Forensic Science International: Genetics 48 (2020) 102344

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Research paper

Forensic evaluation of the Asia Pacific ancestry-informative MAPlex assay

C. Xavier[a,*,1], M. de la Puente[a,b,1], C. Phillips[b], M. Eduardoff[a], A. Heidegger[a], A. Mosquera-Miguel[b], A. Freire-Aradas[b], R. Lagace[c], S. Wootton[c], D. Power[d], W. Parson[a,e], M.V. Lareu[b], R. Daniel[f,g]

[a] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[b] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[c] Human Identification Group, Thermo Fisher Scientific, CA, USA
[d] Thermo Fisher Scientific, Victoria, Australia
[e] Forensic Science Program, The Pennsylvania State University, University Park, PA, USA
[f] Office of the Chief Forensic Scientist, Forensic Services Department, Victoria Police, Macleod, Australia
[g] National Centre for Forensic Studies, Faculty of Science & Technology, University of Canberra, ACT, Australia

ARTICLE INFO

Keywords:
Biogeographical ancestry estimation
Asia-Pacific populations
Massive parallel sequencing
Ion S5 sequencing

ABSTRACT

DNA intelligence, and particularly the inference of biogeographical ancestry (BGA) is increasing in interest, and relevance within the forensic genetics community. The majority of current MPS-based forensic ancestry-informative assays focus on the differentiation of major global populations. The recently published MAPlex (Multiplex for the Asia Pacific) panel contains 144 SNPs and 20 microhaplotypes and aims to improve the differentiation of populations in the Asia Pacific region. This study reports the first forensic evaluation of the MAPlex panel using AmpliSeq technology and Ion S5 sequencing. This study reports on the overall performance of MAPlex including the assay's sequence coverage distribution and stability, baseline noise and description of problematic SNPs. Dilution series, artificially degraded and mixed DNA samples were also analysed to evaluate the sensitivity of the panel with challenging or compromised forensic samples. As the first panel to combine biallelic SNPs, multiple-allele SNPs and microhaplotypes, the MAPlex assay demonstrated an enhanced capacity for mixture detection, not easily performed with common binary SNPs. This performance evaluation indicates that MAPlex is a robust, stable and highly sensitive assay that is applicable to forensic casework for the prediction of BGA.

1. Introduction

The MAPlex assay (Multiplex for the Asia Pacific) is an ancestry-informative multiplex of SNP and microhaplotype markers [1], which are genotyped by massively parallel sequencing (MPS) using the Thermo Fisher Scientific (TFS) Ion S5 System [2]. MAPlex was developed to provide optimal differentiation of East Asian, South Asian and Near Oceanian populations found in the geographically extensive Asia Pacific regions; and comprises 164 markers: 108 bi-allelic SNPs, 36 multiple-allele SNPs and 20 microhaplotypes. The development of the MAPlex assay and its route to an optimized MPS test applicable to forensic DNA analysis consisted of three stages [3,4]: i. the selection and combination of the most powerful ancestry-informative markers (AIMs) for differentiating globally distributed population groups into a single robust MPS multiplex, with a focus on the Asia Pacific region; ii. the evaluation of the genotyping accuracy and forensic performance of the multiplex by its application to analysing the range of DNA samples commonly encountered in forensic casework, comprising low level DNA, degraded DNA, mixed DNA and control DNA, where the genotypes have been previously established independently of the laboratories genotyping the MAPlex markers using MPS; iii. compilation of a broad range of worldwide populations to construct a reference database of genetic variability and the development of enhanced ancestry predictive tools. The first stage has been completed and the description of the MAPlex component AIMs and the compilation of initial sets of reference and test population data formed the first publication [5]. This study describes the evaluation of the forensic effectiveness of MAPlex, combined with genotyping concordance studies to ensure the MPS-based sequence analysis accurately detects the variation present in each component AIM.

* Corresponding author at: Institute of Legal Medicine, Medical University of Innsbruck, Müllerstraße 44, 6020 Innsbruck, Austria.
E-mail address: catarina.gomes@i-med.ac.at (C. Xavier).
[1] These authors contributed equally.

https://doi.org/10.1016/j.fsigen.2020.102344
Received 19 February 2020; Received in revised form 22 April 2020; Accepted 15 June 2020
Available online 20 June 2020

The forensic evaluation steps undertaken form a series of analyses in a standardized compact arrangement of test DNAs designed to measure the forensic effectiveness of MAPlex. These comprise: dilution series of standard control DNAs to gauge sensitivity of MAPlex for accurate genotyping of low level DNA; analysis of artificially degraded DNA to measure how successfully the short amplicons of MAPlex can detect its component loci; artificially mixed DNA at ratios aiming to assess the extent to which MAPlex can detect minor components of ~10 % in simple 2-way mixtures. The ability of MAPlex to detect and analyse mixed DNA was evaluated with regard to 2-way mixtures of individuals with different or identical ancestries, as a forensic ancestry panel tends to compile markers with sharply contrasting allele frequency distributions, this characteristic can enhance or diminish the ability of the assay to detect atypical sequence ratio patterns that signal a mixture, and this was fully explored.

Although the focus of the studies we report here was on forensic performance, the analysis of common control DNA samples (which are commercially available) enabled detailed assessments of the assay's reproducibility - by comparing the results from two MAPlex development laboratories; and genotyping accuracy - from comparison to online databases of genotypes for the same samples, generated with different sequencing technologies. Potentially of more importance is the opportunity to measure the sequencing characteristics of the multiplex, its component loci, and the Ion S5 sequencing system for which MAPlex has been designed. Such characteristics include: base misincorporation rates and the analytical thresholds necessary to discount baseline non-allelic base reads; distribution of sequence coverage and strand bias amongst the MAPlex markers; the allelic balance of sequence ratios in heterozygous genotypes; and alignment issues which may be unavoidable if the context sequence of key markers has polymeric base motifs or unmapped insertion-deletion polymorphisms (Indels) close to the target site. We report detailed evaluation of both the forensic parameters and genotyping precision of MAPlex. Lastly, it is important to note that a key feature of MAPlex that shaped a significant part of the analyses made, is the combination of binary SNPs with multiple-allele SNPs and microhaplotypes for the first time in a forensic MPS multiplex.

2. Materials and methods

2.1. Experimental design, DNA samples and preparation of mixtures

A series of six tests, commonly used in forensics as internal validation of new methods, were used to evaluate the overall performance of the MAPlex assay. To assess the reproducibility of MAPlex, 16 replicates of 1 ng and 3 replicates of 2 ng 2800 M Control DNA (Promega, Madison, US) were sequenced following manufacturer's guidelines. All genotypes were compared and individual locus performance analyzed. To evaluate genotype accuracy a concordance study was conducted using 7 Coriell cell line DNA samples: HG00403; NA18498; NA06994; NA07000; NA07029; NA11200; NA10540. The genotyping concordance framework used allowed the comparison of: a) results from two participating laboratories; and, b) laboratory results with the curated online variant databases of the 1000 Genomes Project Phase III [6] for samples HG00403, NA18498, NA06994, NA07000 and Simons Genome Diversity Project (SGDP) [7] for sample NA11200. Discordances with 1000 Genomes Phase III data were reviewed in detail using the recently released 1000 Genomes Project New York Genome Center high coverage dataset [8].

To estimate the lower limits of DNA input for the MAPlex assay, a sensitivity doubling dilution series (2 ng-0.016 ng) of the 2800 M control DNA (Promega) was prepared and run four times (dilution series 1–4). A second sensitivity dilution series of 2800 M control DNA (10 ng to 10 pg) was made to compare the manufacturer's guideline for library preparation (22 cycles) for the recommended 1 ng DNA input (dilution series 5 and 6) and the manufacturer's guideline for degraded/low-level (< 1 ng input) DNA samples with an increased number of PCR cycles (27 cycles) (dilution series 7). This approach was shown to increase coverage for low quantity samples in previous studies [9,10].

The performance of the MAPlex assay with challenging DNA samples representative of typical forensic analyses, was assessed by running a series of enzymatically-sheared in-house control DNA samples prepared with the Ion Shear Plus kit (TFS) at 20, 25 and 45 min intervals. Degradation status was assessed by common STR typing using AmpF*l*STR™ NGM SElect™ PCR Amplification Kit (TFS). Finally, the capacity of the MAPlex assay to resolve DNA mixtures was evaluated with a set of artificial two-component mixtures prepared at volume ratios of 1:1, 1:3 and 1:9 using the following Coriell DNA samples at ~1 ng/μl (quantified using Qubit® ds DNA HS Assay, TFS): NA18498 (African population of origin: AFR) – HG00403 (East Asian: EAS); HG00403 (EAS) – NA07000 (European: EUR); NA06994 (EUR) – NA18498 (AFR); and NA07000 (EUR) – NA06994 (EUR). Final input amount of DNA for library preparation was ~1 ng of mixed DNA.

2.2. Library and template preparation for Ion S5 sequence analysis

Libraries were prepared from all samples using the DL8 Ion AmpliSeq Kit for Chef (TFS, Waltham, US) and the Ion Chef (TFS) for automated library construction. A final volume of 15 μL of all samples was loaded onto the Ion Chef following manufacturer's guidelines. This produced pooled libraries combining batches of eight samples. The library pools were quantitated using the real time Ion Library TaqMan Quantitation Kit qPCR protocol, diluted to 30 pM, pooled to equimolar concentrations and re-loaded onto the Ion Chef for template preparation using the Ion 520 and Ion 530 Kits (TFS). All protocols followed manufacturer's guidelines. Finally, the pooled libraries were loaded onto Ion 530 Chips (TFS) and sequenced with the Ion S5 System (TFS).

2.3. Data analysis

Data was analyzed using the Torrent Suite Software (TFS) which applies the Torrent Mapping Alignment Program (TMAP) algorithm for aligning the reads to the reference genome (GRCh37/hg19) and genotype calls made with the TFS HID SNP Genotyper version 5.2.2 plugin (herein Genotyper). Genotyper used standard parameters for single-source samples comprising a minimum coverage (min_cov) of 6 reads and a minimum allele frequency for heterozygous call (min_allele_freq) of 0.1 [11,12]. For mixtures, the min_allele_freq threshold was set at 0.02 for an enhanced detection of the minor component alleles [10,13,14]. Genotyper outputs two different files: a variant call format (vcf) and comma separated values file (csv) with all the details per SNP target describing: the genotype call, coverage, number of reads per base, number of forward and reverse reads, and major allele frequency. SNP genotypes and sequence coverage data were retrieved and analyzed further using R v. 3.5.0 [15]. Genotypes and alignments were manually confirmed by loading the BAM/BAI files into the Integrative Genomic Viewer (IGV) [16]. The haplotypes of the microhaplotype loci, in the form of phased SNP alleles, were obtained using the TVC_Microhaplotyper version 8.1 plugin (TFS, herein Microhaplotyper). The parameter 'relative analytical threshold' was kept at the default value of 0.02 for mixtures and set to 0.1 for single-source sample analysis.

3. Results & discussion

3.1. Genotyping concordance

Genotyping concordance was evaluated at two levels: i. inter-laboratory concordance of seven Coriell controls; and ii. concordance between laboratories and online data from 1000 Genomes (4/7 Coriell controls) and SGDP (1/7 Coriell controls). Table 1 lists the discordant genotypes found for both comparisons.

Overall, inter-laboratory concordance reached 99.80 %. A total of 3

Table 1
Discordant genotypes summary, listing genotypes obtained using the MAPlex panel in Lab 1 and Lab 2, and publicly available genotypes from the 1000 Genomes Project. 1kGP3: 1000 Genomes Phase III; 1kGHC: 1000 Genomes high coverage data; "–": not discordant.

	HG00403				NA18498				NA7000					NA11200	
	1kGHC	1kGP3	Lab 1	Lab 2	1kGHC	1kGP3	Lab 1	Lab 2	1kGHC	1kGP3	Lab 1	Lab 2	Lab 1		Lab 2
rs776912	TT	CT	TT	TT		–	–	–		–	–	–	–		–
MH3_rs3111398	CC	CT	CC	CC		–	–	–		–	–	–	–		–
rs2789823		–	–	–	GG	AG	GG	GG		–	–	–	–		–
rs408046	AT	AT	AA	AT		–	–	–	GT	GT	AG	GT	GT		AG
MH20_rs621340		–	–	–	GG	GT	GG	GG		–	–	–	–		–
MH21_rs6517970	AC	AA	AC	AC		–	–	–		–	–	–	–		–

1kGP3: 1000 Genomes Phase 3 data.
1kGHC: 1000 Genomes high coverage data.
–: not discordant.

discordances in 1519 genotypes were found for samples HG00403, NA07000 and NA11200 in SNP rs408046. Visualization of the BAM/BAI files in IGV revealed differences in the alignments causing the discordant genotype calls (Supplementary File S1A). These differences could be due to use of different versions of the TFS Torrent Suite Software (v.5.2.1 in Lab 1 vs. v.5.2.2 in Lab 2), with the genotype calls made with v.5.2.2 being fully concordant with those of 1000 Genomes.

Concordance between laboratories and 1000 Genomes reached a level of 99.54 %. In total, 4 discordances were found in 868 genotypes in different SNPs – three of these were SNPs within microhaplotypes. The cause of the discordances remained unclear after IGV scrutiny of the sequences, and in each case the genotypes between laboratories were fully concordant. These discordances were further investigated by comparing our variant calls to those of the 1000 Genomes New York Genome Center high coverage dataset [8], indicating genotypes were fully concordant with those obtained in both laboratories, reaching 100 % concordance. Full concordance was observed between laboratories and SGDP data.

Microhaplotype component SNP alleles that form the haplotypes were also evaluated in order to assess phasing concordance, reaching an inter-laboratory level of 100 % (no discordances in 140 genotypes). Concordance with 1000 Genomes database reached a level of 96.25 %, the three discordances (in 80 genotypes) were caused by the discrepancies described above. Full concordance was observed between laboratories and SGDP phased available data (no discordances in 14 genotypes).

3.2. Reproducibility

In the 16 replicates of 2800 M at 1 ng and 3 replicates at 2 ng all genotypes were concordant and dropout was not observed. Sequence data from the replicates allowed the analysis of other factors related to the Ion S5™ sequencing performance and the assay design. Detailed analysis of mean coverage, coverage distribution, strand bias and base misincorporation rates are described in the following sections. In order to screen for poor performing SNPs, all positions for all replicates were inspected with IGV, which led to the identification of two SNPs consistently underperforming in all 1 ng replicates: rs2387842 and rs1422656 showed alignment issues due to the presence of repetitive regions and closely positioned Indels [17]. IGV alignments of the amplicons containing each SNP are outlined in Supplementary Files S1B-C. The observation of underperforming SNPs due to homopolymeric tracts and Indels has been previously reported [10,14,18].

3.2.1. Sequence coverage

As expected, 2 ng replicates resulted in higher values of total coverage than 1 ng replicates, with mean values of 1,606,415.33 and 816,688.69, respectively (Supplementary Table S1). The automated library preparation process involves bead equalization of individual libraries prior to generating a pooled library. Although the equalization aims to achieve a 100 pM library prior to pooling, the initial template input amount impacts on the representation of the library in the final pool. This may be observed by variation in sequencing read depth for samples of varying initial template input amounts. A statistic T-test was applied to evaluate the difference between the total coverage values in 1 ng and 2 ng replicates which produced a p-value < 0.05 and thus statistically significant (p-value = 0.00023081), indicating the MAPlex assay is sensitive to DNA input.

Coverage distribution per SNP was analyzed by normalizing the number of reads in each locus by the total coverage recorded for the whole MAPlex panel. This resulted in very similar coverage distribution between all replicates (Fig. 1A and Supplementary Table S2), showing the robustness and stability of the assay for its optimum DNA input. Nevertheless, all three 2 ng replicates show higher median values than the 1 ng set and less variability in coverage distribution (Fig. 1A – the last three samples vs. the others). A Kolmogorov-Smirnov statistical test was applied to assess differences in mean normalized coverage distribution between 1 ng and 2 ng replicates (Fig. 1B and Supplementary Table S2). Fig. 1B also shows SNPs that have low coverage in 1 ng also have low coverage in 2 ng replicates. Similarly, the high coverage outliers at the other extreme of this plot are the same markers in both replicate sets, confirming the assay's stability. The Kolmogorov-Smirnov test produced two statistical values: D, accounting for the absolute maximum difference between the mean normalized coverage values of both sets of replicates (1 and 2 ng); and a p-value. 1 ng and 2 ng replicate mean normalized coverage distributions were found to be statistically different (D = 0.13364, p-value = 0.04148), in agreement with previous results that indicate MAPlex is sensitive to initial DNA input. Such trends can also be graphically observed by plotting the quantiles of both distributions against each other using normal quantile-quantile (Q-Q) plots (Supplementary Fig. S1). Normal Q-Q plots were also produced for each distribution (1 ng and 2 ng) and no linear relation was observed indicating that normalized coverage does not follow a normal distribution pattern (Supplementary Fig. S2).

3.2.2. Strand bias

Strand bias was analyzed by calculating the ratio between SNP target reads on the forward strand divided by the total number of SNP target reads (Fig. 2A). Even though strong strand bias was observed in a previous study [10] of the Ion PGM™ sequencing system, in our analyses of MAPlex sequence data the range of values for mean strand bias ratio per SNP observed was much smaller (mean values of $0.361 - 0.611$). Indeed, absolute strand bias values of 1 and 0 were not observed, indicating that in all MAPlex SNPs both forward and reverse reads made reasonably balanced contributions to the genotype calling, and demonstrating noticeable improvement of sequence analysis in the Ion S5, as well as optimized assay design pipelines. Considering replicate mean strand ratio, the mean value obtained was 0.497, very

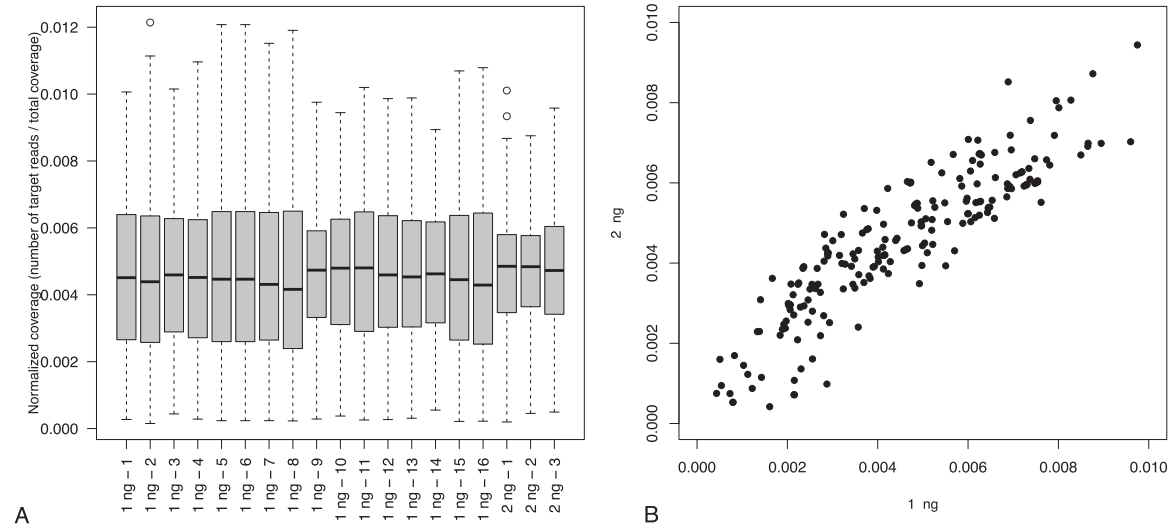

Fig. 1. A: Normalized coverage distribution in all markers per reproducibility replicate of 1 ng and 2 ng DNA input. Normalized coverage was calculated as the ratio between the number of reads at the SNP position and the total coverage of the replicate. B: Normalized coverage per SNP in mean 1 ng replicates plotted against the mean 2 ng replicates.

close to the optimum 0.5. Only 26 SNPs showed mean strand ratio values falling outside the interval $0.45 < x < 0.55$, representing 12 % of the entire panel, and confirming an improved assay design process.

3.2.3. Base misincorporation rates

Base misincorporation analysis detects incorporation of bases not matching the original genotype at individual SNP sites (Fig. 2B and Table S3). A comparison between the insertion of a non-specific base and the incorporation of a base matching the alternative allele in a homozygote distinguishes between random erroneous insertions or base incorporations that could possibly lead to an incorrect genotype call. The misincorporation rates obtained with MAPlex were very similar (T-test p-value = 0.283), showing no observable difference between a described allele base (DAM) over a random base misincorporation (RBM). In fact, when calculating the ratio between both DAM and RBM and the total misincorporation reads, very close values of 0.566 and 0.433, respectively, were obtained. The mean percentage misincorporation observed was 0.175 %. In total, 53 SNPs showed values above the mean (24.4 % of the total panel). Considering the two separate analyses (DAM and RBM), 37 and 14 SNPs showed higher values than the mean (17.05 % and 6.45 %), respectively. Only one SNP presented values above 2 %, with a random base incorporation rate of 0.46 %. These low overall rates of misincorporation allows better mixture detection, as the threshold for minor component allele calling can be set to 0.02.

3.2.4. Baseline noise of the MAPlex assay

A total of fourteen negative controls from reproducibility and sensitivity series were analyzed to establish a baseline noise level. All reads successfully targeting the SNP positions were analyzed and counted. A mean value of numbers of reads in these positions was established considering all samples; with counts distinguished as described allele reads or random base reads. Results are shown in Fig. 2C and Supplementary Table S4. For all SNPs, a mean of 1.56 non-specific reads was observed and a total of 178 SNPs show mean reads below the overall mean, or 82 % of markers in the MAPlex panel. Considering the markers above the mean value, only 5 SNPs consistently present more than 10 reads (2.3 % of markers) and only 2 SNPs displayed more than 20 reads, (0.92 % of markers). The mean maximum non-specific reads observed was 26.3 in SNP rs2715883, however most reads occurred in a single NTC (Supplementary Fig. S3). Supplementary Fig. S3 depicts a boxplot of the read depth per marker for all 14 NTCs, indicating all observations with higher numbers of reads are outliers, whilst the median values are close to zero. This result shows mean values alone are not a valid representation of the baseline noise. As described previously, the number of reads per SNP were also analyzed by considering two alternatives: i. a base that is described as an allele for the specific SNP, but at a low level of sequence reads, or ii. a non-specific base misincorporation. By establishing a ratio between the specific base mean reads and the total mean number of reads, a value of 0.99 was obtained, clearly indicating high levels of specific incorporation of bases at each marker's position. Even though the mean number of non-specific reads in all SNPs is very small and therefore were not expected to influence genotype calling, some caution should be taken when analyzing SNPs that consistently have more than 10–20 non-specific reads, depending on the threshold established for genotype calling, and whether this is an overall value, or applied to each SNP individually.

3.2.5. Forensic sensitivity of the MAPlex assay

The sensitivity dilution series was run four times independently with their libraries prepared and sequenced in individual S5 runs (series 1–4). Fig. 3A and Supplementary Fig. S4A show the percentage of correct genotype calls and the percentage of various inconsistencies found per mean dilution and per replicate. High percentages of correct genotypes were observed as low as 0.063 ng DNA input, with 99.1 % in mean replicates. In addition, the lower DNA input dilutions of 0.032 ng and 0.016 ng showed high percentages of correct genotype calls of 96.8 % and 81.9 % respectively (Fig. 3C). At one 0.125 ng replicate two genotype inconsistencies were observed, comprising one allele dropout and one no call due to a low number of reads (15x). Below 0.125 ng, the frequency of observed inconsistencies increased, showing a mean of 0.9 %, 3.2 % and 18 % at 0.063 ng, 0.032 ng and 0.016 ng, respectively (Supplementary Fig. S4 and Supplementary Table S5). For the lowest input dilutions of 0.032 ng and 0.016 ng, the most frequent inconsistency found was allele dropout with less loci giving no calls (both from Genotyper analysis or no reads). Dropins were found at a very low frequency and no completely incorrect genotype calls were found in all mean dilution replicates. Despite a proportion of loci that underperformed consistently in terms of number of target reads, the distribution of inconsistencies across different SNPs did not appear to be related to mean coverage.

The distribution of target reads in each replicate per dilution series followed a consistent and regular decreasing line for series one and two (Fig. 3B, all replicates). However, series 3 and 4 showed more

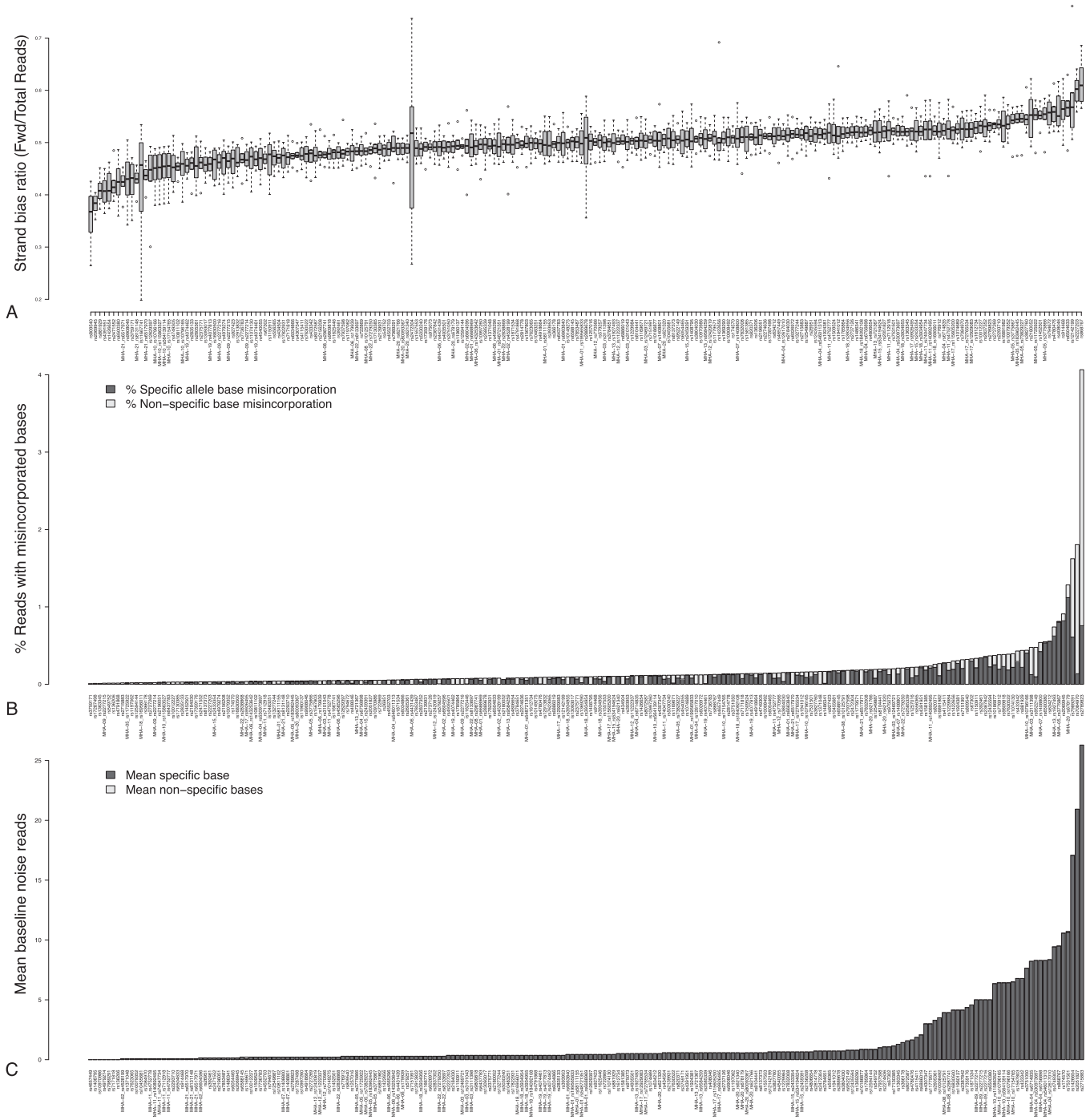

Fig. 2. A: Mean 1 ng replicates strand bias ratio per SNP. Strand bias was calculated as the ratio of the mean number of forward reads to the total number of reads at a specific SNP position. B: Percentage of mean misincorporated bases per SNP divided into expected (specific) allele base misincorporation and random (non-specific) base misincorporation. C: Mean reads per SNP found in fourteen negative controls, commonly described as baseline noise.

variability and higher coverage values in low input replicates (Supplementary Fig. S4B, Supplementary Table S6). Although series 3 and 4 had generally higher coverage values and wider coverage distribution across SNPs, no improvement in the proportion of correct genotypes in the lower DNA input dilutions was observed (Fig. 3A). Indeed, the last dilution of series 4 (4 − 0.016 ng) was the worst performing sample in the sensitivity series, reaching 77.1 % of correct genotype calls, 14.2 % of allele dropouts and 7.3 % of no calls (due to no reads). Nevertheless, it showed a mean coverage of 2063 target reads for all SNPs, whereas the other 0.016 ng replicates showed mean coverage of 283.2, 357.1 and 2249 (series one to three, respectively). Variation in coverage performance between series 1–2 and 3–4 might be due to the fact that those were run by pairs into different chips attaining different quality scores.

Considering the analysis of the microhaplotypes separately, this requires reads that align to the full extent between the bounding component SNPs. Therefore, a decrease in coverage is expected when compared to single SNP genotyping, the latter accounting for shorter reads starting in both directions. However, correct haplotypes were detected down to 0.125 ng. At 0.063 ng, one replicate showed a single allele dropout in MHA-11 and one allele dropin in MHA-20. At 0.032 ng one replicate with 1 allele dropin in MHA-10 (CTAAT) was observed

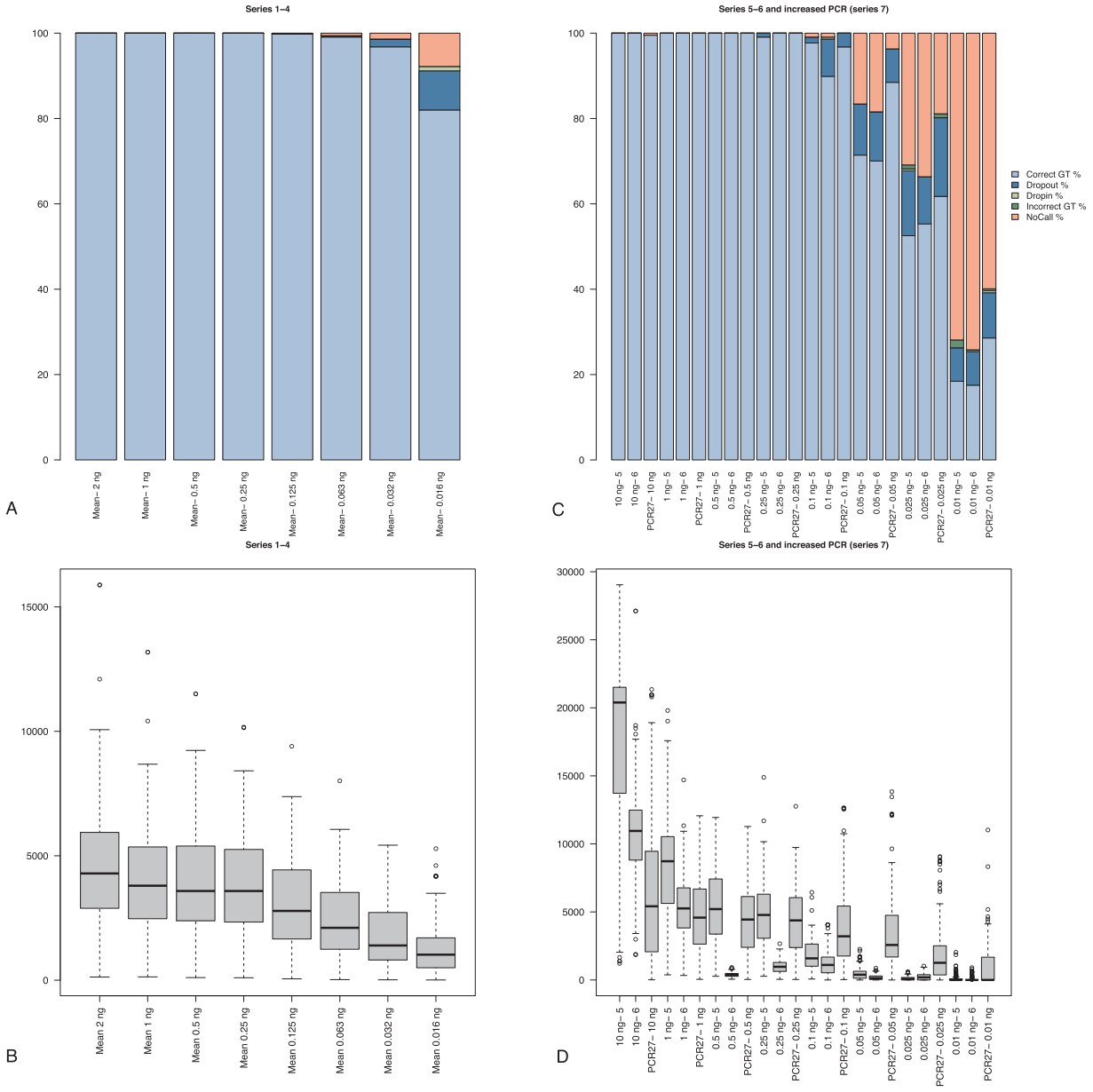

Fig. 3. A: Percentage mean of correct genotypes detected and genotype inconsistencies (allele dropouts and dropins, incorrect genotype calls and no calls) found per sensitivity series 1-4 dilution steps. B: Coverage distribution in mean replicate sensitivity dilution steps (series 1-4). C: Percentage of correct genotypes detected and genotype inconsistencies (allele dropouts and dropins, incorrect genotype calls and no calls) found per replicate of each dilution steps in sensitivity series 5-6 and increased PCR cycles – series 7. D: Coverage distribution per replicate in the sensitivity series 5-6 dilution steps and increased PCR cycles – series 7.

with 262 reads and the same replicate showed an incorrect genotype call in MHA-21 (composed of 3 SNPs only), displaying the haplotypes AGCT and CGTC (4092 and 1654 reads respectively). The expected haplotypes (ACT and CTC) showed no additional G variant site in the middle of the microhaplotype. This base was most likely a PCR error due to the low DNA input that the Microhaplotyper plugin has included as an allele in the MH-allele string. As expected, at 0.016 ng, there was an increase in the number of allele dropouts per replicate (minimum 3 allele dropouts, and one replicate showed 5 allele dropouts), as well as an increased number of allele dropins, incorrect genotype calls and complete locus dropouts. More details are provided in Supplementary Fig. S5 and Supplementary Table S7. These results indicate that the prototype version of the Microhaplotyper plugin used is an effective tool to obtain allele calls for microhaplotypes as phased strings, but requires further improvement and the need for manual revision of component SNP allele calls when analysing compromised samples.

3.3. Additional sensitivity series and initial assessment of increased PCR cycles

3.3.1. Dilution series

Two replicates of a dilution series (series 5 and 6) of 10 ng-0.01 ng were typed with the MAPlex panel and sequenced in one Ion S5 run. Sequence coverage analysis of both dilution series revealed differences between replicates, showing a general trend of lower coverage values in the second dilution series (Fig. 3D and Supplementary Table S8). Although both dilution series showed an expected decrease in number of reads following decreasing DNA input, series 6 displayed a sharp decrease in sequence coverage from 1 ng to 0.5 ng. The 0.5 ng replicate of the series 6 showed more similarities in coverage levels with the lower input DNA dilutions than with the closer dilutions of the series. A possible explanation for this difference is the fact that the dilutions series was prepared separately, and was not based on the same

concentration stocks. Although this result cannot be explained by pipetting variation alone, the sharp drop in coverage did not affect the number of inconsistencies and dropouts found, which appear comparable in both series' replicates.

Correct SNP genotype calls were observed down to 0.5 ng in mean duplicates (Fig. 3C and Supplementary Table S9). At 0.25 ng, two allele dropouts occurred in one of the replicates with a coverage of 2006 reads at the SNP target, while interestingly the second replicate with a lower number of reads (572) generated a full set of marker types. Below 0.1 ng input levels, several allele and locus dropouts occurred as described previously for the operational validation tests performed on similar forensic MPS assays [10,14].

Two dropins were observed: at 0.1 ng in rs7853487, (a CT genotype instead of TT: 261 C reads vs 546 T); and at 0.025 ng in rs12629397 (a GT genotype instead of GG: 50 T reads vs 244 G). In both cases there is evident imbalance beyond the usual proportions of allele reads observed in these SNPs. Such imbalances underline the importance of checking the major allele frequency metric for detecting erroneous genotype calls. Incorrect genotypes were called only at the lowest concentration replicates of 0.025 and 0.01 ng, and can be interpreted to represent random erroneous PCR amplification products from very low input DNA. Supplementary Table S9 details the incorrect calls and dropins observed with the dilution series DNA samples. The overall pattern of dropouts in all samples indicates no direct association with specific MAPlex loci; notably those with below average levels of sequence coverage with higher DNA inputs, but rather as a purely stochastic, and therefore random phenomenon.

3.3.2. PCR cycle number

To further study the sensitivity of the MAPlex assay we tested an increased number of PCR cycles (raised to 27) for a new dilution series (series 7) with the same input concentrations. Coverage values in the higher dilutions (10 ng-0.25 pg) were similar to the same replicates for sensitivity series 5 (Fig. 3D, Supplementary Table S8). However, when analyzing lower dilutions from 0.1 ng to 0.01 ng we observed an increase in coverage values for the increased PCR cycle. Furthermore, the sensitivity replicates with increased number of PCR cycles yielded full genotype results down to the lower level of 0.25 ng (Fig. 3C and Supplementary Table S9); increasing the sensitivity of the assay by one dilution factor. Allele dropouts started to occur only at and below 0.1 ng, further indicating increased sensitivity from additional PCR cycles. Interestingly, no dropins or incorrect genotype calls were observed when the PCR cycles were increased. These results indicate the need to further optimize PCR conditions to enhance the sensitivity of forensic MPS assays analyzing low level DNA, in this sense, increasing the PCR cycle number to 27 is a better approach to low-level DNA analysis.

3.4. Analysis of degraded DNA

Results from an enzymatically degraded time-series of 1 ng input DNA indicated full correct genotypes (Supplementary Fig. S6A, Supplementary Table S10) and high coverage values (Supplementary Fig. S6B, Supplementary Table S11) in all time-series of degradation. To assess the likely state of the DNA in these artificially degraded samples, STR analysis was run in parallel. Despite observing a decline in signal strength from the smaller to the larger amplicons (Supplementary Fig. S7), the enzymatic degradation did not appear to affect amplicons smaller than 100 bp, indicating enzymatic shearing does not provide a complete approximation of DNA degradation, and consequently the performance of the MAPlex assay with degraded DNA requires further evaluation.

3.5. Analysis of mixed DNA

The MAPlex panel was designed to provide enhanced detection of mixtures, comprising microhaplotypes [19,20] and multi-allelic SNPs [21,22] that can assist in identifying mixed DNA. Nevertheless, in routine, operational forensic DNA analysis, biogeographical ancestry inference is generally performed after STR analysis. STRs have higher levels of polymorphism than SNPs and MHs, therefore, STR analysis enables greater resolution in mixture deconvolution. The mixture detection system we propose for MAPlex is based on the combination of three different common approaches that should be taken into account together: (i) heterozygosity levels; (ii) allele imbalance; and (iii) presence of more than two alleles in multi-allelic markers. The results of applying these criteria to the set of mixed DNA samples is detailed in the following section.

The first approach commonly applied to detect mixed DNA is assessment of the heterozygosity levels of the unknown profile with values usually found on single-source samples. However, as AIM markers are specifically chosen to have contrasting frequencies between populations, this approach is much less informative when the mixture contributors are individuals of the same population. Supplementary Fig. S8 shows heterozygosity levels calculated as the percentage of heterozygous calls of the four reference Coriell cell line DNA single-source samples (left-hand and middle portions of the plot) and four pairs of artificial mixtures at three different ratios each. The heterozygosity levels of mixtures in ratios 1:1 (ranging from 54 % to 72 %) reach higher values than single-source samples (equivalent range: 23%–36%) i.e. ~1.5 times higher, including mixtures of individuals from the same population (EUR-EUR). The more imbalanced mixture ratios 1:3 and 1:9 tended to show lower heterozygosity values than the 1:1 mixture due to the lack of detection of the minor alleles when applying default parameter settings for Genotyper. When adjusting the lower threshold for the parameter min_allele_freq (minimum allele frequency, the threshold applied for reliable detection of a heterozygous genotype), the capacity to detect the minor alleles improves and the heterozygosity levels of 1:3 and 1:9 mixture ratios reach values comparable to the 1:1 mixture ratios, as shown in Supplementary Fig. S8 (right-hand portion of the plot).

When considering the genotyping accuracy of mixtures, changing the min_allele_freq parameter improves genotype concordance – calculated taking into account the individual SNP genotypes of the components in each mixture– especially when analyzing the most imbalanced mixture ratios. As shown in Fig. 4, the number of discordant genotypes was not completely reduced when changing this parameter, mostly due to the presence of a third allele in the tri-allelic SNPs in MAPlex [22]. The TFS Genotyper plugin did not allow the calling of three alleles. Additionally, the application of different criteria to manually call tri-allelic SNP alleles could not consistently distinguish the signal of a third allele from base misincorporations. For this reason, detection of mixtures based on calling three alleles will require refinement of the Genotyper plugin to allow specifying the likely number of components of the mixture and the maximum expected number of alleles in any one marker.

A second indicator of mixtures is the departure of the Allele Read Frequency value (ARF, calculated as percentage of allele reads over the total coverage of the marker) from the expected pattern found in single-source samples. Fig. 5 shows reference ARF values for single-source reference samples and four mixtures in three different ratios. Single-source ARF values clustered tightly around 0 % or 100 % for homozygotes and with little spread around the 40–60 % interval for heterozygotes. Mixtures can be easily distinguished, taking into account the broad spread of their ARF values, which aligns more closely to those of single-source samples as the mixture ratio is more extreme (i.e. it is easier to differentiate a 1:1 mixture than a 1:9 mixture).

A third indicator of mixtures is the presence of more than two alleles in multi-allelic markers. As explained above, the detection of three alleles in tri-allelic SNPs does not always properly differentiate allelic sequences from baseline levels of misincorporation. However, this panel included 20 multi-allelic microhaplotype markers, which, can be

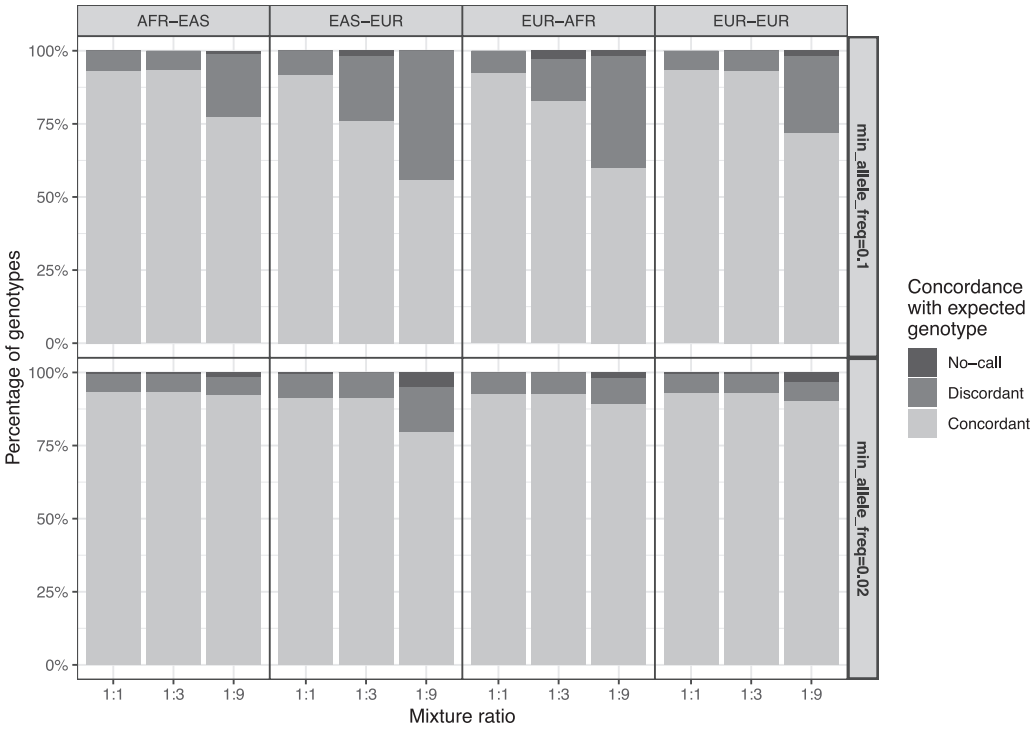

Fig. 4. Bar charts representing the percentage of concordant (light grey) and discordant (dark grey) genotype calls compared to the expected mixture and no-calls (black) in four different mixtures (AFR–EAS, EAS–EUR, EUR–AFR, EUR–EUR) at ratios 1:1, 1:3 and 1:9 genotyped with default parameters (top plot) and min_allele_freq parameter threshold set to 0.02 (bottom plot).

Fig. 5. Reference Allele Read Frequency (ARF), calculated as the percentage of allele reads over the total coverage of the marker; values of four different mixtures (AFR–EAS, EAS–EUR, EUR–AFR, EUR–EUR) at ratios 1:1, 1:3 and 1:9 and their component single-source (SS) reference samples.

genotyped and then reliably phased using the Microhaplotyper plugin. This plugin called all SNP alleles and their haplotype combinations successfully for all sequences at read frequencies higher than a pre-set relative analytical threshold (default value 0.02). For the analysis of mixtures, the relative analytical threshold of Microhaplotyper was left at 0.02 to enhance the detection of the minor component, especially in 1:9 mixtures; where the expected frequency of the minor allele is 0.05. In contrast, when calling haplotypes in single-source samples, the analytical threshold was set at 0.1 to match the equivalent min_allele_freq threshold in Genotyper. The histograms in Supplementary File S2 represent the total read number of called alleles in the 20 microhaplotypes of MAPlex for different mixture ratios and their component single-source reference samples. When considering a sample's profile as a whole, mixtures can be detected in all cases from the presence of more than two alleles in three or more markers. In the absence of a perfect balance of heterozygous markers in single-source samples, proportions of the different alleles of a microhaplotype across the different mixture ratios tended to follow the expected patterns, with the minor component alleles decreasing in proportion from 1:1 through the 1:3 to the 1:9 ratios. However, the AFR-EAS and EAS-EUR 1:1 mixture ratios do not show the expected balance between sample alleles due to the lack of precise quantification prior to mixture preparation, potentially resulting in underrepresentation of the EAS component. In the analyzed mixtures, dropouts of the minor component (indicated by the violet arrows) occurred at a low level at the most imbalanced mixture ratio of 1:9. Also, in the EUR-EUR mixture (Supplementary File 2D), a dropin was identified for MHA-09 (indicated by the red arrows). In this case, the non-matching haplotype is GGC, while the expected haplotypes are AGC and GAT. Supplementary File S1-D shows IGV captures of the microhaplotype region and highlights that the first SNP is embedded in a poly-G tract, likely to cause misalignment and lead to miscalling of a proportion of AGC haplotypes as GGC. Overall, deconvolution of microhaplotype individual profiles for imbalanced mixtures appears to be feasible, taking into account the possibility of a low level of dropouts and dropins.

4. Concluding remarks

High quality sequences were obtained for all SNPs in the MAPlex panel. The reproducibility tests confirmed the assay's sensitivity to low levels of input DNA, but with discernible differences in sequence coverage amongst the component markers. In contrast with previous assays tested on the Ion PGM platform, the Ion S5 System used to sequence the MAPlex panel provided a more balanced ratio of sequences between forward and reverse strands. Low levels of misincorporation and baseline noise appear to be inherent to the sequencing method, not the individual loci of the MAPlex panel, and generally should not influence correct genotype calling. Only the two SNPs rs2387842 and rs1422656 performed below expectations due to misalignment issues that appear not to unduly affect their genotyping accuracy.

Because of the high quality of sequences generated, the assay demonstrated high concordance, reaching levels over 99 % for both comparisons between laboratories and with online data from 1000 Genomes phase III sequence analysis using different chemistries. This validates the use of TFS software pipelines to obtain accurate genotypes from Ion S5 MPS sequence data.

In relation to the ability of the panel to analyze compromised samples, a high sensitivity to low-level DNA input was observed in both sensitivity tests, yielding full correct profiles from 0.125 ng, and more than 80 % of correct genotypes with input DNA below 0.016 ng. The effect of increasing the number of PCR cycles requires further evaluation, however, the data suggests an increase in sensitivity can be achieved by raising the cycle number, without generating dropins and incorrect genotype calls. Although further testing of highly degraded DNA samples, and particularly naturally degraded material, is necessary to gauge the performance of MAPlex, the current analysis with artificially degraded DNA provides an indication of the robustness of short amplicon MPS assays and their improved performance when compared to STRs.

Detection of mixtures was straightforward, taking into account the imbalanced allele read frequencies of SNPs and microhaplotypes with more than two haplotypes. However, improvements in the Genotyper plugin are required to reliably detect three alleles in tri-allelic SNPs. Mixture analysis and deconvolution with microhaplotypes clearly offers a promising way forward for this aspect of forensic MPS analysis, with the additional potential to infer individual biogeographical ancestry of the components of a mixture.

Overall, high genotyping performance of the MAPlex panel was observed in these studies. These results indicate that MAPlex is a balanced, well-designed panel, with high levels of sensitivity and reproducibility. The MAPlex panel will be particularly informative for forensic scenarios where estimation of the biogeographical ancestry of the contributors to a mixture is required. The further testing of MAPlex, with more extensive population samples will assist in evaluating the ancestry predictive performance of the panel, with relevance to its ability to differentiate multiple genotypes in mixed DNA. Lastly, the application of MAPlex to actual casework analyses will help further investigate the lower limits of sensitivity of this panel and forensic sequencing with the Ion S5.

Acknowledgments

MdlP is supported by a postdoctoral fellowship awarded by the Consellería de Cultura, Educación e Ordenación Universitaria and the Consellería de Economía, Emprego e Industria from Xunta de Galicia (Modalidade A, ED481B 2017/088). CP, AFA, AMM, MdlP, MVL are supported by MAPA, Multiple Allele Polymorphism Analysis (BIO2016-78525-R), a research project funded by the Spanish Research State Agency (AEI), and co-financed with ERDF funds. AFA is supported by a post-doctorate grant funded by the Consellería de Cultura, Educación e Ordenación Universitaria e da Consellería de Economía, Emprego e Industria from Xunta de Galicia, Spain (Modalidade B, ED481B 2018/010). The 1000 Genomes high coverage sequence data were generated at the New York Genome Center with funds provided by NHGRI Grant 3UM1HG008901-03S1.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fsigen.2020.102344.

References

[1] Elaine Y.Y. Cheung, Christopher Phillips, Mayra Eduardoff, Maria Victoria Lareu, Dennis McNevin, Performance of ancestry-informative SNP and microhaplotype markers, Forensic Sci. Int. Genet. 43 (2019), https://doi.org/10.1016/j.fsigen.2019.102141 102141, ISSN 1872-4973.

[2] C. Børsting, N. Morling, Next generation sequencing and its applications in forensic genetics, Forensic Sci. Int. Genet. 18 (2015) 78–89.

[3] C. Phillips, Forensic genetic analysis of bio-geographical ancestry, Forensic Sci. Int. Genet. 18 (2015) 49–65.

[4] K.K. Kidd, W.C. Speed, A.J. Pakstis, D.S. Podini, R. Lagace, J. Chang, S. Wootton, E. Haigh, U. Soundararajan, Evaluating 130 microhaplotypes across a global set of 83 populations, Forensic Sci. Int. Genet. 29 (2017) 29–37.

[5] C. Phillips, D. McNevin, K.K. Kidd, R. Lagace, S. Wootton, M. de la Puente, A. Freire-Aradas, A. Mosquera-Miguel, M. Eduardoff, T. Gross, et al., MAPlex - a massively parallel sequencing ancestry analysis multiplex for Asia-Pacific populations, Forensic Sci. Int. Genet. 42 (2019) 213–226.

[6] The Genomes Project Consortium, A global reference for human genetic variation, Nature 526 (2015) 68–74.

[7] S. Mallick, H. Li, M. Lipson, I. Mathieson, M. Gymrek, F. Racimo, M. Zhao, N. Chennagiri, S. Nordenfelt, A. Tandon, et al., The Simons Genome Diversity Project: 300 genomes from 142 diverse populations, Nature 538 (2016) 201–206.

[8] 1000 Genomes FTP - NYGC-GATK. ftp://ftp.1000genomes.ebi.ac.uk/vol1001/ftp/data_collections/1000G_2504_high_coverage/working/20190425_NYGC_GATK/.

[9] M. Al-Asfi, D. McNevin, B. Mehta, D. Power, M.E. Gahan, R. Daniel, Assessment of the precision ID ancestry panel, Int. J. Legal Med. 132 (2018) 1581–1594.

[10] M. Eduardoff, C. Santos, M. de la Puente, T.E. Gross, M. Fondevila, C. Strobl, B. Sobrino, D. Ballard, P.M. Schneider, Á Carracedo, et al., Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM™, Forensic Sci. Int. Genet. 17 (2015) 110–121.
[11] D. Sims, I. Sudbery, N.E. Ilott, A. Heger, C.P. Ponting, Sequencing depth and coverage: key considerations in genomic analyses, Nat. Rev. Genet. 15 (2014) 121–132.
[12] R. Nielsen, J.S. Paul, A. Albrechtsen, Y.S. Song, Genotype and SNP calling from next-generation sequencing data, Nat. Rev. Genet. 12 (2011) 443–451.
[13] M. de la Puente, C. Phillips, C. Santos, M. Fondevila, A. Carracedo, M.V. Lareu, Evaluation of the Qiagen 140-SNP forensic identification multiplex for massively parallel sequencing, Forensic Sci. Int. Genet. 28 (2017) 35–43.
[14] M. Eduardoff, T.E. Gross, C. Santos, M. de la Puente, D. Ballard, C. Strobl, C. Børsting, N. Morling, L. Fusco, C. Hussing, et al., Inter-laboratory evaluation of the EUROFORGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM, Forensic Sci. Int. Genet. 23 (2016) 178–189.
[15] R: a Language and Environment for Statistical Computing, (2020) http://www.R-project.org.
[16] J.T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat. Biotechnol. 29 (2011) 24–26.
[17] L.M. Bragg, G. Stone, M.K. Butler, P. Hugenholtz, G.W. Tyson, Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data, PLoS Comput. Biol. 9 (2013) e1003031.
[18] S.B. Seo, J.L. King, D.H. Warshauer, C.P. Davis, J. Ge, B. Budowle, Single nucleotide polymorphism typing with massively parallel sequencing for human identification, Int. J. Legal Med. 127 (2013) 1079–1086.
[19] K.K. Kidd, A.J. Pakstis, W.C. Speed, R. Lagace, J. Chang, S. Wootton, E. Haigh, J.R. Kidd, Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics, Forensic Sci. Int. Genet. 12 (2014) 215–224.
[20] F. Oldoni, K.K. Kidd, D. Podini, Microhaplotypes in forensic genetics, Forensic Sci. Int. Genet. 38 (2018) 54–69.
[21] C. Phillips, J. Amigo, Á Carracedo, M.V. Lareu, Tetra-allelic SNPs: informative forensic markers compiled from public whole-genome sequence data, Forensic Sci. Int. Genet. 19 (2015) 100–106.
[22] A.A. Westen, A.S. Matai, J.F. Laros, H.C. Meiland, M. Jasper, W.J. de Leeuw, P. de Knijff, T. Sijen, Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples, Forensic Sci. Int. Genet. 3 (2009) 233–241.