



# Inter-laboratory study on standardized MPS libraries: evaluation of performance, concordance, and sensitivity using mixtures and degraded DNA

Petra Müller<sup>1</sup> · Christian Sell<sup>2</sup> · Thorsten Hadrys<sup>3</sup> · Johannes Hedman<sup>4,5</sup> · Steffi Bredemeyer<sup>6</sup> · Francois-Xavier Laurent<sup>7</sup> · Lutz Roewer<sup>6</sup> · Sabrina Achtruth<sup>8</sup> · Maja Sidstedt<sup>4,5</sup> · Titia Sijen<sup>9</sup> · Marc Trimborn<sup>8</sup> · Natalie Weiler<sup>9</sup> · Sascha Willuweit<sup>6</sup> · Ingo Bastisch<sup>2</sup> · Walther Parson<sup>1,10</sup>  · the SeqForSTR-Consortium

Received: 4 September 2019 / Accepted: 29 October 2019 / Published online: 19 November 2019  
© The Author(s) 2019

## Abstract

We present results from an inter-laboratory massively parallel sequencing (MPS) study in the framework of the SeqForSTRs project to evaluate forensically relevant parameters, such as performance, concordance, and sensitivity, using a standardized sequencing library including reference material, mixtures, and ancient DNA samples. The standardized library was prepared using the ForenSeq DNA Signature Prep Kit (primer mix A). The library was shared between eight European laboratories located in Austria, France, Germany, The Netherlands, and Sweden to perform MPS on their particular MiSeq FGx sequencers. Despite variation in performance between sequencing runs, all laboratories obtained quality metrics that fell within the manufacturer's recommended ranges. Furthermore, differences in locus coverage did not inevitably adversely affect heterozygous balance. Inter-laboratory concordance showed 100% concordant genotypes for the included autosomal and Y-STRs, and still, X-STR concordance exceeded 83%. The exclusive reasons for X-STR discordances were drop-outs at DXS10103. Sensitivity experiments demonstrated that correct allele calling varied between sequencing instruments in particular for lower DNA amounts ( $\leq 125$  pg). The analysis of compromised DNA samples showed the drop-out of one sample (FA10013B01A) while for the remaining three degraded DNA samples MPS was able to successfully type  $\geq 87\%$  of all aSTRs,  $\geq 78\%$  of all Y-STRs,  $\geq 68\%$  of all X-STRs, and  $\geq 92\%$  of all iSNPs demonstrating that MPS is a promising tool for human identity testing, which in return, has to undergo rigorous in-house validation before it can be implemented into forensic routine casework.

**Keywords** Short tandem repeats · Massively parallel sequencing · ForenSeq DNA Signature Prep Kit · Inter-laboratory study · Collaborative study

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00414-019-02201-2>) contains supplementary material, which is available to authorized users.

---

✉ Walther Parson  
walther.parson@i-med.ac.at

<sup>1</sup> Institute of Legal Medicine, Medical University of Innsbruck, Müllerstraße 44, 6020 Innsbruck, Austria

<sup>2</sup> Federal Criminal Police Office, Wiesbaden, Germany

<sup>3</sup> Institute of Forensic Sciences, DNA Department, Bavarian State Criminal Police Office, Munich, Germany

<sup>4</sup> Swedish National Forensic Centre (NFC), Linköping, Sweden

<sup>5</sup> Applied Microbiology, Department of Chemistry, Lund University, Lund, Sweden

<sup>6</sup> Institute of Legal Medicine and Forensic Sciences, Charité – Universitätsmedizin Berlin, Berlin, Germany

<sup>7</sup> Institut National de Police Scientifique, Laboratoire de Police Scientifique de Lyon, Ecully Cedex, France

<sup>8</sup> The Police President in Berlin, Forensic Science Institute, Berlin, Germany

<sup>9</sup> Biological Traces, Netherlands Forensic Institute, Laan van Ypenburg 6, 2497 GB The Hague, The Netherlands

<sup>10</sup> Forensic Science Program, The Pennsylvania State University, State College, PA, USA

## Introduction

With the emergence of massively parallel sequencing (MPS) technologies, molecular genetic tools have been developed to characterise the nucleotide sequence of short tandem repeat (STR) markers that have so far been analysed using fragment sizing by capillary electrophoresis (CE) [1–7]. As is common practise in forensic genetics, such novel tools undergo detailed evaluation [3, 8–12] and validation [1, 13, 14] prior to their application in casework. Recent population studies have shown that sequencing of STRs lead to an increased power of discrimination compared to commonly used CE sizing by a) illustrating micro-variant structures and b) identifying sequence variations located in the repeat- (isometric variants; alleles of identical size but different in internal sequence) or flanking region [15–20]. Moreover, current MPS-kits provide the option to multiplex various loci simultaneously within one reaction, such as STRs (autosomal, Y- and X-chromosomal), mitochondrial DNA (control region) as well as single nucleotide polymorphisms (SNPs) that provide estimates of identity, phenotypic traits, biogeographical and ancestry information [21–24]. The inclusion of SNPs adds the benefit [25–27] that these markers allow for shorter amplicon design, which can support human identification of degraded or challenging DNA samples [28–30]. Although published studies have demonstrated that MPS is a promising tool for forensic applications, little is known about the method's variability and potential differences in the limit of detection between sequencing platforms from the same supplier. Inter-laboratory studies are however highly relevant to understand the performance of this new technology when applied to real casework settings. Since the workflow for MPS-based sequence analysis of forensic genetic markers is more complex than for electrophoresis-based systems, we decided to reduce the complexity of this study to libraries prepared at the organising laboratory (OL) from forensically relevant samples that were shipped to the participating laboratories in the following way: In the framework of the SeqForSTRs project (EU Project ISF-Nr. IZ25-5793-2015-30 2017-2020, [31]) the ForenSeq DNA Signature Prep Kit [22] was used for collaborative validation experiments and population studies. As a general benchmark for the comparability of the results between laboratories, and, to reduce the inherent complexity of MPS experiments derived from multiple steps during manual library preparation, the SeqForSTRs consortium performed a collaborative exercise analysing a centrally prepared, standardized sequencing library. The library was distributed between eight participating laboratories located in Austria, France, Germany, The Netherlands, and Sweden and were analysed on their respective MiSeq FGx sequencing platforms (Illumina, San Diego, USA;

[32]). In order to monitor the effect of shipment, one library aliquot was sent to one of the participants, immediately returned to the OL and analysed (run 8) to compare the results to those of the test run prior to shipment (run 0).

We note that only STR genotype calls were considered here in agreement with the aim of the SeqForSTRs project, except for ancient DNA samples, where also SNP information was analysed to evaluate their performance in highly degraded DNA. To rate the performance of iSNPs in compromised samples additional SNP information was examined for standard reference material 2391c (components A–C). The library contained selected DNA samples (control DNAs, reference material, mixtures and degraded DNA) to test for instrument variability, concordance, and sensitivity.

## Materials and methods

This collaborative study combined data generated in nine sequencing runs that were conducted in eight laboratories located in Austria, France, Germany, The Netherlands, and Sweden (Table 1). To ensure uniform starting conditions a standardized sequencing library consisting of 25 selected DNA samples, a negative and a positive amplification control (see “Sample selection” section) was prepared using the ForenSeq DNA Signature Prep Kit (Verogen, San Diego, USA) at the OL. To assess the quality of sequencing products an initial test run (run 0) was performed at the OL (Table 1) prior to shipment. Pooled sequencing libraries were shipped with the United Parcel Service (unchilled and under ambient conditions) and sequenced by all participants using their respective MiSeq FGx sequencers (Illumina). To monitor transport conditions (under ambient temperatures) one aliquot of the library was shipped by the OL to one of the participants, returned to the OL upon receipt and analysed (run 8).

### Sample selection

#### Concordance, mixture, and sensitivity

Concordance was assessed using three single donor standard reference materials (SRM): 2391c, component A–C [33] purchased from the National Institute of Standards and Technology (NIST; Gaithersburg, USA), control DNA 2800M (Promega, Madison, USA), and two samples from previous GEDNAP (German DNA Profiling; <https://www.gednap.org/>) proficiency tests (G49-S4, single source; G49-S1, two-person mixture).

Two-person mixtures were prepared using control DNAs 2800M (major contributor; Promega) and 9947A (minor contributor; Thermo Fisher Scientific (TFS), Waltham, USA) at a stock concentration of 9 ng/μL and five ratios (M500 = 1:1; M166 = 83.3:16.7; M91 = 90.9:9.1; M62.5 = 93.7:6.3 and

**Table 1** Run and quality metrics information for all sequencing runs performed within the SeqforSTRs study. The resulting data were analysed and presented in anonymous form. An initial test run (run 0) was performed at the organising laboratory to assess the quality of

sequencing products. Prior to sequencing each laboratory loaded 7  $\mu\text{L}$  of pooled library on the MiSeq FGx instrument according to the manufacturer's recommendation (Verogen)

Run	Cluster density (K/mm <sup>2</sup> )	Cluster passing filter (%)	Phasing (%)	Pre-phasing (%)	Total no. of reads (performance testing) Single source samples	UAS version	NTC
UAS Guide (range)	400–1650 K/mm <sup>2</sup>	$\geq 80\%$	$\leq 0.25\%$	$\leq 0.15\%$			
0	706	96.13	0.141	0.102	704,959	1.2.16337	Negative
1	684	97.71	0.180	0.099	847,323	1.2.16337	Negative
2	432	97.25	0.168	0.144	438,786	1.2.16337	Negative
3	818	94.79	0.171	0.103	948,395	1.2.16173	Negative
4	603	97.65	0.193	0.084	641,752	1.2.16337	Negative
5	587	96.19	0.166	0.113	579,248	1.2.16337	Negative
6	486	97.76	0.226	0.102	650,416	1.2.16173	Negative
7	862	93.86	0.181	0.100	1,123,576	1.2.16337	Negative
8	625	97.03	0.192	0.130	729,121	1.2.16337	Negative
Mean	645	96.49	0.185	0.109	740,397		
SD	141	1.38	0.020	0.018	205,448		
Median	614	97.03	0.181	0.102	689,769		
Min	432	93.86	0.141	0.084	438,786		
Max	862	97.76	0.226	0.144	1,123,576		

M47.6 = 95.2:4.8). The initial intention was to prepare the mixtures with a male minor contributor however, the female fraction (9947A) was used as minor contributor mistakenly, which affected the ability to test male minor contributions in the mixture analyses. All mixtures were diluted accordingly to 1 ng/ $\mu\text{L}$  in molecular grade water prior to library preparation. Mixture study consisted of M166, M91, M62.5 and M47.6, which were analysed as singletons only. Sensitivity was assessed using mixture M500 prepared to a final DNA input of 1000 pg, 500 pg, 250 pg, 125 pg, 62 pg and 31 pg in molecular grade water and analysed in duplicate.

### Ancient DNA samples

To determine the robustness and performance of large MPS panels on highly degraded DNA, four ancient DNA (aDNA) samples were included. Libraries were prepared from extracted DNA from two bone (femur: FA10013B01A, humerus: FA10026B01A) and two tooth samples (molar: FA10030T01A, incisor: FA10058T01B) according to [22], respectively. Ancient samples dated between the fifth to eighth century ([34]) and were taken from earlier studies [35].

### DNA quantification

The amount of genomic DNA was determined with a real-time PCR assay targeting specific AluYb8 sequences [36]. A spiked in vitro mutagenized and cloned part of the *human retinoblastoma susceptibility protein 1 (RBI)* gene was co-

amplified as internal amplification positive control (pRB1<sub>IPC</sub>) according to [37], updated in [35]. Calibration curve analysis covered a DNA input range from 10 ng to 169.5 fg per reaction and was analysed in duplicate. The reaction volume was 10  $\mu\text{L}$  consisting of 5  $\mu\text{L}$  TaqMan Fast Universal PCR mix (TFS), 3  $\mu\text{L}$  primer probe premix (made in-house) and 2  $\mu\text{L}$  extracted DNA sample. Thermal cycler protocol consisted of an initial denaturation at 95 °C for 20 s followed by 40 cycles of denaturation at 95 °C for 3 s and annealing/elongation at 60 °C for 30 s. Samples were run on an Applied Biosystems 7500 Fast Real-Time PCR Instrument (TFS) using the HID Real-Time PCR Software v 2.3. Kinetic information for the pRB1<sub>IPC</sub> system yielded no indication of inhibition during DNA amplification.

### Library preparation and sequencing

Each laboratory obtained one 2 mL screw-cap tube (Sarstedt, Nümbrecht, Germany; labelled: *SeqForSTRs*, *SeqPool 20  $\mu\text{L}$* , *ForenSeq Kit*, 29.11.2017 PM) containing 20  $\mu\text{L}$  pooled sequencing library. The final library pool was made of 25 selected DNAs, a negative and a positive amplification control. To ensure sufficient library volume OL simultaneously prepared three library pools, including the same samples and index adapters (i5 and i7) using the ForenSeq DNA Signature Prep Kit, primer mix A (Verogen, [22]) in one 96-well plate (Table S1). Library preparation, purification and normalization were performed according to [22]. All samples were amplified with 1 ng DNA according to [22], except for

G49-S4 (2.6 ng DNA), the samples for the sensitivity study (serial dilution: 1000 pg–31 pg DNA) and aDNA samples.

Prior to pooling, 30  $\mu$ L of each library sample (present in pools 1, 2 and 3; Table S1) was joined into a 0.8 mL deep-well storage plate (one sample/well). Libraries were pooled, diluted and denatured following [22] before loading into a MiSeq FGx Reagent Cartridge and sequenced on the MiSeq FGx instrument ([32], Table S2). To achieve uniform designation for all sequencing runs OL provided a text-based sample sheet including relevant information on, e.g. sample name as well as adapter combination needed for demultiplexing and data analysis (Table S3).

## Data analysis

### Capillary electrophoresis

STR typing was performed at OL with PowerPlex ESX 17 (Promega [38]), PowerPlex 16 System (Promega [39]), PowerPlex 21 System (Promega [40]), Investigator Argus X-12 (Qiagen, Hilden, Germany [41]), AmpFISTR Yfiler (TFS [42]), Genderplex (an in-house developed sex-typing assay [35, 43]) or AmpFISTR NGM SElect Kit (TFS [44]). Amplification was performed on an Applied Biosystems GeneAmp 9700 thermal cycler (TFS) following the manufacturer's or recommended protocols [35, 38–44]. PCR fragments were separated and detected using an Applied Biosystems Prism 3500XL Genetic Analyzer (TFS). The analysis of size-based STR fragments was conducted at OL with GeneMapper ID-X software, version 1.2 (TFS) by applying in-house validated dye thresholds: blue – 50 relative fluorescence units (RFU), green – 80 RFU, yellow – 100 RFU, red – 100 RFU.

### Universal Analysis Software

Analysis of MPS data utilized the ForenSeq Universal Analysis Software, version 1.2 (Verogen, exact version see Table 1) for allele and genotype calling by applying the manufacturer's default analysis settings [45]. Universal Analysis Software (UAS) called alleles and genotypes based on the number of reads by applying a specified analytical (AT; 1.5%) and interpretation threshold (IT; 4.5%), except for DYS398II (AT: 5%; IT: 15%), DYS448 (AT: 3.3%; IT: 10%) and DYS635 (AT: 3.3%; IT: 10%) [45]. Single nucleotide polymorphism (SNP) analysis of aDNA samples were manually revised and no-call genotype corrected according to [11]. For example, no-call SNPs that fell below the manufacturer's default IT but in the range of 20 to 29 reads were manually corrected and genotype results put in brackets. No-call SNPs that fell below the manufacturer's default IT but showed reads  $\geq$  30 were manually corrected and called. All laboratories applied UAS using the preinstalled default

settings. Following to sequencing, an Excel-based genotype report-file was generated at each laboratory and sent to the OL for further analysis. Provided data were analysed at OL and reported in anonymous form.

## Statistical analyses

Figures, diagrams and statistical analyses were generated using Microsoft Excel, GraphPad Prism, version 7.04 for Windows and IBM SPSS, version 24 [46].

## Allele frequencies and statistical calculations

Allele frequencies used for genotype frequency calculations were taken from SPSmart v5.1.1 [47–50]. Allele frequencies for SRM 2391c (components A–C) [33] were selected according to the manufacturer's provided information on geographic origin of the respective DNA sample (2391cA: European, 2391cB: American, 2391cC: Oceania). Random match probability (RMP) estimates were calculated according to the formulae 4.1. [51]. Alternatively, likelihood ratios (LR) were used, which is a measure of the power of proof regarding the hypothesis that two DNA profiles were derived from the same suspect and is known as the inverse of the RMP [51, 52].

## Results and discussion

### Sequencing run information

The results of the collaborative exercise were analysed using UAS, which provided the following general information on the sequencing runs (recommended values in brackets [45]; Table 1): cluster density (400–1650 K/mm<sup>2</sup>), cluster passing filter ( $\geq$  80%), phasing ( $\leq$  0.25%) and pre-phasing ( $\leq$  0.15%). Mean cluster density plus standard deviation (SD) was calculated for all runs and amounted on average to 645 K/mm<sup>2</sup> (SD = 141 K/mm<sup>2</sup>) with a minimum of 432 K/mm<sup>2</sup> (run 2) and a maximum of 862 K/mm<sup>2</sup> (run 7; Table 1). The library was analysed at OL after preparation (run 0) and after shipment (run 8). The cluster density of the post-shipment run (run 8) yielded 88.5% of the initial cluster density value (run 0). Still, this cluster density fell within the range of the cluster densities of the other sequencing runs. Mean values (SD in brackets) for cluster passing filter, phasing and pre-phasing amounted on average to 96.49% (1.38%), 0.185% (0.020%) and 0.109% (0.018%) with a minimum of 93.86% (run 7), 0.141% (run 0) and 0.084% (run 4) and a maximum of 97.76% (run 6), 0.226% (run 6) and 0.144% (run 2), respectively (Table 1). Based on these findings, we conducted a more detailed analysis of all datasets to examine possible effects of cluster density variation on the overall data interpretation quality. Below, we present and discuss inter-laboratory concordance ([Inter-](#)



laboratory concordance section), locus coverage (Locus coverage section) and heterozygous balance (Heterozygous balance section). All negative amplification controls yielded no detectable sequences.

### Inter-laboratory concordance

Inter-laboratory concordance was assessed by comparing a) genotype profiles to known reference profiles (SRM 2391c A–C, control DNA 2800M) [22, 33] or previously obtained MPS results (GEDNAP samples G49-S1 and G49-S4; data not shown) and b) genotype from identical library preparation and run in different laboratories. All runs obtained fully concordant autosomal STR (aSTR) and Y-STR results to known reference and between sequencing runs for SRM 2391c A–C, control DNA 2800M, G49-S1 and G49-S4 (Table S4). X-STR analysis revealed full concordance to known reference and between sequencing runs for all runs, except for runs 1 (G49-S1), 2 (2391cA, 2391cC), 4 (2800M) and 6 (2800M, G49-S1). X-STR concordance for control DNA 2800M, SRM 2391cA, 2391cC and G49-S1 was 85.7% (runs 4 and 6: 6 out of 7 alleles; both runs), 92.3% (run 2: 12 out of 13 alleles), 85.7% (run 2: 6 out of 7 alleles) and 91.7% (runs 1 and 6: 11 out of 12 alleles; both runs), respectively (Table S4). All six instances of discordances were due to allelic drop-out at DXS10103 (control DNA 2800M: allele 18 (hemizygous genotype); SRM 2391cA: allele 19 (genotype: 18, 19); SRM 2391cC: allele 19 (hemizygous genotype) and G49-S1: allele 18 (genotype: 18, 20) (Table S4)). Interestingly, for G49-S1 runs 0, 2–5 and 8 obtained a marker drop-out at DX10103 resulting in 83.3% concordance, while only run 7 correctly typed both alleles that were known from reference (Table S4). In addition, DXS10103 was reported to be sensitive to locus drop-out in earlier studies [1, 8, 13, 53, 54] most likely due to underperformance of this marker during amplification. Therefore, further optimization would be required as recommended by [16, 53].

### Performance testing

Cluster density is known to be an important but also critical parameter that affects overall run quality, the total amount of sequencing reads and reads passing filter. For example, while lower cluster density does not necessarily adversely affect data quality, it predominantly leads to decreased data output. In contrast, overclustering potentially leads to poor run performance including the introduction of sequencing artefacts, elevated sequencing error rates together with lowered sequencing data output [55]. Based on the observed cluster density variation (Table 1, “Sequencing run information” section) we checked for possible differences in locus coverage and intra-locus balance (heterozygous genotype). The sample set analysed in

this respect consisted of five single source samples (control DNA 2800M, SRM 2391c A–C and GEDNAP G49-S4).

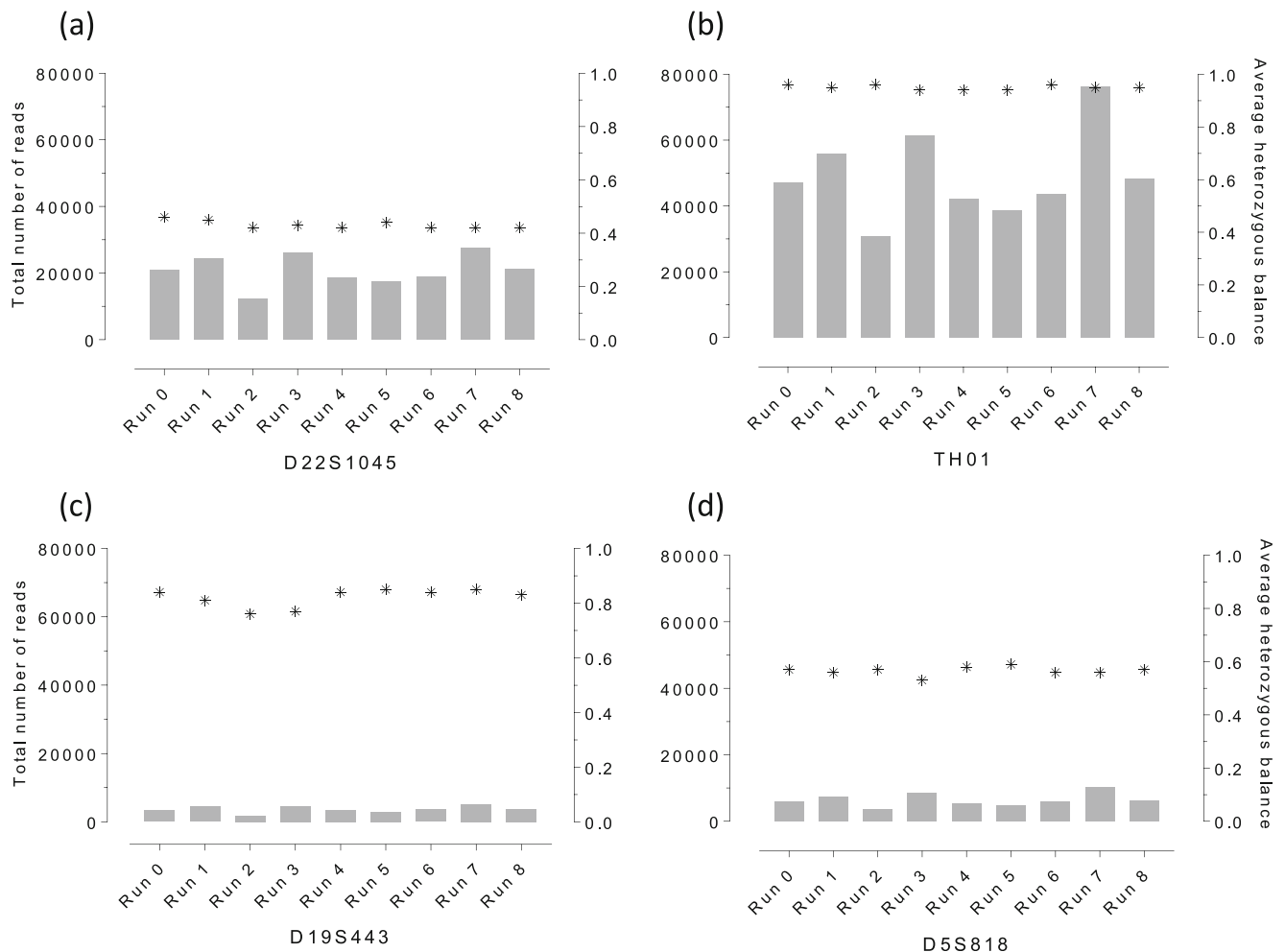
### Locus coverage

Locus coverage was calculated for each STR marker and sequencing run (Table S5). Among aSTRs, TH01 (49,387; 13,501) showed the highest mean number of reads while D19S443 (3691; 1003) performed the least in terms of locus coverage (Table S5). In line with these findings for the ForenSeq DNA Signature Prep Kit, Müller et al., 2018 [4] observed comparable results for TH01 and D19S443 evaluating both Globalfiler MPS-kits (i.e. the Early Access Applied Biosystems Precision ID Globalfiler Mixture ID and the Globalfiler NGS STR Panels) on the Ion S5 instrument [4]. Although using a centrally prepared sequencing library we observed locus coverage variation for single source samples sequenced on the eight MiSeq FGx sequencers (Tables 1 and S5). The mean number of reads/run was 740,397 (205,448) ranging from 438,786 reads (run 2) to 1,123,576 reads (run 7) relating to the afore-mentioned single source samples (Table 1). Run 2 yielded lowest locus coverage over all STRs, whereas run 7 displayed highest locus coverage for nearly all STRs (55 out of 59 STRs; 93.2%) (Table S5).

### Heterozygous balance

Heterozygous balance (HB), also known as intra-locus balance, was calculated (for each sample and locus) for heterozygous- and isometric genotypes by dividing the lower number of allele reads by the higher number of allele reads [56]. Average HB/run and mean HB/locus was calculated for 27 aSTRs, Amelogenin, two Y-STRs and six X-STRs. Results were comparable for each particular locus between all sequencing runs (Table S6). All markers showed  $HB \geq 0.60$  (manufacturer’s default threshold settings) except for D22S1045 and D5S818 that yielded lower HB values, 0.43 (0.02) and 0.57 (0.02), respectively (Table S6). These findings were in line with earlier studies that reported imbalances at D22S1045 [3, 12–15, 17, 22, 54, 57] and D5S818 [53, 54, 57]. Six aSTRs (D4S2408, CSF1PO, D7S820, D8S1179, D9S1122 and TH01), one Y-STR (DYF387S1) and three X-STRs (DXS10135, DXS10074 and HPRTB) displayed  $HB \geq 0.90$  (Table S6). These findings were consistent with previous studies [3, 13]. Results for afore-mentioned aSTRs and one X-STR were in line with Silvia et al., 2017 [54], except for DXS10135 and DXS10074.

Apparently, this analysis showed that low locus coverage did not necessarily increase heterozygous imbalance (Fig. 1). Among aSTRs, TH01 showed highest locus coverage along with highly balanced allele calls (0.95, 0.01). In contrast, D19S443 displayed lowest locus coverages for most sequencing runs with relatively balanced genotype calls (0.82, 0.03)



**Fig. 1** Inter-laboratory investigation of possible correlations between locus coverage differences and heterozygous imbalances. Total locus coverage and average heterozygous balance (HB) were calculated using five single source samples (SRM 2391c A-C, control DNA 2800M and G49-S1; amplified with 1 ng DNA). Against expectations, we found that varying locus coverage did not affect heterozygous balance. For example,

**a** D22S1045 and **d** D5S818 were found to show  $HB \leq 0.60$  but mean locus coverages of 20,838 reads (D22S1045) and 6460 reads (D5S818). In contrast, **b** TH01 and **c** D19S443 were found to be highly balanced forensic markers though showing obvious differences in locus coverage (mean locus coverage of 49,387 reads (TH01) and 3691 reads (D19S443); Table S5)

(Table S6, Fig. 1). Observed differences in locus coverage as a result of cluster density variation had only minor effects on the fluctuation of HB (Tables 1 and S6).

## Sensitivity

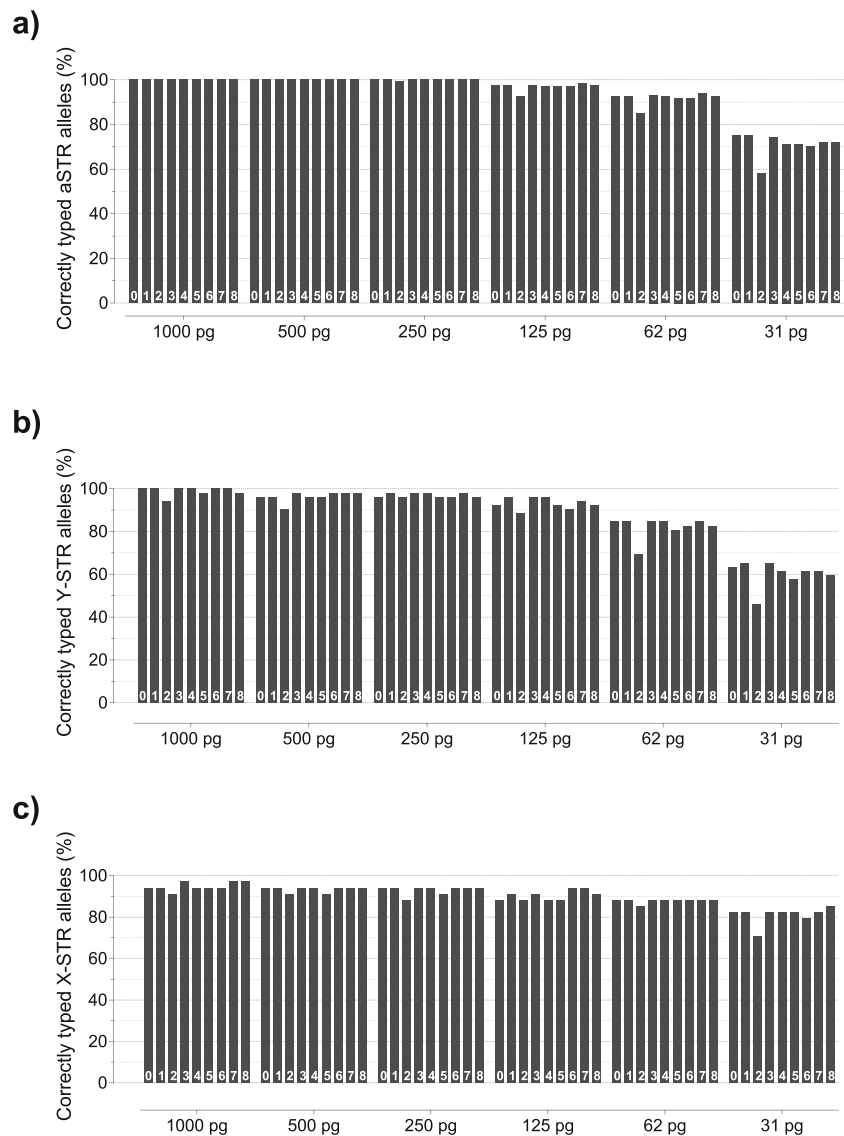
Sensitivity of the ForenSeq DNA Signature Prep Kit was assessed with decreasing DNA input of sample M500 ranging from 1000 pg to 31 pg (final DNA input). Alleles were called based on read counts when they exceeded the manufacturer's default IT [1, 45]. The full profile of M500 consisted of 80 aSTR alleles, 26 Y-STR alleles and 17 X-STR alleles.

The mean aSTR sensitivity success rate over all runs was 100% down to 500 pg and still  $\geq 91.8\%$  (2.6%) down to 62 pg DNA input (Table S7). On average, all sequencing runs successfully typed  $\geq 93.2\%$  (2.8%) of all Y-STRs down to 125 pg DNA (Table S7). Furthermore, X-STR analysis showed

sensitivity levels  $\geq 90.5\%$  (2.5%) using 125 pg DNA (Table S7). For all runs, the average sensitivity level was 71.0% (5.1%), 60.3% (5.8%) and 81.1% (4.2%) for aSTRs, Y-STRs and X-STRs, respectively, using 31 pg DNA (Table S7). These findings were in agreement with Churchill et al., 2016 [3] who reported yields of more than 94% of all alleles using 100 pg DNA amplified with the ForenSeq DNA Signature Prep Kit.

To investigate potential instrument-related differences in sensitivity we carefully examined aSTR, Y-STRs and X-STR results for each sequencing run individually. Results for aSTRs revealed that all runs obtained full-profiles down to 250 pg, except for run 2 (drop-out at D19S443, allele 13, for 250 pg) (Table S8, Fig. 2a). The results showed that all runs obtained comparable sensitivity levels for all classes of markers included in the ForenSeq DNA Signature Prep Kit, except for run 2 (Table S8, Figs. 2 and S1). Particularly for

**Fig. 2** Sensitivity summary showing sensitivity levels for each sequencing run (numbers on the bottom indicate the respective sequencing run) and marker class included in the ForenSeq DNA Signature Prep Kit: **a** autosomal STRs (aSTRs), **b** Y-STRs, and **c** X-STRs. Sample M500 was prepared at ratio 1:1 using control DNAs 9947A and 2800M. A serial dilution was prepared at 1000 pg, 500 pg, 250 pg, 125 pg, 62 pg, and 31 pg total DNA input and amplified in duplicate. Results indicate that all sequencing runs obtained comparable sensitivity levels for all classes of markers, except for run 2. Particularly for DNA input amounts of  $\leq 62$  pg sensitivity levels of run 2 differed distinctly from the other sequencing runs



lower DNA input amounts ( $\leq 62$  pg) sensitivity levels of run 2 differed distinctly from the others (Fig. 2a–c). Still, run 2 success rates were 58.1%, 46.2% and 70.6% for aSTRs, Y-STRs and X-STRs using as little as 31 pg DNA, respectively (Table S7, Figs. 2 and S1). In general, results for the ForenSeq DNA Signature Prep Kit aSTR sensitivity were comparable to CE-based STR kits [58–61] and in line with earlier published MPS studies [1, 3, 4, 14, 54].

## Mixtures

We defined mixtures here as DNA profiles showing more than two alleles in at least two forensic STR markers [62], based on biological material from two or more persons contributing to the sample being tested [63]. Mixture study was performed using samples M166, M91, M62.5 and M47.6 prepared from control DNAs 9947A (minor contributor) and 2800M (major

contributor) at four ratios (83.3:16.7; 90.9:9.1; 93.7:6.3 and 95.2:4.8). The total DNA input amount for each mixture was 1 ng. Within each mixture we identified aSTR alleles that were consistent with the minor contributor genotype (9947A). Note, only those 12 aSTRs with genotype calls that were distinguishable from the major contributor's genotype were included for mixture analysis (D1S1656, TPOX, D2S1338, D3S1358, D5S818, D8S1179, vWA, PentaE, D16S539, D18S51, D21S11 and D22S1045) yielding a total of 41 distinguishable alleles (Table 2 a and b). Alleles were called based on their reads by applying the manufacturer's default thresholds. On average, the ForenSeq DNA Signature Prep Kit was able to detect 100%, 94.4% (1.7%), 76.7% (5.0%) and 45% (5.0%) of the minor contributor's alleles for M166, M91, M62.5 and M47.6, respectively (Table 2b). Furthermore, for M62.5 and M47.6 we observed a noticeable decrease in the proportion of correctly assigned alleles

**Table 2** Summary of mixture analysis showing the total number of alleles observed for mixtures M166, M91, M62.5, and M47.6 prepared at four ratios (83.3:16.7; 90.9:9.1; 93.7:6.3, and 95.2:4.8). Mixtures were prepared using control DNAs 9947A (minor component) and 2800M (major component) amplified with 1 ng total DNA (analysed as singletons). Including only distinguishable autosomal STRs, we were

able to identify 41 potential alleles per mixture at forensic markers D1S1656, TPOX, D2S1338, D3S1358, D5S818, D8S1179, vWA, PentaE, D16S539, D18S51, D21S11, and D22S1045. **a** depicts results obtained for all sequencing runs individually while **b** summarizes the average percentage of correctly typed alleles per minor contributor (dark grey header) and mixture (light grey header)

a)		Mixture ratio 2800M:9947A (%)		Total DNA input/ component (pg)		Total number of observed alleles			Allele drop-out		Total no. of correctly typed alleles (%)	
Run	Reference	Mixture	major:minor	2800M	9947A	Mixture	2800M	9947A	2800M	9947A	Mixture	Minor contributor
aSTRs	0	M166	83.3:16.7	833	167	41	21	20	0	0	41 (100)	20 (100)
	1					41	21	20	0	0	41 (100)	20 (100)
	2					41	21	20	0	0	41 (100)	20 (100)
	3					41	21	20	0	0	41 (100)	20 (100)
	4					41	21	20	0	0	41 (100)	20 (100)
	5					41	21	20	0	0	41 (100)	20 (100)
	6					41	21	20	0	0	41 (100)	20 (100)
	7					41	21	20	0	0	41 (100)	20 (100)
	8	41	21	20	0	0	41 (100)	20 (100)				
	0	M91	90.9:9.1	909	91	40	21	19	0	1	40 (97.6)	19 (95)
	1					40	21	19	0	1	40 (97.6)	19 (95)
	2					40	21	19	0	1	40 (97.6)	19 (95)
	3					40	21	19	0	1	40 (97.6)	19 (95)
	4					39	21	18	0	2	39 (95.1)	18 (90)
	5					40	21	19	0	1	40 (97.6)	19 (95)
	6					40	21	19	0	1	40 (97.6)	19 (95)
	7					40	21	19	0	1	40 (97.6)	19 (95)
	8	40	21	19	0	1	40 (97.6)	19 (95)				
	0	M62.5	93.7:6.3	937.5	62.5	37	21	16	0	4	37 (90.2)	16 (80)
	1					37	21	16	0	4	37 (90.2)	16 (80)
	2					36	21	15	0	5	36 (87.8)	15 (75)
	3					36	21	15	0	5	36 (87.8)	15 (75)
	4					36	21	15	0	5	36 (87.8)	15 (75)
	5					34	21	13	0	7	34 (82.9)	13 (65)
	6					37	21	16	0	4	37 (90.2)	16 (80)
	7					37	21	16	0	4	37 (90.2)	16 (80)
	8	37	21	16	0	4	37 (90.2)	16 (80)				
	0	M47.6	95.2:4.8	952.4	47.6	29	21	8	0	12	29 (70.7)	8 (40)
	1					30	21	9	0	11	30 (73.2)	9 (45)
	2					29	21	8	0	12	29 (70.7)	8 (40)
	3					30	21	9	0	11	30 (73.2)	9 (45)
	4					31	21	10	0	10	31 (75.6)	10 (50)
5	30					21	9	0	11	30 (73.2)	9 (45)	
6	32					21	11	0	9	32 (78.1)	11 (55)	
7	30					21	9	0	11	30 (73.2)	9 (45)	
8	29	21	8	0	12	29 (70.7)	8 (40)					

b)		Overall correctly typed minor contributor's alleles (%)				Overall correctly typed mixture alleles (%)			
Sample	/	Range		Mean (SD)	/	Range		Mean (SD)	
		Min	Max			Min	Max		
M166	/	100	0	100 (0.0)	/	100	0	100 (0.0)	
M91	90	95	5	94.4 (1.7)	95.1	97.6	2.4	97.3 (0.8)	
M62.5	65	80	15	76.7 (5.0)	82.9	90.2	7.3	88.6 (2.4)	
M47.6	40	55	15	45.0 (5.0)	70.7	78.0	7.3	73.2 (2.4)	

(Figure S2a), which is most likely related to the decline in minor contributor's DNA input, while the major contributors DNA input remained almost unchanged (Table 2a, Figure S2b). Referring to M47.6 (ratio: 95.2:4.8) we were able to type 73.2% (2.4%) of all alleles. These findings for mixture analysis were comparable to Jäger et al., 2017 [1], who observed on average 73 allele calls out of 83 alleles (or 87.9%) at 5% minor contributor DNA input including 27 aSTRs. However, when detecting the minor contributors alleles, findings for M91 demonstrated that 5% variation (ranging from 90% to 95%) was obtained between all sequencing runs, which increased to 15% for M62.5 (ranging from 65% to 80%) and M47.6 (ranging from 40% to 55%), respectively (Table 2b, Figure S2a). Still, results for M47.6 showed that on average all runs were able to detect 45% (5%) of the minor contributor's alleles using as little as 47.6 pg DNA (Table 2b).

Again, despite the constant DNA input of 1 ng, clear differences in the total number of sequencing reads were obtained between the participating laboratories (Figure S2b). Figure S2b shows the read-proportion for each mixture component of M166, M91, M62.5, and M47.6 in relation to DNA input. As expected, reads of the minor component decreased with lowered DNA input while those of the major component

accumulated with increasing DNA amount, except for M47.6. Generally, in proportion to DNA input the fraction of the minor contributor 9947A received higher number of reads compared to those of 2800M (Table S9). The observed mixture ratios were equivalent for all mixtures and between sequencing runs, except for M47.6, run 4 (Table S9). Although all mixtures were prepared using 1 ng DNA M47.6 showed an additional drop in the total number of reads for control DNA 2800M. However, this drop approximated the theoretical mixture ratio that would have been expected (Table S9, Figure S2b). It should be noted that the decrease in sequencing reads of control DNA 2800M is more likely related to sample preparation or laboratory procedure than to differences between sequencing instruments.

**Ancient DNA samples**

The analysis of aDNA is associated with possible contamination at various stages of the process, some of which are difficult to control, e.g. contamination prior to receiving the sample in the laboratory. Also, aDNA is characterised by some degree of degradation that depends on various factors such as the general age of the specimen - although this is not following



a strictly linear correlation - and environmental (storage) conditions. Usually ancient samples contain lower DNA quantity compared to contemporary samples [64]. To gather information on the ForenSeq DNA Signature Prep Kit’s performance on highly degraded DNA samples we included four compromised specimens from the early medieval times. As expected, we observed lowered percentage of correctly typed loci compared to high quality samples with broad variation in success rates between the samples (success rates ranged from 1.3% to 96.7%, Table 3, Figure S3). Genotypes were called when they exceeded the manufacturer’s default threshold [45]. Manual revision of the non-called STR loci (automatically by the software) was performed only for aDNA samples (all runs) if the locus information in the provided sample-details-report Excel-table indicated “INC” (genotype) in combination with the comment “interpretation threshold” (QC indicator). INC (inconclusive) indicated ambiguous genotype results that were not reported in the Excel-table. Non-called SNPs that indicated “INC” (genotype) in combination with “interpretation threshold” (QC indicator) and fell in the range of 20 to 29 reads were manually corrected and genotype results put in brackets (Universal Analysis Software section).

Furthermore, the manual revision of homozygous STR genotype calls that indicated the comment “interpretation threshold” (QC indicator) was performed as described above, except for alleles in stutter position.

On average, the results for the ancient samples FA10013B01A, FA10026B01A, FA10030T01A, and FA10058T01B demonstrated that MPS was able to recover 2 out of 152 markers (1.3%), 136 out of 151 markers (90.1%), 116 out of 128 markers (90.6%) and 147 out of 152 markers (96.7%), respectively (Table 3, Figure S3).

The lowest percentage of correctly typed loci was obtained for FA10013B01A dating from the seventh–eighth century showing drop-out rates of 98.7% (150 out of 152 markers failed amplification) (Tables 5 and S10–S12, Figure S3). Only identity SNPs (iSNP) rs1109037 (genotype: GG; all runs) and rs6444724 (genotype: CC; runs 0, 1, 3, 4, and 7) were successfully typed. Manual revision of non-called iSNP genotype calls was performed as described in “Universal Analysis Software” section for rs6444724 (genotype: CC; runs 5, 6, and 8), rs740598 (genotype: A; for runs 1, 6; genotype: G; for run 7) and rs964681 for run 1 (genotype: TT) (Table S13).

**Table 3** Summary of ancient DNA analysis using the ForenSeq DNA Signature Prep Kit amplified for 27 autosomal STRs (aSTRs), 24 Y-STRs, 7 X-STRs, and 94 identity SNPs (iSNPs). The sample set consisted of four compromised samples dating from the fifth to eighth century (for more details see Table S10). Concordance was assessed by comparing genotype results between sequencing runs and to reference

profile (if available, see Tables S10–S12). Overall markers indicate the mean percentage (standard deviation [SD] in brackets) of concordantly typed markers for each ancient DNA sample. The average number of typed markers per sample was calculated based on the mean percentage of concordantly typed markers. The maximum number of markers using the ForenSeq DNA Signature Prep Kit (primer mix A) is 152

Sample	Successfully typed markers per MPS run (%)										Successfully typed markers over all runs			
	Marker	Run										Mean (SD) [%]	Mean no. of typed markers/sample	No. of typed markers vs. total no. of markers
		0	1	2	3	4	5	6	7	8				
FA10013B01A	aSTRs	0	0	0	0	0	0	0	0	0	0	0	/	2/152
	Y-STRs	0	0	0	0	0	0	0	0	0	0	0	/	
	X-STRs	0	0	0	0	0	0	0	0	0	0	0	/	
	iSNPs	2.1	2.1	1.1	2.1	2.1	1.1	1.1	2.1	1.1	1.7 (0.5)	2		
FA10026B01A	aSTRs	92.6	96.3	85.2	96.3	92.6	88.9	92.6	92.6	92.6	92.2 (3.4)	25	136/151 <sup>a</sup>	
	Y-STRs	79.2	83.3	66.7	83.3	83.3	79.2	75.0	79.2	79.2	78.7 (5.3)	19		
	X-STRs	71.4	71.4	57.1	71.4	71.4	71.4	57.1	71.4	71.4	68.2 (6.3)	5		
	iSNPs	93.6	94.6	92.5	93.6	92.5	93.6	92.5	93.6	94.6	93.5 (0.8)	87		
FA10030T01A	aSTRs	85.2	88.9	81.5	92.6	88.9	88.9	81.5	88.9	88.9	87.3 (3.8)	24	116/128 <sup>b</sup>	
	Y-STRs	♀	♀	♀	♀	♀	♀	♀	♀	♀	/			
	X-STRs	85.7	85.7	71.4	85.7	85.7	71.4	85.7	85.7	85.7	82.5 (6.3)	6		
	iSNPs	92.6	92.6	86.2	93.6	92.6	92.6	91.5	93.6	92.6	92.0 (2.3)	86		
FA10058T01B	aSTRs	96.3	100	96.3	100	100	96.3	96.3	96.3	100	97.9 (2.0)	26	147/152	
	Y-STRs	100	100	95.8	100	100	100	95.8	100	100	99.1 (1.9)	24		
	X-STRs	85.7	85.7	85.7	85.7	85.7	85.7	85.7	85.7	85.7	85.7 (0.0)	6		
	iSNPs	96.8	97.8	95.7	97.9	96.8	96.8	98.9	97.9	96.8	97.3 (0.9)	91		

<sup>a</sup> Exclusion of rs722290 from concordance calculation due to tri-allelic genotype; therefore, the total number of markers is 151 instead of 152

<sup>b</sup> Sample derived from a female individual; therefore, the total number of markers recovered with the ForenSeq DNA Signature Prep Kit was 128 instead of 152

The DNA profile of FA10026B01A was concordant between all runs and to known allele calls derived from CE (if available), except for D3S1358 (CE: 15; MPS: 14, 15), D12S391 (CE: 19; MPS: 19, 19.3), D21S11 (CE: 30; MPS: 29, 30) and DYS385a-b (CE: allele 14; MPS: 11, 14) (Tables S10–S11). The careful revision of previously generated CE data for sample FA10026B01A indicated peaks at D3S1358, D12S391, D21S11 and DYS385 that were either located in stutter position and/or fell below the in-house analytical threshold of 50 RFU. Based on these findings, and, because of the comparison of MPS and CE data, the most likely reason for the observed differences is allele drop-out. Results for all runs at locus D13S317 revealed the presence of three different isometric variants of allele 11, viz. (a) sequence identical to the reference sequence as taken from the updated “Forensic STR Sequence Guide v4” file [65] originally published in [66] and available online at <https://strider.online> [67], (b) allele 11 containing rs9546005-T, and (c) allele 11 containing a C>A transversion at nucleotide position GRCh38 CHR13:82148040 together with rs9546005-T, respectively (Tables S10 and S14). Furthermore, run 3 showed two isomeric variants of DYS385a-b, allele 11 (a) sequence identical to the reference sequence as taken from [65, 67], and (b) allele 11 containing a G>A transition located at GRCh38 CHRY:18639770 (18639770-G showing 107 reads and 18639770-A showing 38 reads), respectively (Table S11). A possible reason for sequence variants in degraded samples is based on the misincorporation of nucleotides during enzymatic amplification due to cytosine deamination during DNA degradation [68–71]. One instance of discordance between sequencing runs was found for run 2 showing allelic drop-out of the longer allele at D19S443 (MPS: 13, 14) (Table S10). Identity SNP results for FA10026B01A were fully concordant between sequencing runs, except for rs1015250 and rs722290 (Table S13). Manual revision of rs1015250 no-call iSNP revealed a potential G allele for runs 3–7 showing reads that accounted for 20%, 16%, 16%, 18%, and 14% of the called iSNP-allele reads, respectively. However, non-called iSNP, allele G, was found for rs1015250 for runs 0–2 and 8 too with reads  $\leq$  19. Data analysis revealed one tri-allelic genotype (iSNP genotype: C, G, (A)) at rs722290 for runs 0, 1 and 7. Therefore, rs722290 was excluded from concordance evaluation (Table S13). Observed reads for rs722290, allele A, added up to 4.6%, 4.2%, and 5.6% of allele-reads with highest intensity for runs 0, 1, and 7, respectively (data not shown). Locus drop-out was found for all runs at rs13182883, rs354439, rs719366, and at rs13218440 (runs 0, 2, 3, 5, and 6), rs1736442 (runs 0, 2, 4 and 5), and rs7041158 (run 2) (Table S13). The average mean over all runs for FA10026B01A was 92.2% (3.4%),

78.7% (5.3%), 68.2% (6.3%), and 93.5% (0.8%) of typed aSTRs, Y-STRs, X-STRs and iSNPs, respectively (Table 3, Figure S3).

Results for FA10030T01A were fully concordant between runs and to known reference (if available), except for D7S820 (CE: 8; MPS: 8, 10), D19S433 (CE: 12; MPS: 12, 13). Results for D22S1045 revealed the drop-out of the longer allele within all runs, except for run 3 (CE: 16, 17) (Table S10). In addition, all runs showed drop-out of the shorter allele at DXS8378 (CE: 11, 13) (Table S12). According to previously obtained CE results and to [34], human remains of FA10030T01A were derived from a female individual. Data analysis revealed one Y-STR call at DYS391, allele 11, most likely due to contamination during handling of the sample or laboratory procedure (Table S11). However, the total number of allele-reads for FA10030T01A at DYS391, allele 11, was  $\leq$  49 (Table S11). Identity SNP results for FA10030T01A were fully concordant between runs, except for rs1294331, rs354439, and rs1382387 showing the drop-out of allele A, allele T and allele G, respectively (Table S13). Locus drop-out was observed at rs2920816 (all runs), rs719366 (runs 0, 2–8), rs13182883 (runs 0–2, 4, 6–8) and rs13218440 (runs 1–6, 8). Overall, profile of FA10030T01A consisted of 87.3% (3.8%), 82.5% (6.3%), and 92.0% (2.3%) concordantly typed aSTRs, X-STRs, and iSNPs, respectively (Table 3, Figure S3).

The highest percentage of correctly typed loci was found for FA10058T01B, which was the eldest aDNA sample, dating from the fifth century showing 97.9% (2.0%), 99.1% (1.9%), 85.7% (0.0%), and 97.3% (0.9%) of concordantly typed aSTRs, Y-STRs, X-STRs, and iSNPs, respectively. Results were fully concordant between runs and to known reference (if available) for all 152 markers (Tables S10–S13, Figure S3). Drop-out was obtained for runs 2 and 5 at rs13218440 (Table S13). In addition, MPS was able to identify homozygous genotype as isometric heterozygous genotype for FA10058T01B at D8S1179, allele 13 (repeat structure variants [TCTA] [TCTG] [TCTA]<sub>11</sub> and [TCTA]<sub>13</sub> containing a G>A transition located at GRCh38 CHR8:124894872) and DYF387S1, allele 38 (repeat structure variants [AAAG]<sub>3</sub> [GTAG] [GAAG]<sub>4</sub> [AAAG]<sub>2</sub> [GAAG] [AAAG]<sub>2</sub> [GAAG]<sub>9</sub> [AAAG]<sub>8</sub> [AAAA] [AAAG]<sub>7</sub> and [AAAG]<sub>3</sub> [GTAG] [GAAG]<sub>4</sub> [AAAG]<sub>2</sub> [GAAG] [AAAG]<sub>2</sub> [GAAG]<sub>10</sub> [AAAG]<sub>7</sub> [AAAA] [AAAG]<sub>7</sub> both containing a G>A transition located at GRCh38 CHRY:23785484) (Table S14) [65, 67]. Earlier studies reported sequence variants at D8S1179 [4, 15, 72, 73], which was in line with our findings for this marker. However, at the time of writing, no references were found for DYF387S1. Discordances between sequencing runs were observed for three iSNPs at rs1355366, rs13182883, and rs13218440 due to the drop-out of allele G, allele A and allele A, respectively (Table S13). During manual revision of non-called iSNPs we observed reads that were below 1.8% of the called iSNP-allele reads at rs251934 (allele T) and rs1821380

(allele A) and were thus considered as noise. Runs 2 and 5 obtained locus drop-out at rs13218440 (Table S13).

In general, our results support the findings of earlier publications [74–76] that the application of MPS for STR analysis of degraded and challenging DNA samples can be beneficial. As expected and due to the nature of SNP-containing amplicons [25, 26], the ancient DNA data showed that, in contrast to STRs, the majority of sequencing runs obtained constantly high SNP recovery rates, which was in line with previous reports [8, 75] (Table 3). To evaluate the validity of the obtained results and to assess the impact of iSNPs vs. aSTRs we calculated the RMPs and LR for each aDNA sample (if applicable) as well as for reference material SRM 2391c, components A–C (Tables S15–S18). RMPs for iSNPs were calculated from a subset of 49 iSNP that are included in the validated SNPforID 52-plex panel [77]. Note that rs1355366 was excluded for RMP calculation due to the ambiguous genotype results. The calculated RMPs for 27 aSTRs and 49 iSNPs for aDNAs ranged from  $1.65E-32$  to  $7.15E-27$  and from  $1.37E-29$  to  $8.55E-28$ , respectively (Tables S15–S17), whereas the corresponding RMP values for SRM 2391c A–C ranged from  $3.54E-43$  to  $2.48E-34$  (aSTRs) and from  $9.89E-31$  to  $1.40E-28$  (iSNPs) (Table S18). As expected, the data of the reference material SRM 2391c A–C showed clearly higher LR levels for aSTRs than for the iSNP subset (Figure S4) [52, 78–80]. In the aDNA samples, however, the STR-based LR were roughly in the same range as those obtained for iSNPs. Furthermore, there was no pronounced difference in iSNPs LR levels between reference and aDNA samples (Figure S4) [25, 52]. This relative decline of the STR-based LR compared to the constant levels of SNP-based LR can be an indication for a noticeably better performance of SNPs compared to STRs when analysing old and compromised DNA samples with MPS approaches. However, it is important to mention that the LR for SRM 2391cB was markedly higher than those of components A and C representing an above average LR value and that only a small number of compromised samples were included. Therefore, these data can be regarded as preliminary results indicating that SNPs are a preferable choice of markers, especially for heavily degraded DNA. Nevertheless, this has to be shown in more comprehensive MPS-studies.

Additionally, we contrasted CE-based with MPS-based aDNA genotyping. For this comparison, we calculated LR (if applicable) for the AmpFISTR NGM Select Kit, the AmpFISTR NGM Select loci as included in the ForenSeq DNA Signature Prep Kit and for the entire set of aSTR as well as for a subset of 49 iSNP loci included in the latter kit (Tables S15–S17). The results clearly show the benefit of the large number of markers included in the MPS-based assay that resulted in much higher LR than CE did (Tables S15–S17). In general, we observed comparable MPS-based LR for all aDNA samples and classes of markers suggesting stable performance of the ForenSeq DNA Signature Prep Kit on compromised samples (Figure S5). Overall, the predominant reason for STR discordance between

MPS and CE results within aDNA samples was due to additional allele calls obtained using MPS that lead to an increased power of discrimination, except for FA10013B01A, which showed total drop-out in MPS analysis but a partial profile using CE (Tables S10–S12). Possible reasons could be that the CE analysis was performed eight years earlier and the remaining DNA extract meanwhile frozen, stochastic effects due to low DNA input in general (Table S10) or inhibition of the MPS approach as described in [28, 81]. For instance, using MPS the average profile of FA10026B01A and FA10058T01B was composed of 25 aSTRs, 19 Y-STRs, 5-XSTRs and 87 iSNPs and 26 aSTRs, 24 Y-STRs, 6 X-STRs and 91 iSNPs, respectively (Table 3, Figure S5). However, CE analysis was able to recover 4 aSTRs, 5 Y-STRs and 15 aSTRs, 12 Y-STRs plus 1 X-STR for FA10026B01A and FA10058T01B, respectively (Tables S10–S11, Figure S5). Therefore, it seems important to note that such high numbers of different loci cannot be typed with CE-based STR kits within a single run. This indicates the usefulness of MPS-based technologies for the analysis of compromised samples especially if the original material is limited.

## Conclusions

The presented results are the first MPS-data collected in a collaborative exercise performed among eight independent European laboratories using sequencing instruments of the same supplier. This inter-laboratory study conducted in the framework of the SeqForSTRs project [31] described validation experiments for MPS STR analysis using a standardized sequencing library prepared with the ForenSeq DNA Signature Prep Kit and sequenced on eight different MiSeq FGx instruments. The primary intention of this study was to reduce the inherent complexity derived from multiple steps during manual library preparation to a minimum to test for instrument variation about which knowledge is scarce (including the effect of transport conditions). This was enabled by centralized library preparation using the ForenSeq DNA Signature Prep Kit (primer mix A). Despite broad observed variation in instrument performance, all laboratories obtained run quality metrics that fell within the manufacturer's recommended range. Importantly, obtained locus coverage differences did not necessarily adversely affect heterozygous balances. Inter-laboratory concordance revealed 100% concordance for autosomal- and Y-STRs and yet 85.7% for X-STRs due to drop-out of one allele plus 83.3% due to locus drop-out at marker DXS10103. Mean success rates of all sensitivity runs with 125 pg DNA input were 96.9% (1.7%), 93.2% (2.8%), and 90.5% (2.5%) for aSTR, Y-STR, and X-STR alleles and were comparable to [1, 3, 53, 54] as well as to results using MPS-kits from other suppliers [9, 82].

Interestingly, sensitivity results for each particular run showed that especially for DNA input amounts of < 125 pg the ability to correctly type a STR profile might be dependent on the particular detection limit of the sequencing instrument. Sensitivity results showed that between laboratories the variation in successfully typed aSTRs (highest vs. lowest number of correctly typed alleles) increased with decreasing DNA input and added up to 9% (62 pg) and 17% (31 pg), respectively. These results indicate the importance of performing rigorous in-house validation before implementing MPS into forensic casework applications. Results showed that MPS is a promising tool for the identification of compromised DNA samples and were in line with [8, 75]. The analysis of degraded DNA samples using the ForenSeq DNA Signature Prep Kit showed the total drop-out of FA10013B01A; however, MPS was able to successfully type  $\geq 87\%$  of all aSTRs,  $\geq 78\%$  of all Y-STRs,  $\geq 68\%$  of all X-STRs, and  $\geq 92\%$  of all iSNPs for the remaining three compromised DNA samples. Again, this high number of successfully typed loci was not achieved using conventional CE-based technologies.

**Acknowledgements** The authors would like to thank Burkhard Berger, Harald Niederstätter, and Martin Bodner for helpful discussions, as well as Alexandra Kaindl-Lindinger (all Institute of Legal Medicine, Medical University of Innsbruck) and Angelika Fürst (Bavarian State Criminal Police Office, Munich) for technical support.

**Funding information** Open access funding provided by University of Innsbruck and Medical University of Innsbruck. This study received support from the European Union, grant agreement number IZ25-5793-2015-30 2017-2020 - Evaluierung des Next Generation Sequencing (NGS) für die Anwendbarkeit zur DNA-Analytik in der polizeilichen Praxis von Bund und Ländern – Sequencing of forensic STRs (SeqForSTRs) and grant agreement number HOME/2014/ISFP/AG/LAWX/4000007135 under the Internal Security Funding Police programme of the European Commission-Directorate General Justice and Home Affairs (DNASEqEx).

**Compliance with ethical standards** Research involving human participants and/or animals: does not apply; the study is using standard reference material and DNA extracted from old human remains (> 800 years);

**Conflict of interest** The authors declare that they have no conflict of interest.

**Disclaimer** This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use, which may be made of the information contained therein.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Jäger AC, Alvarez ML, Davis CP, Guzman E, Han Y, Way L et al (2017) Developmental validation of the MiSeq FGx Forensic Genomics System for targeted next generation sequencing in forensic DNA casework and database laboratories. *Forensic Sci Int Genet* 28:52–70
- Churchill JD, Chang J, Ge J, Rajagopalan N, Wootton SC, Chang CW, Lagacé R, Liao W, King JL, Budowle B (2015) Blind study evaluation illustrates utility of the Ion PGM system for use in human identity DNA typing. *Croat Med J* 56:218–229
- Churchill JD, Schmedes SE, King JL, Budowle B (2016) Evaluation of the Illumina Beta Version ForenSeq DNA Signature Prep Kit for use in genetic profiling. *Forensic Sci Int Genet* 20:20–29
- Müller P, Alonso A, Barrio PA, Berger B, Bodner M, Martin P, Parson W, DNASEQEX Consortium (2018) Systematic evaluation of the early access applied biosystems precision ID Globalfiler mixture ID and Globalfiler NGS STR panels for the ion S5 system. *Forensic Sci Int Genet* 36:95–103
- Van Neste C, Van Nieuwerburgh F, Van Hoofstat D, Deforce D (2012) Forensic STR analysis using massive parallel sequencing. *Forensic Sci Int Genet* 6:810–818
- Scheible M, Loreille O, Just R, Irwin J (2014) Short tandem repeat typing on the 454 platform: strategies and considerations for targeted sequencing of common forensic markers. *Forensic Sci Int Genet* 12:107–119
- Scheible M, Loreille O, Just R, Irwin J (2011) Short tandem repeat sequencing on the 454 platform. *Forensic Sci Int Genet Suppl Ser 3*: e357–e3e8
- Xavier C, Parson W (2017) Evaluation of the Illumina ForenSeq DNA Signature Prep Kit - MPS forensic application for the MiSeq FGx benchtop sequencer. *Forensic Sci Int Genet* 28:188–194
- Guo F, Zhou Y, Liu F, Yu J, Song H, Shen H, Zhao B, Jia F, Hou G, Jiang X (2016) Evaluation of the Early Access STR Kit v1 on the Ion Torrent PGM platform. *Forensic Sci Int Genet* 23:111–120
- Strobl C, Eduardoff M, Bus MM, Allen M, Parson W (2018) Evaluation of the precision ID whole MtDNA genome panel for forensic analyses. *Forensic Sci Int Genet* 35:21–25
- Eduardoff M, Gross TE, Santos C, de la Puente M, Ballard D, Strobl C, Børsting C, Morling N, Fusco L, Hussing C, Egyed B, Souto L, Uacyisrael J, Syndercombe Court D, Carracedo A, Lareu MV, Schneider PM, Parson W, Phillips C, EUROFORGEN-NoE Consortium (2016) Inter-laboratory evaluation of the EUROFORGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM. *Forensic Sci Int Genet* 23:178–189
- Just RS, Moreno LI, Smerick JB, Irwin JA (2017) Performance and concordance of the ForenSeq system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. *Forensic Sci Int Genet* 28:1–9
- Köcher S, Müller P, Berger B, Bodner M, Parson W, Roewer L, Willuweit S, DNASEqEx Consortium (2018) Inter-laboratory validation study of the ForenSeq DNA Signature Prep Kit. *Forensic Sci Int Genet* 36:77–85
- Moreno LI, Galusha MB, Just R (2018) A closer look at Verogen's ForenSeq DNA Signature Prep kit autosomal and Y-STR data for streamlined analysis of routine reference samples. *Electrophoresis* 39:2685–2693
- Devesse L, Ballard D, Davenport L, Riethorst I, Mason-Buck G, Syndercombe CD (2017) Concordance of the ForenSeq system and characterisation of sequence-specific autosomal STR alleles across two major population groups. *Forensic Sci Int Genet* 34:57–61
- Phillips C, Devesse L, Ballard D, Weert L, Mil P, Melis S et al (2018) Global patterns of STR sequence variation: sequencing the



- CEPH human genome diversity panel for 58 forensic STRs using the Illumina ForenSeq DNA Signature Prep Kit. *Electrophoresis* 39:2708–2724
17. Churchill JD, Novroski NMM, King JL, Seah LH, Budowle B (2017) Population and performance analyses of four major populations with Illumina's FGx Forensic Genomics System. *Forensic Sci Int Genet* 30:81–92
  18. Gettings KB, Borsuk LA, Steffen CR, Kiesler KM, Vallone PM (2018) Sequence-based U.S. population data for 27 autosomal STR loci. *Forensic Sci Int Genet* 37:106–115
  19. Novroski NMM, King JL, Churchill JD, Seah LH, Budowle B (2016) Characterization of genetic sequence variation of 58 STR loci in four major population groups. *Forensic Sci Int Genet* 25: 214–226
  20. van der Gaag KJ, de Leeuw RH, Hoogenboom J, Patel J, Storts DR, Laros JFJ et al (2016) Massively parallel sequencing of short tandem repeats—Population data and mixture analysis results for the PowerSeq system. *Forensic Sci Int Genet* 24:86–96
  21. Faith SA, Scheible M. Analyzing data from next generation sequencers using the PowerSeq Auto/Mito/Y System. Promega Corporation Web site. <http://at.promega.com/resources/profiles-in-dna/2016/analyzing-data-from-next-generation-sequencers-using-the-powerseq-automitoy-system/> (Updated 2016). 2016.
  22. Illumina. ForenSeq DNA Signature Prep Kit, Reference Guide (Part # 15049528 v01). 2015.
  23. ThermoFisherScientific. Precision ID GlobalFiler NGS STR Panel v2 with the HID Ion S5/HID Ion GeneStudio S5 System: template preparation and sequencing, Quick Reference (Pub. No. MAN0016252, Rev. B.0). 2018.
  24. ThermoFisherScientific (2018) Expand your forensics workflow with the Precision ID NGS System, Brochure
  25. Butler JM, Coble MD, Vallone PM (2007) STRs vs. SNPs: thoughts on the future of forensic DNA testing. *Forensic Sci Med Pathol* 3:200–205
  26. Brookes AJ (1999) The essence of SNPs. *Gene* 234:177–186
  27. Kidd KK, Pakstis AJ, Speed WC, Grigorenko EL, Kajuna SL, Karoma NJ et al (2006) Developing a SNP panel for forensic identification of individuals. *Forensic Sci Int* 164:20–32
  28. Elwick K, Zeng X, King J, Budowle B, Hughes-Stamm S (2018) Comparative tolerance of two massively parallel sequencing systems to common PCR inhibitors. *Int J Legal Med* 132:983–995
  29. Zhang Q, Zhou Z, Liu Q, Liu L, Shao L, Zhang M, Ding X, Gao Y, Wang S (2018) Evaluation of the performance of Illumina's ForenSeq system on serially degraded samples. *Electrophoresis* 39:2674–2684
  30. Zeng X, Elwick K, Mayes C, Takahashi M, King JL, Gangitano D et al (2018) Assessment of impact of DNA extraction methods on analysis of human remain samples on massively parallel sequencing success. *Int J Legal Med*
  31. Bundeskriminalamt, Fachbereich-IZ-25, Berlin, Germany. EU Project ISF, Nr. IZ25-5793-2015-30 2017-2020 Evaluierung des Next Generation Sequencing (NGS) für die Anwendbarkeit zur DNA-Analytik in der polizeilichen Praxis von Bund und Ländern – Sequencing of forensic STRs (SeqforSTRs)
  32. Illumina (2015) MiSeq FGx Instrument, Reference Guide (Part# 15050524, Rev. C)
  33. NationalInstituteofStandards&Technology. Standard Reference Material 2391c PCR-based DNA profiling standard (Certificate of Analysis). 2011, Certificate Revision History: 03 April 2015
  34. McGlynn G (2007) Using  $^{13}\text{C}$ -,  $^{15}\text{N}$ -, and  $^{18}\text{O}$  stable isotope analysis of human bone tissue to identify transhumance, high altitude habitation and reconstruct palaeodiet for the early medieval Alpine population at Volders. Ludwig-Maximilians-University at München, Austria
  35. Bauer CM, Niederstätter H, McGlynn G, Stadler H, Parson W (2013) Comparison of morphological and molecular genetic sex-typing on mediaeval human skeletal remains. *Forensic Sci Int Genet* 7:581–586
  36. Walker JA, Hedges DJ, Perodeau BP, Landry KE, Stoilova N, Laborde ME, Shewale J, Sinha SK, Batzer MA (2005) Multiplex polymerase chain reaction for simultaneous quantitation of human nuclear, mitochondrial, and male Y-chromosome DNA: application in human identification. *Anal Biochem* 337:89–97
  37. Niederstätter H, Köchl S, Grubwieser P, Pavlic M, Steinlechner M, Parson W (2007) A modular real-time PCR concept for determining the quantity and quality of human nuclear and mitochondrial DNA. *Forensic Sci Int Genet* 1:29–34
  38. Promega. PowerPlex ESX 17 Fast System for use on the Applied Biosystems Genetic Analyzers, Technical Manual. September, 17
  39. Promega. PowerPlex 16 System, Technical Manual. Revised 5/16
  40. Promega. PowerPlex 21 System for use on the Applied Biosystems Genetic Analyzers, Technical Manual. Revised 4/17
  41. Qiagen (2013) Investigator Argus X-12 Handbook
  42. ThermoFisherScientific. AppliedBiosystems AmpFISTR Yfiler PCR Amplification Kit, User Guide (Pub.no. 4358101, Rev. J). 2014:234
  43. Esteve Codina A, Niederstatter H, Parson W (2009) “GenderPlex” a PCR multiplex for reliable gender determination of degraded human DNA samples and complex gender constellations. *Int J Legal Med* 123:459–464
  44. ThermoFisherScientific (2015) AppliedBiosystems AmpFISTR NGM Select PCR Amplification Kit, User Guide (Pub. no. 4458841, Rev. F)
  45. Illumina (2016) ForenSeq Universal Analysis Software Guide (Document # 15053876 v01)
  46. IBM-Corp (2016) IBM SPSS Statistics for Windows, Version 24.0. Armonk
  47. Amigo J, Salas A, Phillips C, Carracedo A (2008) SPSmart: adapting population based SNP genotype databases for fast and comprehensive web access. *BMC Bioinformatics* 9:428
  48. Amigo J, Phillips C, Salas T, Formoso L, Carracedo A, Lareu M (2009) pop.STR—An online population frequency browser for established and new forensic STRs. *Forensic Sci Int Genet Suppl Ser* 2:361–362
  49. Amigo J, Phillips C, Lareu M, Carracedo A (2008) The SNPforID browser: an online tool for query and display of frequency data from the SNPforID project. *Int J Legal Med* 122:435–440
  50. Amigo J, Phillips C, Salas A, Carracedo A (2009) Viability of in-house datamarting approaches for population genetics analysis of SNP genotypes. *BMC Bioinformatics* 10(Suppl 3):S5
  51. National Research C (1996) The evaluation of forensic DNA evidence. The National Academies Press, Washington, DC
  52. Butler JM (2005) Forensic DNA typing: biology, technology, and genetics of STR markers, 2nd edn. Elsevier Academic Press, New York
  53. Guo F, Yu J, Zhang L, Li J (2017) Massively parallel sequencing of forensic STRs and SNPs using the Illumina ForenSeq DNA Signature Prep Kit on the MiSeq FGx Forensic Genomics System. *Forensic Sci Int Genet* 31:135–148
  54. Silvia AL, Shugarts N, Smith J (2017) A preliminary assessment of the ForenSeq FGx System: next generation sequencing of an STR and SNP multiplex. *Int J Legal Med* 131:73–86
  55. Illumina (2016) Optimizing cluster density on Illumina Sequencing Systems - understanding cluster density limitations and strategies for preventing under- and overclustering
  56. Gill P, Sparkes R, Fereday L, Werrett DJ (2000) Report of the European Network of Forensic Science Institutes (ENSFI): formulation and testing of principles to evaluate STR multiplexes. *Forensic Sci Int* 108:1–29
  57. Almalki N, Chow HY, Sharma V, Hart K, Siegel D, Wumbach E (2017) Systematic assessment of the performance of illumina's MiSeq FGx forensic genomics system. *Electrophoresis* 38:846–854



58. Martin P, de Simon LF, Luque G, Farfan MJ, Alonso A (2014) Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci. *Forensic Sci Int Genet* 13:68–78
59. Oostdik K, Lenz K, Nye J, Schelling K, Yet D, Bruski S, Strong J, Buchanan C, Sutton J, Linner J, Frazier N, Young H, Matthies L, Sage A, Hahn J, Wells R, Williams N, Price M, Koehler J, Staples M, Swango KL, Hill C, Oyerly K, Duke W, Katzilierakis L, Ensenberger MG, Bourdeau JM, Sprecher CJ, Krenke B, Storts DR (2014) Developmental validation of the PowerPlex Fusion System for analysis of casework and reference samples: a 24-locus multiplex for new database standards. *Forensic Sci Int Genet* 12:69–76
60. Ensenberger MG, Lenz KA, Matthies LK, Hadinoto GM, Schienman JE, Przech AJ, Morganti MW, Renstrom DT, Baker VM, Gawrys KM, Hoogendoorn M, Steffen CR, Martin P, Alonso A, Olson HR, Sprecher CJ, Storts DR (2016) Developmental validation of the PowerPlex Fusion 6C System. *Forensic Sci Int Genet* 21:134–144
61. Wang DY, Gopinath S, Lagace RE, Norona W, Hennessy LK, Short ML et al (2015) Developmental validation of the GlobalFiler Express PCR Amplification Kit: a 6-dye multiplex assay for the direct amplification of reference samples. *Forensic Sci Int Genet* 19:148–155
62. Schneider PM, Fimmers R, Keil W, Molsberger G, Patzelt D, Pflug W, Rothämel T, Schmitter H, Schneider H, Brinkmann B (2009) The German Stain Commission: recommendations for the interpretation of mixed stains. *Int J Legal Med* 123:1–5
63. Butler JM, Kline MC, Coble MD (2018) NIST interlaboratory studies involving DNA mixtures (MIX05 and MIX13): variation observed and lessons learned. *Forensic Sci Int Genet* 37:81–94
64. Marciniak S, Klunk J, Devault A, Enk J, Poinar HN (2015) Ancient human genomics: the methodology behind reconstructing evolutionary pathways. *J Hum Evol* 79:21–34
65. Phillips C, Gettings KB, King JL, Ballard D, Bodner M, Borsuk L, Parson W (2018) "The devil's in the detail": Release of an expanded, enhanced and dynamically revised forensic STR Sequence Guide. *Forensic Sci Int Genet* 34:162–169
66. Parson W, Ballard D, Budowle B, Butler JM, Gettings KB, Gill P, Gusmão L, Hares DR, Irwin JA, King JL, Knijff P, Morling N, Prinz M, Schneider PM, Neste CV, Willuweit S, Phillips C (2016) Massively parallel sequencing of forensic STRs: considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements. *Forensic Sci Int Genet* 22:54–63
67. Bodner M, Bastisch I, Butler JM, Fimmers R, Gill P, Gusmão L et al (2016) Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER). *Forensic Sci Int Genet* 24:97–102
68. Dabney J, Meyer M, Pääbo S (2013) Ancient DNA damage. *Cold Spring Harb Perspect Biol* 5
69. Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res* 29:4793–4799
70. Stiller M, Green RE, Ronan M, Simons JF, Du L, He W et al (2006) Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc Natl Acad Sci U S A* 103:13578–13584
71. Briggs AW, Stenzel U, Johnson PL, Green RE, Kelso J, Prufer K et al (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A* 104:14616–14621
72. Gettings KB, Aponte RA, Vallone PM, Butler JM (2015) STR allele sequence variation: current knowledge and future issues. *Forensic Sci Int Genet* 18:118–130
73. Gettings KB, Kiesler KM, Faith SA, Montano E, Baker CH, Young BA, Guerrieri RA, Vallone PM (2016) Sequence variation of 22 autosomal STR loci detected by next generation sequencing. *Forensic Sci Int Genet* 21:15–21
74. Calafell F, Anglada R, Bonet N, Gonzalez-Ruiz M, Prats-Munoz G, Rasal R et al (2016) An assessment of a massively parallel sequencing approach for the identification of individuals from mass graves of the Spanish Civil War (1936–1939). *Electrophoresis* 37:2841–2847
75. Fattorini P, Previderé C, Carboni I, Marrubini G, Sorçaburu-Cigliero S, Grignani P, Bertoglio B, Vatta P, Ricci U (2017) Performance of the ForenSeq™ DNA Signature Prep kit on highly degraded samples. *Electrophoresis* 38:1163–1174
76. Kulstein G, Hadrys T, Wiegand P (2018) As solid as a rock—comparison of CE- and MPS-based analyses of the petrosal bone as a source of DNA for forensic identification of challenging cranial bones. *Int J Legal Med* 132:13–24
77. Sanchez JJ, Phillips C, Borsting C, Balogh K, Bogus M, Fondevila M et al (2006) A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis* 27:1713–1724
78. Gill P (2001) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. *Int J Legal Med* 114:204–210
79. Gill P, Werrett DJ, Budowle B, Guerrieri R (2004) An assessment of whether SNPs will replace STRs in national DNA databases—joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDM). *Sci Justice* 44:51–53
80. Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B (1999) The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems. *Electrophoresis* 20:1682–1696
81. Sidstedt M, Steffen CR, Kiesler KM, Vallone PM, Radstrom P, Hedman J (2019) The impact of common PCR inhibitors on forensic MPS analysis. *Forensic Sci Int Genet* 40:182–191
82. Wang Z, Zhou D, Wang H, Jia Z, Liu J, Qian X et al (2017) Massively parallel sequencing of 32 forensic markers using the Precision ID GlobalFiler NGS STR Panel and the Ion PGM System. *Forensic Sci Int Genet* 31:126–134

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