)

were missed in amplicon mt_129. Likewise, in sample DB4588, a gap of 91 bp spanning from amplicons mt_12 to mt_13 was observed at np 10,062–10,152. Moreover, a lower quality and filtering scores were observed around the poly-C stretches np 303–310 (HVS-II) and np 16,184–16,193 (HVS-I), possibly because of an inaccurate flow-call and pH fluctuation in the proton-based system. Also, a specific region comprising amplicons mt_79 to mt_80 (np 8248–8256) was found to be affected by extreme overall read strand bias which could be the result of amplification or alignment errors or contamination by nuclear mitochondrial DNA (NUMT) [19]. To assess the potential impact of the short reads filtering on the coverage of these regions, we reanalysed the samples without applying the read size cut-off. No substantial improvement in coverage of these regions was observed.

Overall, a total of 459 sequence variants were identified by NGS. Fully concordant mtDNA profiles were obtained for all the replicates of the 2800 M control DNA, as well as for the control region of the 8 samples previously analysed by Sanger sequencing. Some discrepancies were instead observed in the profiles of two samples: BO08 and 9947A. In the BO08 sample, which was analysed in two different labs by different NGS protocols and platforms (see “Sample description” in “Materials and methods”), multiple insertions at the homopolymeric regions 513–525 (524.AC) were detected by the Ion S5 (85.6% variant frequency) but not by the Ion PGM. STS based-typing of the CRs (see STS protocol in supplementary data), confirmed the presence of the AC repeat variant.

In all the replicates of the 9947A control DNA, the haplotypes were concordant with the NGS data described in Riman et al. [30], except for two of the three known point heteroplasmies (PHP). Although PHP at position 7861Y (T/C) was correctly reported in all the 9947A replicates (mean variant frequency 17.1%), the other two known heteroplasmic sites at positions 1393R (G/A) and 3242R (G/A) were never identified with the 10% default threshold. Again, to rule out the potential impact of the short-read filtering on the identification of these PHP, we reanalysed the samples without this filtering step and we obtained the same results. The differences in PHPs observed in 9947A control DNA are described below. With the mtDNA sequence variants detected, a phylogenetic check was performed in HaploGrep2; haplogroups are reported in Table S4. The mtDNA haplogroup composition of the 12 different samples (and replicates) revealed a typical Western European example of Eurasian haplogroups. Indeed, the samples analysed were assigned to 6 different mtDNA lineages and sub-lineages belonging to the H, HV, R, T, U, and X macro-haplogroups (Table S4).

Repeatability and reproducibility

Repeatability and reproducibility of the PrecisionID mtDNA Whole Genome Panel were evaluated by comparing sequencing results from libraries prepared by a single scientist from identical samples on the same Ion S5 instrument, and by two different scientists on the same Ion S5 platform (Table S1 and S3). The depth of coverage pattern was similar among all the replicates: the average base coverage across samples was 2507 reads (median, 2224; [IQR] 1650–2766, with an average of 96.31% uniformity of base coverage), the average amplicon coverage across the samples was 2752 (median, 2034; [IQR] 1495–2853) reads, and the mean number of mapped reads across all the samples was 368,202 (median 329,142; [IQR] 241,234–403,842) (Table S3). The haplotype calls for the 18 full mtDNA were completely concordant with the reported 9947A and 2800 M sequences [12, 30, 31] except for the above mentioned PHPs 1393R and 3242R in 9947A replicates.

Case type study

Forensic specimens that contain few copies of DNA by their nature, such as hair shafts, human skeletal remains, or highly degraded samples, remain a challenge to the forensic DNA typing community. In order to evaluate the performance of the Precision ID mtDNA Whole Genome Panel in typical forensic specimens, a total of 6 hair shafts (9–1, 24–1, 24–2, 30–2, 51–2, and 51–3) were analysed twice in two different runs with their putative reference samples (49–28 blood sample and 53–1 buccal swab). All those samples were previously subjected to STR analysis by conventional CE-based typing and sequenced at their mtDNA CRs using STS. Full STR profiles were obtained only for the two reference samples, while from STS analysis, partial HVS-I sequences were obtained for 5 hair shafts (9–1, 24–1, 24–2, 51–2, 51–3). Using NGS, hair samples 30–2 and 51–3 yielded very partial mtGenome sequences with large uncovered regions, possibly because of the severely degraded mtDNA. Therefore, these two samples were excluded from further interpretation results. Full mtDNA sequences were obtained for the two reference samples (49–28, 53–1) and for four (9–1, 24–1, 24–2, 51–2) of the six hair shaft samples. For these latter four casework samples, the median of average base coverage in NGS data was 1254 reads ([IQR] 1142–1607), the median of average amplicon coverage was 1182 ([IQR] 1075–1507) reads, while the median of total number of mapped reads was 188,089 ([IQR] 173,292–241,014) (Table S5).

MtDNA sequence obtained via the Precision ID mtDNA Whole Genome Panel was consistent with the available partial HVS-I Sanger data for each sample (Table S4). In addition, an identical mtDNA (haplogroup N9a1) was obtained from hair samples 24–1, 24–2, and 51–2, and the reference

sample 49–28 (blood). Analogously, hair shaft 9–1 showed the same haplotype (haplogroup D4e1a1) of the reference sample 53–1 (buccal swab), indicating a complete concordance of the results obtained with the two methods and that the Precision ID mtDNA Whole Genome Panel is fit for the purpose to use with typical forensic specimens.

Analysis of variant frequency distribution

Previous studies suggested that the haplotype purity (i.e. the number and percentage of reads supporting a sequence variant) may be diagnostic in ascertaining the correctness of that variant call [8]. High levels of noise due to PCR artefacts, alignment errors, or NUMTs can, therefore, complicate the analysis, leading to false-positive or false-negative calls, as well as false heteroplasmy detection.

The distribution of variant frequencies (VarFreq) in the 9947A and 2800 M standard controls was then analysed and compared within- and between-run replicates used for the concordance study. No significant differences in the overall VarFreq distribution were observed within and between runs (6 replicates in 2 runs for each control DNA) thus confirming the reproducibility of the system. The empirical cumulative distribution function of VarFreq was computed considering all replicates divided into dilution series to obtain the proportion of positions \leq VarFreq in a range of 0–99 VarFreq. Of the 16,648 mtDNA positions, 503 (3%) showed a VarFreq $<$ 99% and 53 (0.3%) $<$ 90%. The positions with VarFreq $<$ 90% showed a remarkable recurrence in multiple replicates of the same sample as well as across different samples. In particular, some of them (i.e. positions 310 and 16,189) are associated to *C-stretch* variations, and the difficulties in variant calling within homopolymer stretches have already been reported for Ion systems [8, 30]. Further positions corresponded to known NUMTs and artefact sequences reported in previous studies, such as region 8248–8256 np [19, 32] while others showed a consistent noise level across samples and runs (Fig. 2; Table 2). However, it is known that NUMTs are more apt to be co-amplified by short-amplicon PCR approaches, like the ion technique, in respect of the long-amplicon PCR [33]. Despite neither false-positives nor negative calls being identified in our samples, the recurrent noise at these positions could be the source of erroneous calls and PHP detection; therefore, variants in these positions should be interpreted with caution.

Sensitivity study

Sensitivity of the Precision ID mtDNA Whole Genome Panel was tested using the 9947A forensic standard control with different input amounts of gDNA ranging from 100 to 0.009375 pg (Table S1). Each DNA quantity was amplified in triplicate using the same PCR cycle conditions (24 PCR

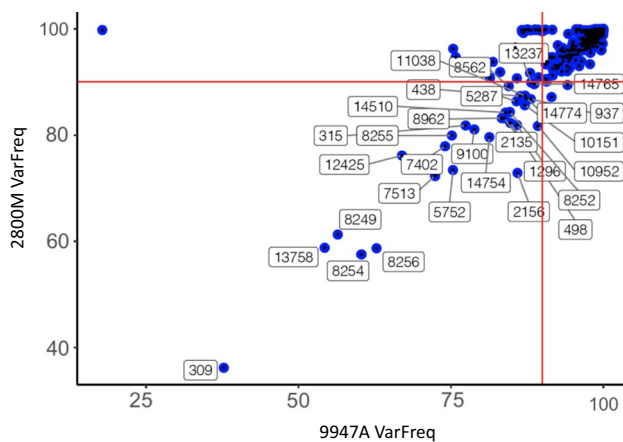


Fig. 2 The scatterplot shows the mean VarFreq of each positions across all the replicates of 9947A (*x* axis) and 2800 M (*y* axis). The red line indicates a VarFreq of 90% (see text). Positions with a VarFreq below 90% in both controls are labelled

cycles). The results of the sensitivity study on the samples are summarized in Table 3. Full mtDNA coverage (100%) was successfully achieved for all the replicates of 9947A down to 0.15 pg input gDNA (X9 dilution), except for one replica at 10 pg (X3 dilution), which resulted with a loss of about 39% of the mtDNA sequence probably due to pipetting errors during amplification. Down to 0.6 pg (X7), all the replicates covering full mtDNA matched the known 9947A mtDNA sequence, with the exception of the two PHP positions mentioned above (1393R and 3242R). For the remaining dilutions with full DNA coverage, 0.3 pg (X8) and 0.15 pg (X9) input gDNA, in one out of the three replicates the expected 7861Y heteroplasmy was also missed. PCR replicates from X10 (0.075 pg) to X13 (0.009375 pg) showed stochastic variation and several gaps on mtDNA with a coverage ranging from 47.3 to 96.7%. Sequencing results of these replicates showed evident stochastic variation (with drops of expected calls and introduction of false-positive calls) when less than 0.15 pg of gDNA was used as PCR input.

Interestingly, the analysis of the cumulative VarFreq distribution of these dilution series reflected the obtained results with clear differences between likely stochastic and non-stochastic outputs (Fig. 3). In fact, all replicates with a good outcome (from X1 to X7) showed a similar VarFreq distribution, while all the replicates with false negative and/or false positive calls (from X8 to X13) displayed a decrease of VarFreq in many positions which corresponds to a loss of haplotype purity. This data could be used to have an immediate picture of the overall quality of the sequencing data. Indeed, independently from the quantitation of mtDNA or gDNA, the cumulative VarFreq distribution could be a good parameter in assessing whether a sequencing run is reliable. However, further studies with

larger sample sizes would be needed to validate and implement this approach.

Overall, the sensitivity study using the 9947A control DNA demonstrated that the Precision ID mtDNA Whole Genome Panel provided full and reliable mtDNA sequences down to 0.6 pg, although we have to consider the lot-to-lot variability in the cell lines used as controls [34]. Using the standard equation provided by the Thermo Fisher Scientific protocol (0.1 ng of gDNA ~ 2900 copies of mtDNA), 0.6 pg of gDNA would contain about 17 copies of mtDNA, clearly that is a rather unrealistic approximation. A specific mtDNA quantification would be required to identify the usage limit of the Precision ID mtDNA Whole Genome Panel in terms of the mtDNA copy number.

Heteroplasmy

Point heteroplasmy detection was assessed in both the high-quantity replicates and the dilution replicates of the sensitivity study of the 9947A control DNA. As previously mentioned, when a standard 10% VarFreq cut-off was applied, only PHP 7861Y (T/C) was correctly called in all the high-quantity replicates. In this case, we obtained a median heteroplasmy of $17.1\% \pm 1.2\%$ slightly higher than the one reported by Riman et al. (about 12%) [30]. The same PHP was correctly called in all the replicates of the dilution series up to X7. However, as reported in Fig. 4A, starting from the X6 dilution, the heteroplasmy quantification became more variable as gDNA input decreased up to a completely random detection and quantification in the most extreme dilutions (X9–X13). On the contrary, the 1393R PHP, expected at a frequency of about 15%, was never called either in high-quantity replicates (median $3.2\% \pm 0.7\%$) or in all the dilution series (Fig. 4B). Of note, the short-read filtering step did not affect the frequency estimate of this PHP. As suggested in recent studies, a possible explanation of these differences may be related to the NGS platform, chemistry, software setting, and the lot-to-lot variability of the control DNAs [30, 35]. Similarly, the low-level heteroplasmy at site 3242, expected at a frequency of about 3%, showed a VarFreq far below 1% in all high-input DNA replicates (median $0.7\% \pm 0.1\%$) as well as in most of the dilution series (Fig. 4C). Lowering the heteroplasmy cut-off to correctly call these low-level PHPs produced a number of false positives, as expected from the previous analysis of the VarFreq distribution of all the mtDNA calls.

Conclusions

We here describe the evaluation of Precision ID Whole mtDNA Genome Panel performed on an Ion S5 platform using control DNAs 2800 M and 9947A as well as typical forensic specimens.

Table 2 Positions with VarFreq < 90% across multiple replicates of 9947A and 2800 M of samples

Position	rCRS base	Sample	VarFreq mean	VarFreq standard deviation	Second most common variant
309*	C	9947A	37.76	23.81	DEL
		2800 M	36.18	5.69	
315	C	9947A	77.40	10.21	INS
		2800 M	81.84	4.08	
438	C	9947A	86.36	6.83	DEL
		2800 M	87.33	4.34	
498	C	9947A	84.75	2.13	DEL
		2800 M	82.32	2.74	
937	T	9947A	88.09	6.60	DEL
		2800 M	86.77	3.67	
1296*	A	9947A	85.84	6.01	DEL
		2800 M	81.91	5.73	
2135	A	9947A	83.75	4.54	DEL
		2800 M	83.21	5.95	
5287*	A	9947A	85.76	6.11	DEL
		2800 M	86.34	1.48	
5752*	A	9947A	75.32	14.18	DEL
		2800 M	73.43	9.89	
7402	C	9947A	74.03	6.31	DEL
		2800 M	77.93	9.17	
7513	T	9947A	72.39	13.67	DEL
		2800 M	72.28	9.02	
8249	G	9947A	56.47	14.13	INS
		2800 M	61.26	4.32	
8252	C	9947A	84.66	4.62	G
		2800 M	84.34	5.54	
8254	C	9947A	60.31	5.32	DEL
		2800 M	57.51	5.16	
8255*	G	9947A	75.14	11.57	DEL
		2800 M	79.94	4.71	
8256	T	9947A	62.78	10.49	DEL
		2800 M	58.67	5.08	
8962	A	9947A	83.35	5.73	DEL
		2800 M	83.22	5.42	
9100	A	9947A	78.83	4.36	G
		2800 M	81.09	3.85	
10,151	A	9947A	87.12	5.91	DEL
		2800 M	85.59	4.16	
11,038	A	9947A	87.17	5.73	DEL
		2800 M	87.41	3.10	
12,425	A	9947A	66.95	12.42	DEL
		2800 M	76.17	4.26	
13,237*	A	9947A	88.15	4.95	DEL
		2800 M	89.89	2.88	
13,758	C	9947A	54.32	6.28	DEL
		2800 M	58.76	4.84	
14,510	A	9947A	83.97	7.57	DEL
		2800 M	84.23	5.51	

Table 2 (continued)

Position	rCRS base	Sample	VarFreq mean	VarFreq standard deviation	Second most common variant
14,754*	C	9947A	81.30	12.77	DEL
		2800 M	79.64	5.44	
14,774*	C	9947A	87.38	10.76	DEL
		2800 M	86.89	3.94	

*NUMT positions reported in Li et al. [32]

Table 3 Proportion of mtDNA sequenced for dilution series

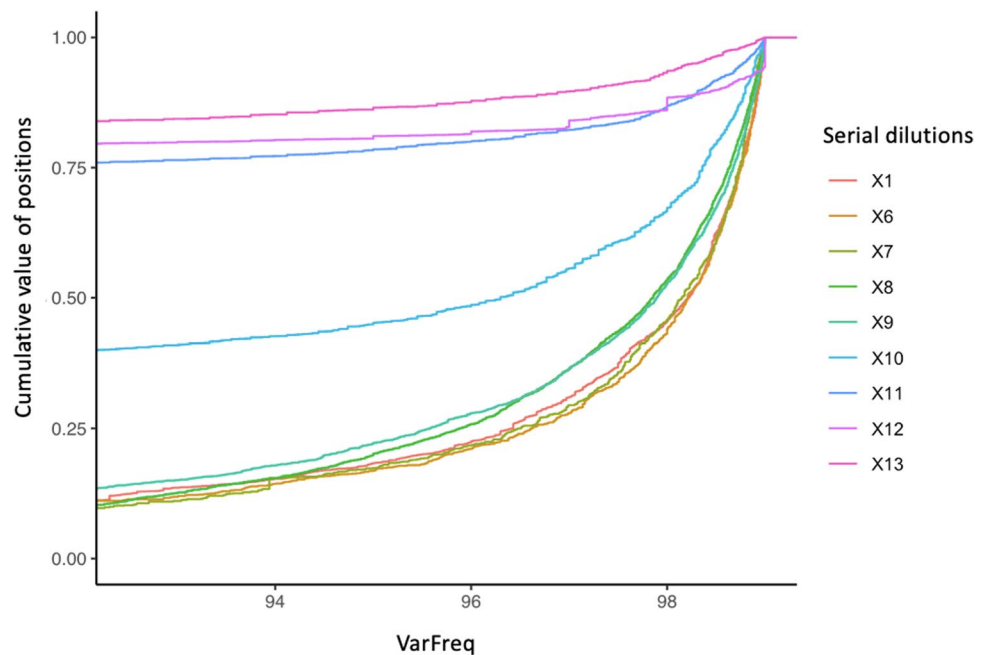
Sample	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
	100%	100%	100%	100%	100%	100%	100%	100%#	100%	90.8%*	96.6%*	80.7%*§	72.4%*
9947A	100%	100%	61.5%*	100%	100%	100%	100%	100%	100%#	91.7%*	63.4%*	91.6%*	47.3%*§
	100%	100%	100%	100%	100%	100%	100%	100%	100%	93.6%*	96.7%*	62.4%*§	49.9%*§

*Incomplete mtDNA profile.

#Undetected 7861Y heteroplasmy.

§Presence of false positive variant calls.

Fig. 3 Cumulative VarFreq distribution computed in the range of 0–99% of each dilution series (considering all replicates). For a better interpretation, the 92.5–99% VarFreq range is shown



A preliminary analysis of negatives highlighted the presence of a non-negligible level of background noise related to the alignment of very short reads. The introduction of a filtering step for reads > 80 bp in the analytical workflow greatly reduced the noise level and improved the accuracy of variant calling. The analysis of VarFreq along the entire mtGenome highlighted the presence of positions with a systematic low VarFreq across different runs and samples. Based on this information, variants and heteroplasmy calls in these positions require

special attention and should be carefully considered by the analysts. Regarding PHP detection, the system has the potential to be extremely sensitive; however, the VarFreq distribution makes it difficult to distinguish between real and false calls when low PHP are considered. Further bioinformatic solutions are desirable to improve overall PHP interpretation. From this perspective, the use of a probabilistic approach based on the analysis of the VarFreq distribution would allow the estimation of probabilities for each PHP call and set a more appropriate PHP threshold

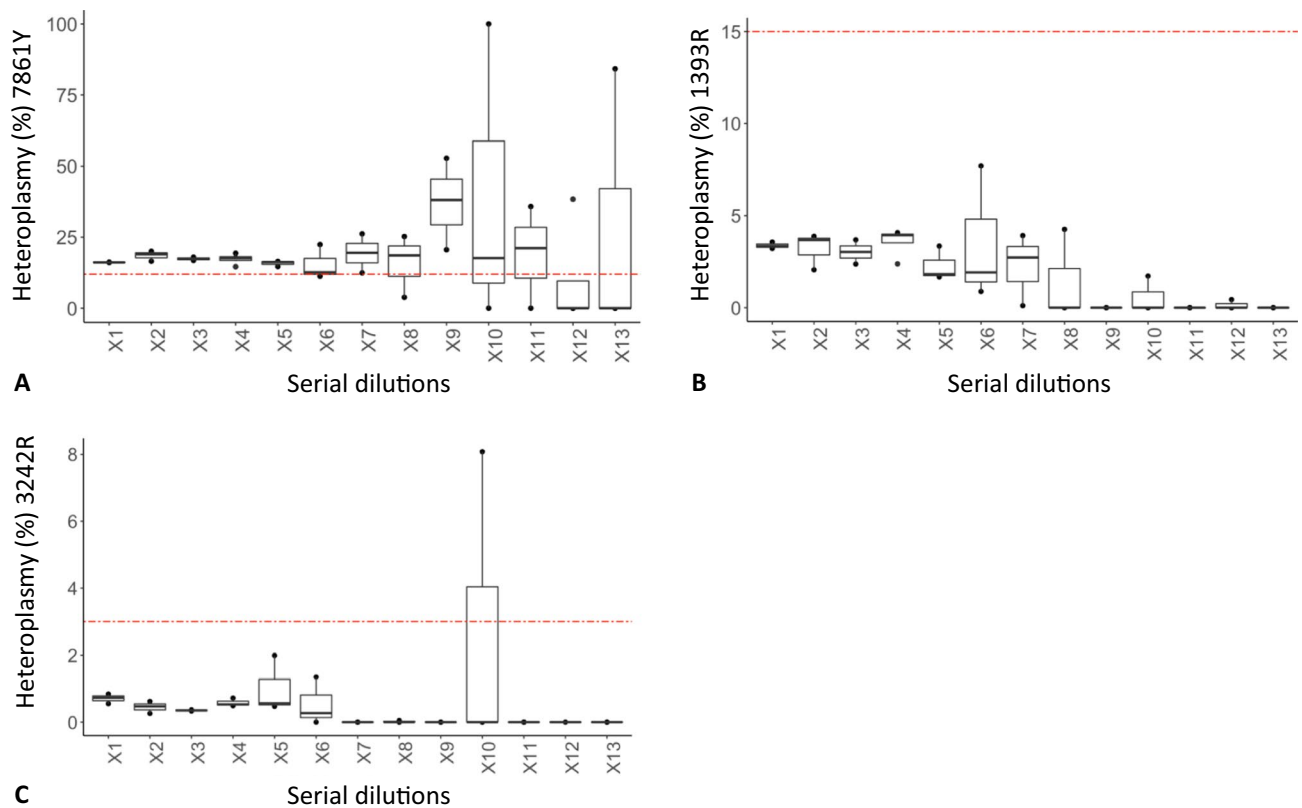


Fig. 4 The plot shows the values of point heteroplasmy in the 9947A dilution series. **A** PHP 7861Y (T/C); **B** PHP 1393R (G/A); **C** PHP 3242R (G/A). The dashed red line indicates the expected heteroplasmy value

for each sample based on sequencing results. Overall, validation experiments confirmed that the Precision ID Whole mtDNA Genome Panel is able to generate accurate, reproducible, and reliable whole mitochondrial genome sequences and is more sensitive than STS on casework forensic samples.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00414-021-02686-w>.

Author contribution All authors contributed to the study conception and design.

Conceptualization: Christian Faccineto, Giovanni Vazza.

Formal analysis and investigation: Daniele Sabbatini, Patrizia Serventi, Martina Rigato, Sara De Fanti, Stefania Sarno, Donata Luiselli.

Writing, review, and editing: Christian Faccineto, Patrizia Serventi, Daniele Sabbatini, Cecilia Salvo, Gianluca Casamassima, Gianluca Margiotta, Sara De Fanti, Stefania Sarno, Nicola Staiti, Donata Luiselli, Alberto Marino, Giovanni Vazza.

Funding Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement.

Data availability The data that support the findings of this study are available from the corresponding authors, G.V. and C.F., upon reasonable request.

Declarations

Ethics approval All procedures performed in studies involving human participants were approved by the Bioethics Committee of the University of Bologna and were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amorim A, Fernandes T, Taveira N (2019) Mitochondrial DNA in human identification: a review. *PeerJ* 7:e7314. <https://doi.org/10.7717/peerj.7314>
- Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am J Hum Genet* 46:828–842
- Parson W, Bandelt HJ (2007) Extended guidelines for mtDNA typing of population data in forensic science. *Forensic Sci Int Genet* 1:13–19. <https://doi.org/10.1016/j.fsigen.2006.11.003>
- Parson W, Gusmão L, Hares DR et al (2014) DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Sci Int Genet* 13:134–142. <https://doi.org/10.1016/j.fsigen.2014.07.010>
- Parson W, Strobl C, Huber G et al (2013) Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM). *Forensic Sci Int Genet* 7:543–549. <https://doi.org/10.1016/j.fsigen.2013.06.003>
- Templeton JEL, Brotherton PM, Llamas B et al (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig Genet* 4:26. <https://doi.org/10.1186/2041-2223-4-26>
- Just RS, Irwin JA, Parson W (2015) Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Sci Int Genet* 18:131–139. <https://doi.org/10.1016/j.fsigen.2015.05.003>
- Woerner AE, Ambers A, Wendt FR et al (2018) Evaluation of the precision ID mtDNA whole genome panel on two massively parallel sequencing systems. *Forensic Sci Int Genet* 36:213–224. <https://doi.org/10.1016/j.fsigen.2018.07.015>
- Pereira V, Longobardi A, Børsting C (2018) Sequencing of mitochondrial genomes using the Precision ID mtDNA Whole Genome Panel. *Electrophoresis* 39:2766–2775. <https://doi.org/10.1002/elps.201800088>
- Strobl C, Eduardoff M, Bus MM et al (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. *Forensic Sci Int Genet* 35:21–25. <https://doi.org/10.1016/j.fsigen.2018.03.013>
- Wu J, Mamidi T, Zhang L, Hicks C (2019) Integrating germline and somatic mutation information for the discovery of biomarkers in triple-negative breast cancer. *Int J Environ Res Public Health* 16:1055. <https://doi.org/10.3390/ijerph16061055>
- Brandhagen MD, Just RS, Irwin JA (2020) Validation of NGS for mitochondrial DNA casework at the FBI Laboratory. *Forensic Sci Int Genet* 44:102151. <https://doi.org/10.1016/j.fsigen.2019.102151>
- Scientific Working Group on DNA Analysis Methods (2016) Scientific working group on DNA analysis methods validation guidelines for DNA analysis methods. https://docs.wixstatic.com/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf
- European Network of Forensic Science Institutes (ENFSI) (2010) Recommended minimum criteria for the validation of various aspects of the DNA profiling Process. http://enfsi.eu/wp-content/uploads/2016/09/minimum_validation_guidelines_in_dna_profiling_-_v2010_0.pdf
- De Fanti S, Vianello D, Giuliani C et al (2017) Massive parallel sequencing of human whole mitochondrial genomes with Ion Torrent technology: an optimized workflow for Anthropological and Population Genetics studies. *Mitochondrial DNA Part A DNA Mapping, Seq Anal* 28:843–850. <https://doi.org/10.1080/24701394.2016.1197218>
- Llamas B, Valverde G, Fehren-Schmitz L et al (2017) From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR Sci Technol Archaeol Res* 3:1–14. <https://doi.org/10.1080/20548923.2016.1258824>
- Anslinger K, Bayer B, Rolf B et al (2005) Application of the BioRobot EZ1 in a forensic laboratory. *Leg Med* 7:164–168. <https://doi.org/10.1016/j.legalmed.2005.01.002>
- Buś MM, Lembring M, Kjellström A et al (2019) Mitochondrial DNA analysis of a Viking age mass grave in Sweden. *Forensic Sci Int Genet* 42:268–274. <https://doi.org/10.1016/j.fsigen.2019.06.002>
- Strobl C, Churchill Cihlar J, Lagacé R et al (2019) Evaluation of mitogenome sequence concordance, heteroplasmy detection, and haplogrouping in a worldwide lineage study using the Precision ID mtDNA Whole Genome Panel. *Forensic Sci Int Genet* 42:244–251. <https://doi.org/10.1016/j.fsigen.2019.07.013>
- Andrews RM, Kubacka I, Chinnery PF et al (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147–147. <https://doi.org/10.1038/13779>
- Roth C, Parson W, Strobl C et al (2019) MVC: an integrated mitochondrial variant caller for forensics. *Aust J Forensic Sci* 51:S52–S55. <https://doi.org/10.1080/00450618.2019.1569150>
- Smith TF, Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol* 147:195–197. [https://doi.org/10.1016/0022-2836\(81\)90087-5](https://doi.org/10.1016/0022-2836(81)90087-5)
- van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30:E386–E394. <https://doi.org/10.1002/humu.20921>
- Parson W, Dür A (2007) EMPOP-A forensic mtDNA database. *Forensic Sci Int Genet* 1:88–92. <https://doi.org/10.1016/j.fsigen.2007.01.018>
- Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>
- Kloss-Brandstätter A, Pacher D, Schönherr S et al (2011) HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat* 32:25–32. <https://doi.org/10.1002/humu.21382>
- Park G, Park JK, Shin SH et al (2017) Characterization of background noise in capture-based targeted sequencing data. *Genome Biol* 18:136. <https://doi.org/10.1186/s13059-017-1275-2>
- Loman NJ, Misra RV, Dallman TJ et al (2012) Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30:434–439. <https://doi.org/10.1038/nbt.2198>
- Bragg LM, Stone G, Butler MK et al (2013) Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Comput Biol* 9:e1003031. <https://doi.org/10.1371/journal.pcbi.1003031>
- Riman S, Kiesler KM, Borsuk LA, Vallone PM (2017) Characterization of NIST human mitochondrial DNA SRM-2392 and SRM-2392-I standard reference materials by next generation sequencing. *Forensic Sci Int Genet* 29:181–192. <https://doi.org/10.1016/j.fsigen.2017.04.005>
- Lee EY, Lee HY, Oh SY et al (2016) Massively parallel sequencing of the entire control region and targeted coding region SNPs of

- degraded mtDNA using a simplified library preparation method. *Forensic Sci Int Genet* 22:37–43. <https://doi.org/10.1016/j.fsigen.2016.01.014>
32. Li M, Schroeder R, Ko A, Stoneking M (2012) Fidelity of capture-enrichment for mtDNA genome sequencing: influence of NUMTs. *Nucleic Acids Res* 40:e137–e137. <https://doi.org/10.1093/nar/gks499>
 33. Sturk-Andreaggi K, Parson W, Allen M, Marshall C (2020) Impact of the sequencing method on the detection and interpretation of mitochondrial DNA length heteroplasmy. *Forensic Sci Int Genet* 44:102205. <https://doi.org/10.1016/j.fsigen.2019.102205>
 34. Cihlar JC, Amory C, Lagacé R et al (2020) Developmental validation of a MPS workflow with a PCR-based short amplicon whole mitochondrial genome panel. *Genes (Basel)* 11:1345. <https://doi.org/10.3390/genes11111345>
 35. Cihlar JC, Peters D, Strobl C et al (2020) The lot-to-lot variability in the mitochondrial genome of controls. *Forensic Sci Int Genet* 47:102298. <https://doi.org/10.1016/j.fsigen.2020.102298>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.