Paolo Fattorini[1]
Carlo Previderé[2]
Ilaria Carboni[3]*
Giorgio Marrubini[4]
Solange Sorçaburu-Cigliero[1]
Pierangela Grignani[2]
Barbara Bertoglio[2]
Paolo Vatta[5]
Ugo Ricci[3]

[1]Department of Medicine, Surgery and Health, University of Trieste, Trieste, Italy
[2]Department of Public Health, Experimental and Forensic Medicine, Section of Legal Medicine and Forensic Sciences, University of Pavia, Pavia, Italy
[3]SOD Diagnostica Genetica, A.O-U Careggi, Firenze, Italy
[4]Department of Drug Sciences, University of Pavia, Pavia, Italy
[5]Scuola Internazionale Superiore di Studi Avanzati (SISSA), Functional and Structural Genomics sector, Trieste, Italy

Accepted January 2, 2017

Performance of the ForenSeq™ DNA Signature Prep kit on highly degraded samples

Next generation sequencing (NGS) is the emerging technology in forensic genomics laboratories. It offers higher resolution to address most problems of human identification, greater efficiency and potential ability to interrogate very challenging forensic casework samples. In this study, a trial set of DNA samples was artificially degraded by progressive aqueous hydrolysis, and analyzed together with the corresponding unmodified DNA sample and control sample 2800 M, to test the performance and reliability of the ForenSeq™ DNA Signature Prep kit using the MiSeq Sequencer (Illumina). The results of replicate tests performed on the unmodified sample (1.0 ng) and on scalar dilutions (1.0, 0.5 and 0.1 ng) of the reference sample 2800 M showed the robustness and the reliability of the NGS approach even from sub-optimal amounts of high quality DNA. The degraded samples showed a very limited number of reads/sample, from 2.9–10.2 folds lower than the ones reported for the less concentrated 2800 M DNA dilution (0.1 ng). In addition, it was impossible to assign up to 78.2% of the genotypes in the degraded samples as the software identified the corresponding loci as "low coverage" (< 50x). Amplification artifacts such as allelic imbalances, allele drop outs and a single allele drop in were also scored in the degraded samples. However, the ForenSeq™ DNA Sequencing kit, on the Illumina MiSeq, was able to generate data which led to the correct typing of 5.1–44.8% and 10.9–58.7% of 58 of the STRs and 92 SNPs, respectively. In all trial samples, the SNP markers showed higher chances to be typed correctly compared to the STRs. This NGS approach showed very promising results in terms of ability to recover genetic information from heavily degraded DNA samples for which the conventional PCR/CE approach gave no results. The frequency of genetic mistyping was very low, reaching the value of 1.4% for only one of the degraded samples. However, these results suggest that further validation studies and a definition of interpretation criteria for NGS data are needed before implementation of this technique in forensic genetics.

Keywords:
DNA degradation / Forensic genetics / Next generation sequencing

1 Introduction

Human identification in forensics is conventionally performed by PCR analysis of a selected core of autosomal STR markers [1, 2]. In addition to this standard approach, and depending on the particular casework, lineage markers (i.e., Y-chromosome and mitochondrial DNA markers), X-chromosome STRs and autosomal SNPs can provide useful tools to complement genetic identifications, to determine exclusions and to generate investigative leads based on phenotypic or ethnicity-predicting DNA markers [3]. All these DNA typing systems can be multiplexed together to maximize the discriminating power required to address

Correspondence: Dr. Paolo Fattorini, Department of Medicine, Surgery and Health, University of Trieste, Italy, Strada per Fiume 447, 34149, Italy
E.mail: fattorin@univ.trieste.it
Fax: +39-04-03996265

*Current address: Institute of Public Health, Section of Legal Medicine, Catholic University, Rome, Italy

Colour Online: See the article online to view Fig. 2 in colour.

most problems of human identification. The final step is then represented by capillary electrophoresis (CE), the primary methodology to separate and detect alleles and base sequences.

During the last ten years, a new technology has been developed, based on the use of high-throughput platforms [4]. The Next Generation Sequencing (NGS) technology, in fact, offers the opportunity to sequence entire genomes or hundreds of selected genes/loci simultaneously in a very efficient way [5–8]. In the last few years, new NGS platforms, based on different sequencing chemistries, were developed to an extraordinary high-throughput level leading to the reduction of the DNA sequencing cost per sample. For example, in ancient DNA (*aDNA*) studies, a discipline with many overlapping issues with forensic genetics [9], the NGS approach, either in the form of direct (shot-gun) or targeted enrichment or sequence capture sequencing, now represents the "gold standard" in technology [10, 11]. More recently, the first kits designed for forensic identification purposes have been developed and customized [12].

Reliability of the results obtained from degraded/damaged DNA samples is a critical concern in forensic genetics as well as in *aDNA* analyses [9]. In fact, in *post mortem* tissues, the degree of DNA degradation increases at a rate, which mostly depends on environmental conditions [13]. It is well known that DNA damage hinders polymerization; consequently, the main feature of degraded samples is that they are quite refractory to PCR amplification [13, 14]. It is therefore a common finding to achieve partial or no profiles at all from these kinds of damaged/degraded DNA samples which can be characterized, in addition, by the presence of PCR artifacts potentially affecting the correct assignment of the genotypes. To overcome this concern, several analytical strategies have been developed such as the increase of the number of PCR cycles [14, 15], the reduction of the size of the amplicons [16, 17], the post-PCR cleanup of the amplified products [18], and, more recently, the repair of the DNA template in a pre-PCR step [19]. Currently, NGS represents the technology with the highest sensitivity; for this reason, there is a strong interest in understanding its ability to recover as much genetic information as possible from highly degraded forensic DNA samples for which the conventional PCR/CE approach failed [12]. This potential clearly emerged from the first study on this topic which was performed using the HID-Ion AmpliSeq™ Identity Panel kit on sonicated DNA samples [20].

In this paper, a trial set of DNA samples was artificially enriched in Apurinic-Apirymidinic sites (A-P sites) [21], the most frequent lesion found in aged and probably even in forensic DNA samples [22]. This set of damaged DNA samples was used to evaluate the performance and reliability of the ForenSeq™ DNA Signature Prep kit (in its beta version) using the Illumina MiSeq NGS sequencer. This kit allows the simultaneous typing of 29 autosomal STR markers, 25 Y-STR markers, 9 X-STR and 95 identity informative SNP markers when using primers mix A [23]. The aim of this study is to evaluate the ability of this new technology to produce reliable genetic results from difficult samples, in comparison with the conventional PCR/CE approach.

2 Materials and methods

2.1 Samples

Sample FM was obtained from the buffy coat derived from 500 mL of a male donor's blood who provided informed consent in agreement with the Ethical Committee of the A.U-O of Trieste (Italy). Briefly, after phenolic extraction and ethanol precipitation, the sample was resuspended in bi-distilled water, quantified by NanoDrop ND-1000 (Thermo Fisher Scientific Inc., MA, USA), aliquoted in Eppendorf tubes and stored at –20°C until use.

To enrich the samples in A-P sites, a method previously published was followed [24]. Briefly, 200 μL aliquots, each one containing about 13 μg of DNA, were incubated in a water bath at 70°C for 6.0, 8.0 and 10.0 hours (namely sample FM-6, FM-8 and FM-10, respectively). After incubation, each sample was immediately centrifuged in TE (10 mM Tris pH 7.4, 0.1 mM Na_2EDTA pH 8.0)-equilibrated Ultracel 3 K Amicon Ultra columns (Millipore, MA, USA) for 30 min at 12 000 rpm. The retained samples were washed once with TE, centrifuged at 12 000 rpm for 20 min and finally recovered following the manufacturer's recommendations. The samples were then adjusted with TE to a final volume of 300 μL and stored at –20°C until further use. A negative control of sample preparation (NCSP), represented by TE filtered through Ultracel 3 K column, was also included.

To assess the molecular weight (MW) of the samples, aliquots of each sample were run through 1.2% agarose gels stained with Ethidium Bromide (0.5 μg/mL). λ/HindIII and Easy Ladder (purchased from Bioline, UK) were run simultaneously as MW markers.

2.2 DNA quantification

One-microliter aliquots from each sample were quantified using NanoDrop spectrophotometry. In addition, qPCR-based methods were used to test the amplifiability of the samples. Quantifiler Human DNA Quantification kit and Quantifiler Human DNA Trio Quantification kit (Thermo Fisher Scientific) were used to test 1.0 μL aliquots of each sample in triplicate assays, following the manufacturer's recommendations, on a 7500 Real Time PCR System (Thermo Fisher Scientific). Calibration was performed in triplicate. No template control and NCSP were included. The degradation index (D.I.) for the Quantifiler Trio was calculated by the ratio between the SA (small amplicon) and LA (large amplicon) quantification data. The UV/qPCR ratios were calculated as previously published [24].

DNA control sample 2800 M was provided with the ForenSeq™ DNA Signature kit at a nominal concentration of 10 ng/μL.

2.3 NGS experiment

2.3.1 Library preparation and sequencing

The *beta*-version of the ForenSeq™ DNA Sequencing kit (Illumina) was used to prepare the libraries. As recommended in the ForenSeq™ DNA Signature Prep Guide [23], one nanogram of the unmodified control sample FM DNA, together with three scalar amounts (1.0, 0.5 and 0.1 ng) of sample 2800 M were used as positive controls.

The maximum input DNA volume (5.0 μL) was utilized for the degraded samples FM-6, FM-8 and FM-10. The resulting DNA amounts were 1.68, 0.840 and 0.055 ng for samples FM-6, FM-8 and FM-10, respectively, as assessed by the *Quantifiler Trio* SA probe. Similarly, two dilutions of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3) containing 0.840 and 0.420 ng, respectively, were also prepared and included in the experiment. In total, nine samples were analyzed in duplicate tests. A negative PCR amplification control (no template—NTC) and the NCSP were processed alongside the steps of the experiment, according to the manufacturer's recommendation [23].

Briefly, the DNA samples were amplified and tagged through 18 cycles in a GeneAmp® PCR System 9700 thermal cycler, according to the protocol described in [23]. The DNA primer mix A (DPMA) selected for this experiment allowed the simultaneous amplification of 62 STR markers (28 autosomal, 25 Y-chromosome, 9 X-chromosome) plus Amelogenin and 95 identity informative SNPs. The target enrichment procedure was completed by adding index adapters to the tagged DNA templates and setting up 15 additional PCR cycles. Each library was then purified and normalized by the bead-based recommended procedure. Five microliter aliquots of each one of the twenty normalized libraries were pooled together. Seven microliters of the pooled libraries were diluted with the hybridization buffer (HT1), added to the Human Sequencing Control (HSC) and denatured following the recommended procedures.

2.3.2 NGS run and data analysis

The sequencing run was performed on the MiSeq® system (Illumina) converted into the MiSeq FGx Forensic Genomic System version according to the manufacturer's protocol [23, 25]. Data analysis was initially performed using the provided *beta* version of the ForenSeq Universal Analysis Software. Quality metrics (intensity, i.e., number of reads and the corresponding alleles) for each run were then converted into a Microsoft Office Excel 2007 spreadsheet and analyzed with Excel and R, version 3.0.1 (2013-05-16, Copyright 2013 The R Foundation for Statistical Computing). The following analyses were then performed.

2.3.3 Molecular analysis

The mean number of reads/sample was calculated from the overall number of reads for the STR and SNP panels in the two experiments. For example, the calculation performed for the 1.0 ng dilution of sample 2800 M is the following: (511 850 + 246 002)/2 = 378 926 (see Supporting Information Table S1). To evaluate the intra-experiment variability (e.g., between the data recovered from the two replicates of each sample), the number of reads for each marker was transformed in two-axes plot diagrams. The concordance was then calculated by r^2.

The average depth of coverage (DoC)/locus was calculated by dividing the overall number of reads generated for a specific panel (in the two runs) by the number of markers of the same panel which gave at least 50 reads (that is the cut-off value for locus call). For example, the 0.1 ng dilution of sample 2800 M gave an overall number of 191 732 reads generated by 125 markers producing at least 50 reads. Therefore, the average DoC/locus was calculated as follows: 191 732/125 = 1534 (see Supporting Information Table S1). The same raw datasets were also used to calculate the median and the min/max values.

The relative DoC (*r*DoC) of each locus was calculated for samples FM and the 1.0 ng dilution of sample 2800 M by dividing the average number of reads of that specific marker by the average number of reads of the corresponding panel. For example, a total of 324 939 and 336 841 reads were generated for the STR panel in the two replicates of control sample FM, respectively (mean = 330 890). In those tests, 22 541 and 23 115 reads were assigned to locus D22S1045, respectively (mean = 22 828). Therefore, the *r*DoC of locus D22S1045 in sample FM was calculated as 22 828/330 890 = 0.0689.

For the STR panel, the percentage of stutter reads was calculated by dividing the number of reads identified as stutter products (e.g., -1 repeat) by the total number of reads of the corresponding STR panel.

2.3.4 Genetic analysis

The default settings of the ForenSeq Universal Analysis Software Guide [26] were applied to the data generated in the experiment. The software was programmed to provide analysis metrics, describing the quality of the data generated locus per locus.

The minimum DoC to assign a genotype depends on the NGS technology and on the end-point of the study [27–30]. The only pivot study performed with the ForenSeq™ DNA Sequencing kit reported a minimum DoC of 10x [31]. Our preliminary data analysis, performed with a cut off of 30x, evidenced mistyping error rates of up to 2.2% in these three trial samples [32]. For this reason, we tried to reduce these error rates by increasing the minimum coverage for "locus call" to 50x. Below this cut-off value, each specific locus was tagged by the software with the flag "low coverage".

The analytical (AT) and interpretation (IT) thresholds were set up as follows: for loci with DoC from 50x to 650x: 10 and 30 reads, respectively; for loci with DoC ≥ 651x: 1.5% and 4.5% of the total reads of the locus, respectively.

The heterozygote balance for each STR and SNP locus was calculated by dividing the number of reads of the lower covered allele by the higher one. Values < 0.60 were tagged by the software with the flag "imbalance".

Stutter bands for the STR loci were identified by the software according to predefined % intensity values of the true allele, as reported in the ForenSeq Universal Analysis Software Guide [26]. These values show a high degree of variability, ranging from 7.5% for D4S441 and Penta E to 50% for DYS481. The software tagged the locus with a "stutter flag" when stutters were above the predefined % intensity values.

The correctness of the allelic calls was initially verified comparing the genotypes recovered from each replicate of sample 2800 M to the expected genotypes of this reference control sample reported in the ForenSeq™ DNA Signature Prep Guide [23]. The allelic calls of each replicate of samples FM-6, FM-8 and FM-10 were compared to the genotypes obtained from the reference sample FM. The comparison allowed an easy identification of the amplification artifacts in each replicate.

The interpretation guidelines for genotyping the NGS data obtained from the two replicate analyses of each sample are based on the "*consensus*" approach [33] and are described as follows, according to [24] with minor changes:

(i) no result (NR) for a given locus was defined when at least one of the two replicates gave either low coverage (e.g., < 50x) or allelic reads below the interpretation threshold;
(ii) the genotype for a given locus was considered correct (C) if both tests gave the expected genotype;
(iii) when different genotypes are obtained at a given locus in the two replicate analyses, the result was considered unreliable (U), as it was not possible to unambiguously assign a genotype for that given locus;
(iv) the genotype for a given locus was considered wrong (W) if both replicates showed the same incorrect genotype.

In order to identify sequence variants, STR alleles generated by NGS were verified with STRait Razor [34] when possible.

2.4 Conventional PCR/CE

The same maximum input DNA volume (5.0 μL) utilized for the NGS analysis of degraded samples FM-6, FM-8 and FM-10, was used for the conventional STR typing approach amplifying the 16 STR loci plus Amelogenin contained in the commercial kit PowerPlex® ESI 17 system (Promega, USA), in duplicate tests. Five-hundred picograms of the unmodified control samples FM and 2800 M were tested. Both NTC and NCSP were included. Thirty cycles of PCR were performed under standard conditions. The resulting amplified PCR products were then separated by capillary electrophoresis using a 3500 Genetic Analyser Sequencer (Thermo Fisher Scientific). The analytical (AT) and interpretation (IT) thresholds were set at 100 and 300 rfu respectively. The genetic typing of the samples was performed following the "*consensus*" approach described above.

3 Results and discussion

3.1 DNA degradation

In order to produce damaged DNA samples useful to test the reliability of the NGS approach, three trial DNAs (namely FM-6, FM-8 and FM-10) were artificially degraded by aqueous hydrolysis, according to a simple and well-defined

Table 1. Molecular features of the trial DNA samples. NanoDrop: amount of DNA as assessed by UV spectrophotometry at 260 nm (mean ± s.d.). qPCR-based quantification: amount of DNA (mean ± s.d.) as assessed by real-time PCR using the *Quantifiler* and *Quantifiler Trio* kits, respectively. The calibration data showed r^2 values > 0.998 in the ranges of quantification (from 50.0 ng to 23.0 pg for *Quantifiler* and from 50.0 ng to 5.0 pg for *Quantifiler Trio*). hTERT: human telomerase reverse transcriptase. Y: Y-specific probe. SA: small autosomal probe. LA: large autosomal probe. In round brackets, the length of the molecular probes, in bp. In square brackets, the UV/qPCR ratios. IPC: C_t values of the Internal Positive Control

Sample	Incubation time at 70 °C (hours)	NanoDrop (ng/μL)	qPCR-based quantification (ng/μL)					
			Quantifiler		Quantifiler Trio			
			hTERT (62 bp)	IPC	Y (75 bp)	SA (80 bp)	LA (214 bp)	IPC
FM (ctrl)	0	32.9 ± 0.7	31.384 ± 2.721 [1.0]	25.2 ± 0.2	29.063 ± 2.454 [1.1]	32.946 ± 3.094 [1.0]	44.531 ± 4.251 [0.7]	27.3 ± 0.1
FM-6	6.0	38.0 ± 0.8	1.305 ± 0.158 [29.1]	25.3 ± 0.1	0.274 ± 0.021 [138.7]	0.336 ± 0.030 [113.1]	< LOQ	26.9 ± 0.1
FM-8	8.0	32.9 ± 1.2	0.399 ± 0.037 [416.5]	25.2 ± 0.2	0.124 ± 0.011 [1,096.7]	0.168 ± 0.007 [1,134.4]	undetermined	26.8 ± 0.2
FM-10	10.0	38.8 ± 1.0	0.045 ± 0.010 [862.5]	25.4 ± 0.2	0.008 ± 0.001 [4,850.0]	0.011 ± 0.001 [3,227.0]	undetermined	26.7 ± 0.1

depurination protocol [24]. This procedure enriches the templates in A-P sites [21], promoting DNA degradation *via* β-elimination [35]. In addition, deamination of the nucleobases occurs in these samples even though at low level [36]. The main molecular features of these three trial samples and of the control sample FM are summarized in Table 1. Samples FM-6, FM-8 and FM-10 exhibited a mean molecular weight (MW) $<$ 500 bp, with a degree of degradation proportional to the length of the treatment (see Supporting Information Fig. S1).

Another important feature is that all treated samples were quite refractory to polymerization as shown by the limited amount of DNA assessed by the four qPCR probes used in this study. This finding emerged clearly when the UV quantification data were compared to the amount of DNA assessed by each molecular probe [24], thus calculating each probe-related UV/qPCR ratio. In addition, it was impossible to calculate the degradation index for the *Quantifiler Trio* data as the LA gave no results for all the trial DNA samples. Negative control samples (NCSP and NTC) gave the expected results (*undetermined* for all the probes). The untreated sample FM showed UV/qPCR ratios close to one (0.739–1.132) and a degradation index of 0.74. These values are typical of high-molecular-weight DNA [24]. No inhibition was detected in all samples as shown by the C_t values of the IPC (Internal Positive Control) probes of the qPCR kits. As shown in Table 1, those values were in fact consistent with the average IPC C_t values of the quantification standards (25.0 $\pm$ 0.2 for *Quantifiler* and 27.2 $\pm$ 0.7 for *Quantifiler Trio*).

In conclusion, the three samples, whose molecular features are described above, can be used as models to mimic problematic forensic DNA samples.

3.2 NGS analysis

3.2.1 MiSeq run data

The MiSeq sequencing run showed a mean cluster density of 1112 K/mm^2 and produced about 13.2 M reads. Of these, about 92.8% passed the quality filters. An overall value of 0.196% of the molecules had a cluster phasing while about 0.088% had a cluster pre-phasing. Therefore, all the parameters of this run completely fulfilled the recommended parameters of MiSeq platforms [26, 31].

3.2.2 Control DNAs (samples 2800 M and FM) and negative controls

The beta version of the ForenSeq™ DNA Signature kit contained four STRs (namely DXS10148, DXS8377, DYS456 and SE33) which are now not present in the current commercial version of the ForenSeq™ DNA Signature kit. In this paper, the genetic results for these four markers were not considered.

3.2.2.1 Control sample 2800 M

The NGS results obtained for the duplicate tests of scalar amounts (1.0, 0.5 and 0.1 ng) of control sample 2800 M are reported in Supporting Information Table S1. The mean number of reads/sample was 378 926, 301 731 and 138 926 for DNA input of 1.0, 0.5 and 0.1 ng, respectively. These values exceeded by 37.7–46.5% those reported elsewhere for the same amounts of this control sample [31]. The most likely explanation for this finding is that in our experiment 20 samples per run were pooled, whereas Churchill et al. [31] pooled 32 samples per run. In addition, it was observed that about 67.1–69.0% of the reads were generated for the STR panel and no more than 2.2–4.5% of the reads of the STR panel were identified as stutter products (–1 repeat).

Reads were generated for all the markers of both panels, in each one of the three 2800 M DNA dilutions. The 0.1 ng DNA dilution showed the highest variability between the two replicates, as indicated by lower r^2 values denoted on the concordance column. The average DoC across STR and SNP panels for the three 2800 M DNA dilutions are shown in Supporting Information Figs. S2 and S3, respectively. No correlation between the average DoC found for each marker and the length of the corresponding amplicon was found for the two panels of markers. This difference in the average DoC within and between STR and SNP markers has already been described [31]. According to previous reports it could be ascribed mainly to a different efficiency of the target enrichment procedure caused by the peculiar melting temperature of each one of the multiplex-couples of primers used in the experiments [14, 31, 37].

Out of the 58 STR loci (plus Amelogenin) considered in this study, low coverage of the locus (e.g., DoC $<$ 50x) was found only for the 0.1 ng 2800 M DNA dilution, with a frequency of 0.009. The imbalance of heterozygous loci was observed mainly in the 0.1 and 0.5 ng dilutions of sample 2800 M (frequency of 0.482 and 0.089, respectively, versus 0.036 of the 1.0 ng dilution), as previously reported [31]. In addition, the most unbalanced allele was the longer one with a frequency of 0.764, a finding already observed in previous studies [38–40]. Two stutter reads above the defined parameters were found: the first one affected one of the two 1.0 ng replicates at locus DYS576 and the second one affected one of the two 0.1 ng replicates at locus D12S391. No other artifact was scored for the STR panel of markers. Sequence variance (e.g., repeat-motif variations) was observed at locus D9S1122 [31] in all the replicates.

Out of the 95 SNP markers considered here, locus rs2342747 gave the homozygous genotype A/A versus the expected A/G in all the tests performed. This locus has been already described to be prone to locus and allele drop outs [31]. Low coverage of the loci was found only in the 0.1 ng dilution with a frequency of 0.016. In addition, for the SNP panel of markers, imbalance was observed mainly in the 0.1 and 0.5 ng dilutions (frequency of 0.428 and 0.128, respectively, versus 0.071 of the 1.0 ng dilution). Reads below the IT were

scored in the 0.1 ng sample with a frequency of 0.028. No other drop outs were observed.

The results described here show the robustness and reliability of the NGS approach even from sub-optimal amount of high-quality DNA [12, 31]. However, it is likely that the low frequency of artifacts found in these dilutions of the control sample 2800 M is due to high number of reads/sample which were obtained, thanks to the reduced number of samples ($n = 20$) pooled in the library.

3.2.2.2 Control sample FM

As described in paragraph 2.3.3, the NGS analysis of 1.0 ng of the control sample FM generated a mean value of 472 502 reads, among which 70.0% were obtained for the STR panel with 1.3% of the reads of this panel identified as stutters (see Tables 2 and 3). The total number of the reads of sample FM exceeded by 19.8% the value found for the same input DNA (1.0 ng) of control sample 2800 M. A variation (in reads/sample) of up to 34.5% has been already described in a

Table 2. NGS analysis of the STR panels for samples FM, FM-6, FM-8 and FM-10. Rea**ds**: overall number of reads obtained in the two tests. In bold characters, the reads assigned to markers with ≥ 50x coverage. In brackets, the overall number of markers with ≥ 50x coverage in the two tests. Concordance: r^2 values of the X-Y plots representing the concordance between the two replicate tests. DoC: depth of coverage/locus (reads/locus) calculated on markers with ≥ 50 reads. The mean values (± standard deviation) are reported in italic while the median values are in bold characters. The min/max values are in brackets. Low coverage: frequency of markers showing < 50 reads not considering the locus DXS8377 (*) (see also paragraph 3.2.2). Imbalance: frequency of heterozygous imbalanced loci. In brackets, the ratio between the reads of the two alleles (mean ± standard deviation). Drop out: frequency of drop outs for heterozygous markers. I.T.: frequency of reads below the interpretation threshold. Stutter: percentage of the reads attributed to stutter products out of the total number of reads. In brackets, the frequency of stutter products whose number of reads exceeded the defined thresholds (see Materials and Methods for details). Drop in: frequency of drop ins

Sample	Reads	Concordance	DoC	Low coverage	Imbalance	Drop out	I.T.	Stutter	Drop in
FM (ctrl)	661 780 **661 707** (124/126)	0.965	*5 336 ± 5 415* **3817** (143; 32 339)	0*	0.057 (0.562 ± 0.026)	0	0	1.27 % (0)	0
FM-6	131 734 **131 482** (69/126)	0.915	1 905 ± 2 873 **725** (51; 12 703)	0.436	0.323 (0.429 ± 0.112)	0.117	0.029	1.91 % (0)	0.014
FM-8	44 571 **44 492** (44/126)	0.807	*1 011 ± 1 328* **468** (54; 5 765)	0.634	0.440 (0.361 ± 0.112)	0.320	0	3.73 % (0.045)	0
FM-10	18 804 **18 734** (28/126)	0.627	*669 ± 802* **337** (56; 3 126)	0.761	0.277 (0.302 ± 0.120)	0.444	0.111	0.31 % (0.027)	0

Table 3. NGS analysis of the SNP panels obtained from samples FM, FM-6, FM-8 and FM-10. Reads: overall number of reads obtained in the two tests. In bold characters, the reads assigned to markers with ≥ 50x coverage. In brackets, the overall number of markers with ≥ 50x coverage in the two tests. Concordance: r^2 values of the X-Y plots representing the concordance between the two replicate tests. DoC: depth of coverage/locus (reads/locus) calculated on markers with ≥ 50 reads. The mean values (± standard deviation) are reported in italic while the median values are in bold characters. The min/max values are in brackets. Low coverage: frequency of markers showing < 50 reads, not considering the *rs2342747 and 430046* markers (*) (see paragraph 3.2.2.2). Imbalance: frequency of heterozygous imbalanced loci. In brackets, the ratio between the reads of the two alleles (mean ± standard deviation). Drop out: frequency of drop outs for heterozygous markers. I.T.: frequency of reads below the interpretation threshold

Sample	Reads	Concordance	DoC	Low coverage	Imbalance	Drop out	I.T.
FM (ctrl)	283 224 **283 155** (186/190)	0.904	*1 522 ± 1 044* **1 205** (128; 5 784)	0*	0.045 (0.477 ± 0.077)	0	0
FM-6	128 789 **128 129** (137/190)	0.977	*935 ± 1 865* **303** (54; 11 630)	0.257	0.326 (0.407 ± 0.135)	0.102	0.073
FM-8	66 647 **66 026** (92/190)	0.958	*718 ± 1 740* **216** (53; 11 518)	0.494	0.242 (0.357 ± 0.156)	0.242	0.054
FM-10	30 015 **29 322** (42/190)	0.949	*698 ± 1 400* **167** (63; 7 290)	0.757	0.166 (0.357 ± 0.018)	0.333	0.050

previous paper [31], for the same amount (1.0 ng) of two different DNA control samples (cell lines 2800 M and 9947A). Therefore, it is likely that the inter-sample variation found here could be ascribed to poor precision in the quantification of the samples as well as to intrinsic chemical features of the templates [31, 35]. Reads were generated for all markers of both panels, with low variability between the two replicates ($r^2 = 0.965$ and 0.904 for STR and SNP markers, respectively).

Very similar variations in DoC for STR markers were found between sample FM and the control DNA 2800 M. This emerged by the comparison of the relative DoC (*r*DoC) of each locus in the two samples (see Supporting Information Fig. S4). The mean ratio value ± s.d. was, in fact, 1.015 ± 0.412 (median value = 0.991; min = 0.078; max = 3.370). It is interesting to note that the markers characterized by the lowest ratios (DY456 and DXS8377) were removed from the commercial version of the kit, while marker DYS392, showing a ratio of 3.370, has been already described [31] as the most variable in terms of DoC among different samples. The frequency of imbalance of heterozygote STR loci was 0.057, for sample FM, similar to the one found for control sample 2800 M (0.036). Marker DYF387S1 showed a tri-allelic pattern (37/38/39) which is a frequent finding for this duplicated marker [41].

Also, very similar variations in DoC were found for SNP loci between samples FM and 2800 M (see Supporting Information Fig. S5). The mean ratio value ± s.d. was, in fact, 1.056 ± 0.274 (median value = 1.045; min = 0.005; max = 1.632). Also for this panel of markers, it is interesting to note that the markers characterized by the lowest ratios (*rs2342747* and *rs430046*) have been already described [31] as those most affected by artifacts. In addition, these two loci showed low coverage in all the replicates. For this reason, they were not further considered in this study. Allelic imbalance of SNP loci showed a frequency of 0.045 similar to the one (0.071) calculated for the same amount of DNA of sample 2800 M.

In conclusion, control DNA 2800 M and the high-molecular-weight DNA sample FM showed comparable features in terms of NGS typing results for both panels of STR and SNP markers.

3.2.2.3 Negative controls

No reads were scored for the STR panel in the two negative controls. A total of 14 and 18 reads were scored in the NTC and, respectively, the NCSP for the SNP panel. Below the AT, no more than 2 reads/locus were found.

3.2.3 Samples FM-6, FM-8 and FM-10

Quantification of degraded samples is a critical concern in forensic and ancient DNA analysis [13, 22]. Molecular quantification of such samples provides, in fact, only an assessment of the DNA amount "relative" to the qPCR probe selected [24]. In addition, in this study, the degradation index of samples FM-6, FM-8 and FM-10 could not been calculated thus supporting the evidence of a severe modification of the templates. For these reasons, we did not follow the recommendation of using 1.0 ng of genomic DNA as reported in the ForenSeq™ DNA Sequencing kit protocol [23] but rather decided to employ the maximum allowed DNA volume for the NGS reaction, corresponding to 5.0 μL (equivalent to 1.64–0.055 ng of DNA, as assessed by the SA probe).

The NGS results achieved for these degraded samples are shown in Tables 2 and 3, for STR and SNP markers, respectively. A significant decrease in the number of reads was evidenced in these damaged templates in comparison to the unmodified sample FM. This reduction was 35.2 and 9.4 folds lower compared to the control FM, for the STR and the SNP panels, respectively, and 10.2 and 2.9 folds lower than the 2800 M 0.1 ng dilution, for the STR and the SNP panels, respectively. Finally, the number of markers which produced reads in both replicates was dramatically reduced.

As expected, the mean DoC/locus decreased severely, especially for the STR loci, in relation to the increase in degradation of the templates. The most likely explanation is that the target enrichment procedure (performed by PCR) was inhibited by the molecular damage of the samples [14, 24, 42, 43]. This general trend is clearly described in Supporting Information Figs. S6 and S7 where the number of reads obtained for the three degraded FM samples and the unmodified FM control were reported and compared for each marker of the STR and SNP panels. An interesting finding was that NGS produced a higher number of reads in sample FM-6 than in control FM for three STR (namely, *DYS385A-B*, *DYS505* and *TPOX*) and ten SNP markers (*rs8037429, rs1413212, rs2269355, rs2040411, rs1493232, rs576261, rs1058083, rs1028528, rs221956* and *rs1490413*) out of the total number of 155 markers. A possible explanation for this result can be the DNA amount of sample FM-6 employed for the tests (1.68 ng as assessed by the SA). Nevertheless, other different mechanisms could be involved. Even in sample FM-10 (where 0.055 ng of DNA was used), in fact, markers *rs2269355* and *rs576261* gave higher DoCs than in control sample FM.

A possible explanation for this result is that the primer annealing sequences for the amplification of these loci could be less damaged than those of the other markers, being poorly represented in purinic compounds. For this reason, an unknown high number of intact annealing sequences, preferentially amplified in a general situation of damaged sequences, could be the explanation for the increase of the DoC in these markers. Even the nucleotide position of the purinic lesion within the primer sequence could affect the PCR efficiency. However, more studies are necessary to elucidate this finding.

As reported in Tables 2 and 3, the analysis of STR and SNP data showed a high percentage of loci which gave none or less than 50 reads in the three trial samples (up to 75.7–76.1% in sample FM-10) and a positive trend between the extent of DNA degradation and the % of low-covered loci. Although this setting of the cut-off value is arbitrary, and leads to a loss of genetic information of up to 5.7% (when compared to the cut-off of 30 x) [32], we believe that a conservative approach

focused on the reliability of the results is what is needed in forensic genetics when validation studies are underway to better understand the potential of an innovative technique.

Last, the percentage of low-covered loci was higher in the STR panels that in the SNP ones, for samples FM-6 and FM-8 (43.6% versus 25.7%, in FM-6 and 63.4% versus 49.4%, in FM-8).

The allelic imbalance varied from 0.277 to 0.440 and from 0.166 to 0.326, for STR and SNP markers, respectively. In 70.3% of the STR markers, the allele with higher molecular weight was less amplified.

Allelic reads below the 30 reads interpretation threshold (IT) were scored for loci with a DoC between 50 and 650 reads. In detail, three STR and seventeen SNP markers showed allelic reads below the IT in the trial DNA samples.

Allelic drop outs were seen for both STR and SNP markers. As expected, the frequency of these artifacts increased proportionally to the extent of degradation in the trial samples, reaching the highest frequency in the most degraded sample FM-10 (up to 44.4% and 33.3% for STRs and SNPs, respectively). In addition, the longer allele of the heterozygous genotype dropped out in 60.8% of the cases, for STR markers. In Fig. 1, the number of reads of the surviving alleles for the dropped out genotypes are shown, subdivided in four intervals. Interestingly, allelic drop out occurred even for STR markers characterized by a DoC $\geq$ 601x, reaching the considerable coverage of 1445, for allele 7 of the TH01 locus, in sample FM-8. For SNP markers this misleading result was scored only for loci with DoCs $\leq$ 244x. These results highlight that a potentially misleading amplification artifact such as an allelic dropout could be a crucial concern even for NGS technology especially when the analyses are performed on highly degraded samples.

Stutter products above the defined setting parameters do not constitute an important artifact, as they were found only in three cases. However, this spurious PCR product could originate potentially misleading results. In fact, the NGS results of locus D17S10301 for sample FM-8 showed a stutter product (allele 11) whose number of reads (326) was even higher than the one of the expected real genotype 12/12 (75 reads). Again, the NGS results of locus TH01 (expected real genotype 7/9.3) for sample FM-10 showed instead allele 6 with a 35x coverage (and then above the interpretation threshold), the surviving allele 7 with a 265x coverage and allele 9.3 dropped out.

A single PCR artifact identifiable as "drop in" was observed in the damaged trial samples. This occurred at locus D22S1045 of sample FM-6 which even showed a genotype characterized by alleles 16 (35x) and 17 (57x) instead of the expected true genotype 11/12.

In conclusion, these results clearly point out the need for replicates, at least in two different experiments carried out under the same analytical conditions, to prevent the risk of mistyping, especially for poorly covered loci.

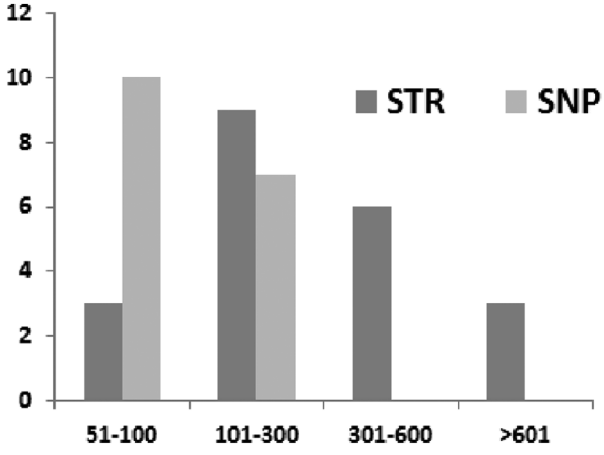

Figure 1. Number of surviving alleles recorded in the trial samples for the STR and SNP panels. X-axis: number of reads subdivided in four categories; Y-axis: number of observations.

3.2.4 Dilutions of sample FM-6

The NGS analyses of two dilutions of 5.0 μL each of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3), are reported in Supporting Information Tables S2 and S3 (for STR and SNP markers, respectively). A trivial result was that the total mean number of reads/sample and the mean DoC/locus were always lower than the ones described for sample FM-6. The frequency of loci with DoC $<$ 50x increased dramatically reaching the values of 0.587 and 0.594 for STRs and SNPs, respectively, for the FM-6 dilution 1:3.

Several allelic reads below the IT were scored. This last finding was observed for loci with a low depth of coverage (mean 99x $\pm$ 38x for STRs and 65x $\pm$ 19x for SNPs). In two cases, spurious reads substituted the dropped out allele of a heterozygous locus (D2S441 and D3S1358). Other spurious reads were sporadically generated for both STR and SNP markers. Allelic drop outs were observed for both STR and SNP markers. The DoC of the surviving allele was $\leq$234x and $\leq$74x, for STR and SNP markers, respectively.

An interesting amplification artifact was also found at locus D7S820 (expected genotype 8/8) of dilution FM-6 (1:1). In detail, the NGS approach displayed 93 reads of the expected $(GATA)_8$ repeat sequence and 38 reads of a mutated $(AATA)(GATA)_7$ one. The molecular origin of this repeat sequence variation is only presumptive but it is well known that the presence of A-P sites can promote preferential and unspecific incorporation of Adenine [43]. In addition, another possible explanation for this finding is that deamination of the Cytosine to Uracil can lead to G$\rightarrow$A transitions [44] as the one observed for the first base of the tetrameric repeat.

In conclusion, even the NGS data results obtained from two dilutions of the less damaged trial sample FM-6 are consistent in identifying the risk of mistyping when the molecular data of poorly covered loci are used for personal identification purposes, in particular if a single NGS experiment is performed.

3.2.5 Genetic typing of the samples

The ability of the NGS approach in recovering reliable genetic results from artificially degraded samples was checked by comparing the genetic typing obtained for samples FM-6, FM-8, FM-10 with the corresponding reference DNA FM. Fifty-eight STRs and ninety-three SNPs were the bulk set of markers for the comparisons.

The interpretation guidelines for genotyping the NGS data obtained from the two replicate analyses of each sample are based on the "*consensus*" approach [24, 33] and have been already described in paragraph 2.3.4. The concordance of results among the replicates of the degraded samples and the final comparison with the expected genotypes of reference sample FM allowed us to assign each STR and SNP genotype to one of the following four categories: correct (C), no result (NR), unreliable (U) and wrong (W) result. Figures 2A and 2B show the distribution of the frequencies of each category, for STR and SNP markers respectively, obtained from the analysis of the three trial samples.

The frequency of correctly typed loci was inversely related (r^2 from 0.971 to 0.999) to the degree of chemical damage of the templates (see Supporting Information Fig. S8). As expected, the less damaged sample FM-6 showed the highest frequency of correctly typed loci (44.8 and 58.7% for the STR and the SNP panels, respectively) while the most degraded sample FM-10 revealed the lowest number of markers successfully typed (5.1 and 10.9%, for the STR and the SNP panels, respectively). All these data together indicate that the SNPs panel is two times more efficient than the STR panel in providing correct genotypes, at least for the extremely degraded sample FM-10.

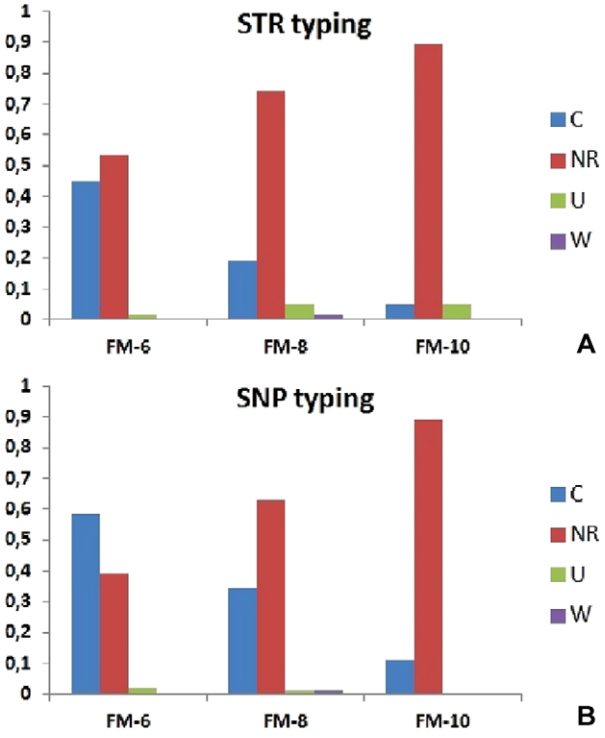

Figure 2. 2A and 2B. Results of the genetic typing for trial samples, according to the interpretation guidelines described in paragraph 2.3.4 (C: correct typing; NR: no result; U: unreliable typing; W: wrong typing). A: STR panel; B: SNP panel. X-axis: samples FM-6, FM-8 and FM-10; Y-axis: frequency of each genotyping category.

In Figs. 3 and 4 are represented the STR and SNP markers which showed correct results for the degraded samples FM-6, FM-8, FM-10, and their relative amplicon sizes. No correlation between the amplified markers and the corresponding amplicon size was found, at least for STR loci. Markers DYS505 and DYS385A-B were correctly typed for the most degraded sample FM-10, in fact, even though they are characterized by quite different amplicon sizes (174 and 335 bp, respectively). Differently, for the most degraded sample FM-10, the correctly typed SNPs seem to cluster on the left side of Fig. 4, characterized by loci with short amplicon sizes.

The STR and SNP markers correctly typed in sample FM-10 also gave the expected correct genotypes in the less damaged samples FM-6 and FM-8, as shown in Figs. 3 and 4.

A positive trend between the number of loci which provided "no result" and the extent of degradation of the samples was also observed. The frequency of the loci which gave "unreliable results", that is different typing results in the two replicates, was higher for the STR panel, reaching the value of 5.2% for sample FM-10 (for the SNP panel no more than 2.2% of the markers gave this outcome) (see Fig. 2A and B).

According to the conservative criteria described in interpretation guidelines for genotyping, only two loci gave incorrect ("wrong") results in both replicate tests, when compared to the expected genotypes, thus showing that a cut-off value of 50x eliminated genotyping errors in samples FM-6 and FM-10. The STR locus TH01 typing gave in both replicates of sample FM-8 only the reads of allele 7 (1,445x in the first test and 917x in the second one) instead of the expected genotype 7/9.3. The SNPs typing of locus rs722290 (expected genotype C/G) gave in both replicates of sample FM-8 only the reads of allele G (64x in the first test and 244x in the second one). This point can be extremely relevant when forensic casework will be routinely approached with the NGS technology.

The genetic typing of the two dilutions of sample FM-6, namely FM-6 (1:1) and FM-6 (1:3), showed that an increased dilution correlates negatively to the number of correctly typed loci. In addition, it was observed that the most diluted sample FM-6 (1:3) provided the highest frequencies of genotypes assigned to the category of "*no result*", according to the criteria described in paragraph 2.3.4 (see Supporting Information Fig. S9A and B). No incorrect genotype was scored while the frequency of "unreliable results" was $\leq$0.017. In Supporting Information Figs. S10 and S11 are represented the STR and SNP markers which showed correct results for the diluted samples FM-6 (1:1) and FM-6 (1:3). As already observed for the degraded samples FM-6, FM-8, FM-10, only a limited number of markers gave a good coverage thus allowing a correct typing in the diluted samples.

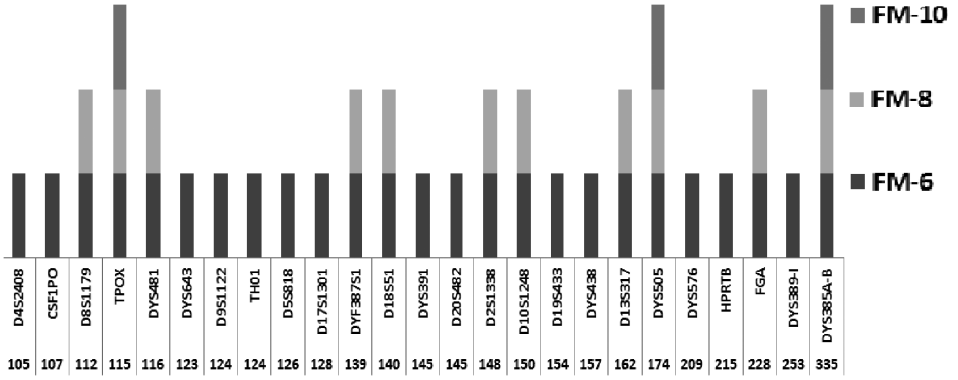

Figure 3. STR markers correctly typed for trial sample FM-6, FM-8 and FM-10. The STRs are sorted by the mean molecular weight of the corresponding amplicons.

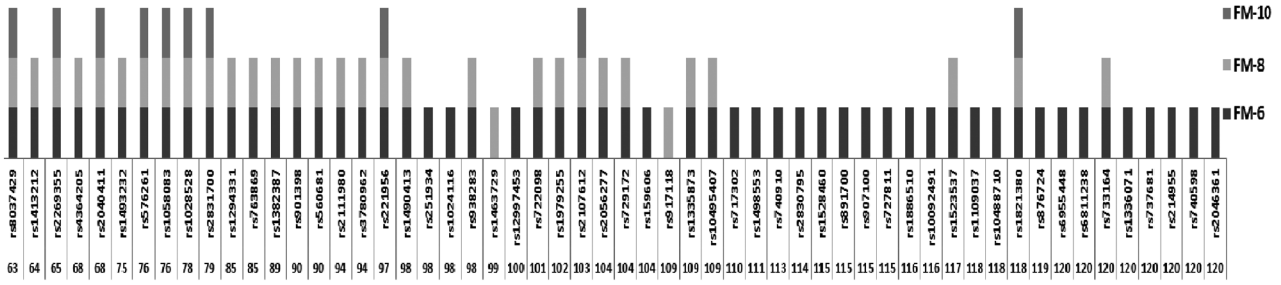

Figure 4. SNP markers correctly typed for trial sample FM-6, FM-8 and FM-10. The SNPs are sorted by the molecular weight of the corresponding amplicons.

3.3 STR Analysis by CE

The same amounts of DNA utilized for the NGS analysis of degraded samples FM-6, FM-8 and FM-10, were used also for the conventional STR typing approach amplifying the 16 STR loci plus Amelogenin contained in the commercial kit PowerPlex® ESI 17 system, in duplicate tests. The results showed all the peculiar features of degraded/modified DNA samples, characterized by partial profiles with loss of high-molecular-weight markers and the presence of amplification artifacts such as allelic imbalances and drop outs (data not shown). The STR typing of sample FM-6 allowed the correct genetic typing of 8 out of 16 STRs (50%) plus Amelogenin, while the number of correctly scored STRs was 5 out of 16 (31.5%) plus Amelogenin, for sample FM-8. The conventional PCR-CE analysis provided reliable results from three low molecular weight loci (D3S1358, D16S539 and vWA) which gave no result by employing NGS-based tests. More interestingly, PCR-CE analysis gave the correct genotype (7/9.3) at locus sample TH01 of sample FM-8 while NGS provided the incorrect 7/7 genotype (see Supporting Information Table S4). No marker could be typed for sample FM-10. No amplicon was found in the negative controls (NTC and NCSP).

4 Concluding remarks

The results reported in this study are the first data gathered on the reliability of the NGS analysis using the ForenSeq™ DNA Sequencing kit on degraded samples. At present, only a single study [31] approached challenging DNA samples but this study reported the results focusing mainly on the concordance of the data with the conventional PCR/CE methods. In the present paper, a DNA sample (namely sample FM) was artificially enriched in A-P sites, thus producing three differently degraded trial DNA samples (FM-6, FM-8 and FM-10). Even if the limit of this assay is that the DNA samples were enriched in specific lesions [21, 24, 35, 36], despite the large number of different DNA damages which can be found in real case-work samples [13], we believe that this approach can confidently represent a model to study the performance and the limits of this new NGS technology in forensic genetics. The results of the three degraded samples were then compared with those of the undegraded sample FM and those of three dilutions (1.0, 0.5 and 0.1 ng) of reference sample 2800 M. Finally, also conventional PCR/CE analysis was performed on the three trial samples as well as on the undegraded control.

In order to increase the number of reads/sample, the libraries of a reduced number of samples (twenty) were pooled together in this study, a strategy suggested to maximize the depth of coverage when challenging samples are considered [23, 37]. However, despite this approach, the number of reads obtained from the three trial DNA samples was 2.9–10.2 folds lower than the ones reported for the less concentrated control DNA 2800 M dilution (0.1 ng). This observation is consistent with findings already reported for other difficult samples [31]. In addition, it was impossible to assign up to 78.2% of the genotypes in the trial samples,

as the corresponding loci were flagged by the software as "low-covered" markers (< 50x). Both the decrease of the number of reads/sample and the increase of the percentage of low-covered loci were proportional to the extent of the degradation in each trial DNA sample. Other analytical parameters, such as the number of PCR cycles of the target enrichment procedure and library preparation should be evaluated in a future prospective in order to optimize the NGS approach to highly degraded samples [31]. In addition, since our data highlighted higher amplification efficiency for a specific set of STR and SNP markers in the trial samples, it could be interesting to investigate the amplification efficiency of these markers in other difficult DNA samples, potentially driving toward the selection of the best set of markers for forensic applications.

The ForenSeq™ DNA Sequencing kit run on the Illumina MiSeq was able to generate data which led to the correct typing of 44.8%, 18.9%, 5.1% of the STRs and 58.7%, 34.7%, 10.9% of the SNPs of trial samples FM-6, FM-8 and FM-10, respectively. These data support the common sense conclusion that the ability to recover reliable genotypes is inversely proportional to the extent of degradation of the samples. However, the SNP markers showed in all trial samples, that they had higher chances to be typed correctly compared to the STRs.

This NGS approach seems to offer very promising performances on heavily degraded DNA samples. In fact, while the conventional approach (PCR/CE) on the most degraded trial sample (sample FM-10) gave no results, three STRs (one autosomal and two Y-chromosome STRs) and ten autosomal SNPs gave reportable genotypes with this NGS assay (see Figs. 3 and 4). The 1/random match probability value (that is the LR) calculated for the autosomal loci [45] gave a result of 1.7×10^5, which can be considered an excellent achievement, for such a highly damaged DNA sample.

The NGS assay showed very reliable results. The frequency of the genetic mistyping was very low, reaching the value of 1.4% only for sample FM-8. It is interesting to note that, in two cases of erroneous typing, the same allele dropped out in the two replicates while the surviving alleles showed a considerable number of reads, far above the interpretation threshold value (1,445x and 917x for allele 7 of TH01 locus; 64x and 244x for allele G of rs722090 locus). This finding points out that even for NGS the risk of mistyping (that is, in these two cases, to type as homozygous a genotype instead of the correct heterozygous one) could not be eliminated by duplicate tests.

In addition, our results point out that validation studies and the definition of standard operating procedures for the interpretation of the NGS DNA typing results are needed in order to evaluate carefully the application of this technology in forensic human identification even if very promising results have been already achieved [12].

Special thanks to Dr. F. Tomasella, Department of Transfusional Medicine, A-U-O of Trieste, Italy, for providing the sample FM, and to Dr. M. Vatta, PhD, FACMG, and A. Cochrane, MPH, Department of Medical and Molecular Genetics Department of Medicine Indiana University School of Medicine Indianapolis, IN, 46202, for their invaluable help in the revision of the manuscript. B.B. is a fellow of the Genetics, Molecular and Cellular Biology PhD program of the University of Pavia.

The authores have declared no conflict of interest.

5 References

[1] Welch, L., Gill, P., Phillips, C., Ansell, R., Morling, N., Parson, W., Palo, J.U., Bastisch, I., *Forensic Sci. Int. Genet.* 2012, *6*, 819–826.

[2] Hares, D. R., *Forensic Sci. Int. Genet.* 2012, *6*, e52–e54.

[3] Roewer, L., *Investig. Genet.* 2013, *4*, 22–30.

[4] Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F., Rothberg, J. M., *Nature* 2005, *437*, 376–380.

[5] Metzker, M. L., *Nat. Rev. Genet.* 2010, *11*, 31–46.

[6] Korlach, J., Bjornson, K. P., Chaudhuri, B. P., Cicero, R. L., Flusberg, B. A., Gray, J. J., Holden, D., Saxena, R., Wegener, J., Turner, S. W., *Methods Enzymol.* 2010, *472*, 431–455.

[7] Lam, H. Y., Clark, M. J., Chen, R., Natsoulis, G., O'Huallachain, M., Dewey, F. E., Habegger, L., Ashley, E. A., Gerstein, M. B., Butte, A. J., Ji, H. P., Snyder, M., *Nat. Biotechnol.* 2012, *30*, 78–82.

[8] Haque, F., Li, J., Wu, H. C., Liang, H. J., Guo, P., *Nano Today* 2013, *8*, 56–74.

[9] Capelli, C., Tschentscher, F., Pascali, V. L., *Forensic Sci. Int.* 2003, *131*, 59–64.

[10] Orlando, L., Gilbert, M. T., Willerslev, E., *Nat. Rev. Genet.* 2015, *16*, 395–408.

[11] Schaefer, N. K., Shapiro, B., Green, R. E., *Mol. Ecol.* 2016, *25*, 2398–2412.

[12] Børsting, C., Morling, N., *Forensic Sci. Int. Genet.* 2015, *18*, 78–89.

[13] Alaeddini, R., Walsh, S. J., Abbas, A., *Forensic Sci. Int. Genet.* 2010, *4*, 148–157.

[14] Eckert, K. A., Kunkel, T. A., *PCR Methods Appl.* 1991, *1*, 17–24.

[15] Forster, L., Thomson, J., Kutranov, S., *Forensic Sci. Int. Genet.* 2008, *2*, 318–328.

[16] Butler, J. M., Shen, Y., McCord, B. R., *J. Forensic Sci.* 2003, *48*, 1054–1064.

[17] Coble, M. D., Butler, J. M., *J. Forensic Sci.* 2005, *50*, 43–53.

[18] Smith, P. J., Ballantyne, J., *J. Forensic Sci.* 2007, *52*, 820–829.

[19] Diegoli, T. M., Farr, M., Cromartie, C., Coble, M. D., Bille, T. W., *Forensic Sci. Int. Genet.* 2012, *6*, 498–503.

[20] Gettings, K. B., Kiesler, K. M., Vallone, P. M., *Forensic Sci. Int. Genet.* 2015, *19*, 1–9.

[21] Lindahl, T., Nyberg, B., *Biochemistry.* 1972, *11*, 3610–3618.

[22] Pääbo, S., Poinar, H., Serre, D., Jeenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., *Ann. Rev. Genet.* 2004, *38*, 645–679.

[23] Illumina, ForenSeq™ DNA Signature Prep Guide, Part # 15049528 Rev. D, February 2015.

[24] Fattorini, P., Previderè, C., Sorçaburu Cigliero, S., Marrubini, G., Alù, M., Barbaro, A. M., Carnevali, E., Carracedo, A., Casarino, L., Consoloni, L., Corato, S., Domenici, R., Fabbri, M., Giardina, E., Grignani, P., Baldassarra, S. L., Moratti, M., Nicolin, V., Pelotti, S., Piccinini, A., Pitacco, P., Plizza, L., Resta, N., Ricci, U., Robino, C., Salvaderi, L., Scarnicci, F., Schneider, P. M., Seidita, G., Trizzino, L., Turchi, C., Turrina, S., Vatta, P., Vecchiotti, C., Verzeletti, A., De Stefano, F., *Electrophoresis.* 2014, *35*, 3134–3144.

[25] Illumina, CASAVA v1.8.2 User Guide, December 2011.

[26] Illumina, ForenSeq Universal Analysis Software Guide, Part # 15053876 Rev. B, February 2015.

[27] Ajay, S. S., Parker, S. C., Abaan, H. O., Fajardo, K. V., Margulies, E. H., *Genome Res.* 2011, *21*, 1498–1505.

[28] Eduardoff, M., Santos, C., de la Puerte, M., Gross, T. E., Fondevila M., Strobl, C., Sobrino, B., Ballard, D., Schneider, P. M., Carracedo, Á., Lareu, M. V., Parson, W., Phillips, C., *Forensic Sci. Int. Genet.* 2015, *17*, 110–121.

[29] Gettings, K. B., Kiesler, K. M., Faith, S. A., Montano, E., Baker C. H., Young, B. A., Guerrieri, R. A., Vallone, P. M., *Forensic Sci. Int. Genet.* 2016, *21*, 15–21.

[30] Goodwing, S., McPeron, J. D., McCombie, W. R., *Nature Rev. Genet.* 2016, *17*, 333–351

[31] Churchill, J. D., Schmedes, S. E., King, J. L., Budowle, B., *Forensic Sci. Int. Genet.* 2016, *20*, 20–29.

[32] Carboni, I., Fattorini, P., Previderè, C., Sorçaburu Ciglieri, S., Iozzi, S., Nutini, A. L., Contini, E., Pescucci, C., Torricelli, F., Ricci, U., *Forensic Sci. Int. Genet.* 2015, *Suppl. Series 5*, e83–e85.

[33] Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, S. Escaravage, N., Waits, L. P., Bouvet, J., *Nucleic Acids Res.* 1996, *24*, 3189–3194.

[34] Warshauer, D. H., King, J. L., Budowle, B., *Forensic Sci. Int. Genet.* 2015, *14*, 182–186.

[35] Lindahl, T., Andersson, A., *Biochemistry* 1972, *11*, 3618–3623.

[36] Lindahl, T., Nyberg, B., *Biochemistry* 1974, *13*, 3405–3410.

[37] Sims, D., Sudbery, I., Ilott, N. E., Heger, A., Ponting, A. P., *Nature Rev. Genet.* 2014, *15*, 121–132.

[38] Sommer, S., Courtiol, A., Mazzoni, C. J., *BMC Genomics* 2013, *14*, 542–552.

[39] Rockenbauer, E., Hansen, S., Mikkelsen, M., Børsting, C., Morling, N., *Forensic Sci. Int. Genet.* 2014, *8*, 68–72.

[40] Fordyce, S. L., Mogensen, H. S., Børsting, C., Lagace, R. E., Chang, C. W., Rajagopalan, N., Morling, N., *Forensic Sci. Int. Genet.* 2015, *14*, 132–140.

[41] Ballantyne, K. N., Goedbloed, M., Fang, R., Shaap, O., Lao, O. Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vézina, H., Knoblauch, H., Lessig, R., Roewer, L., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M. R., Kayser, M., *Am. J. Hum. Genet.* 2010, *87*, 341–353.

[42] Brisco, M. J., Latham, S., Bartley, P. A., Morley, A. A., *BioTechniques* 2010, *49*, 893–897.

[43] Heyn, P., Stenzel, U., Briggs, A. W., Kircher, M., Hofreiter, M., Meyer, M., *Nucleic Acids Res.* 2010, *38*, e161.

[44] Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., Pääbo, S., *Nucleic Acids Res.* 2001, *29*, 4793–4799.

[45] The ALlele FREquency Database, https://alfred.med.yale.edu.