Performance of a massive parallel sequencing microhaplotypes assay on degraded DNA

Chiara Turchi[a,*], Filomena Melchionda[a], Mauro Pesaresi[a], Paolo Fattorini[b], Adriano Tagliabracci[a]

[a] Section of Legal Medicine, Polytechnic University of Marche, Ancona, Italy
[b] Department of Medicine, Surgery and Health, University of Trieste, Italy

ARTICLE INFO

Keywords:
Microhaplotypes
Degraded DNA
LT-DNA
MPS

ABSTRACT

Massively parallel sequencing (MPS) has allowed to analyze a new type of forensic genetic marker, known as microhaplotypes (MHs). MHs appear to be useful for identification purposes, reconstruction of family relationships, ancestry prediction and DNA mixtures deconvolution. Moreover, MHs are potentially suitable for the analysis of degraded DNA samples.

We designed a new panel of 29 MHs for MPS assay, with amplicons sizes below 180 bp and we investigated its effectiveness with low amounts of degraded samples. We genotyped a set of real forensic samples together with a set of artificially degraded DNAs. Also, a sensitivity test was assessed by a set of 2800 M DNA dilutions. The Depth of Coverage (DoC) were uniform across all 29 loci, in spite of amplicons size. Genotyping results shown that full profiles can be obtained even in highly degraded samples when the amount of template range from 0.1 to 5.0 ng. Finally, the increase of the number of PCR cycles did not provide an improvement in typing results of low amounts of degraded samples as, in front of higher number of typed loci, higher frequencies of artefacts leading to mistyping are found at 25 cycles.

1. Introduction

The recent advances in modern DNA technology have confirmed the idea that SNPs can be valuable complementary markers to the gold standards STRs. Massive parallel sequencing (MPS) offers the opportunities to genotype thousands of SNPs from multiple samples in a single experimental run and allows to explore a novel type of genetic marker known as microhaplotypes (microhaps or MH). A microhaplotype locus is defined by at least two single nucleotide polymorphisms (SNPs) closely genetically linked within the length of a sequence read and the expectation of a very low recombination rate [1]. The alleles at a microhap locus are defined by the allelic combinations of the SNPs; the alleles are referred to as haplotypes.

The values of MH marker in forensic genetic field focusing on individual identification, ancestry prediction, familial relationship reconstruction and DNA mixtures deconvolution. Moreover, due to reduced amplicons size where the MH loci lies, microhaps are potentially suitable for the genotyping of degraded DNA, with a discrimination power approaching that of commonly used STRs [2].

In previous studies, we selected 87 MH loci annotated in the Allele FREquency Database for evaluated their genetic variation in 100 Italian individuals using MPS, in order to make inference about its usefulness in forensic genetic [3,4]. To test the forensic ability of this microhaps panel for individual identification the matching probability (PI) was calculated. The resulting combined matching probability value of 87 MH loci was equal to 5.7×10^{-63} [4].

In order to better understanding the performance of microhaplotypes to challenging samples as low amounts of degraded samples, we selected a subset of 29 MH among the 87 MH previously explored. MPS panel were designed keeping the amplicons size below 180 bp, to investigate their usefulness in a set of real forensic samples together with artificially degraded DNAs.

2. Material studied, methods, techniques

We selected 29 microhaps from the previous study [4] that matched the following criteria: 1- comprised from two to five SNPs; 2- displayed Global Average Effective Number of Alleles (A_e) > 1.8; 3- maximum distance between the two distal SNPs below 115 bp. The panel comprised a total of 74 SNPs, spread across 15 of 22 human autosomes, and displayed a combined matching probability equals to 1.3×10^{-21}. PCR primers for MPS library were designed on Ion AmpliSeq Designer tool (Thermo Fisher Scientific, https://ampliseq.com/), keeping the amplicon size range between 125 to 175 bp, with the essential criterion

* Corresponding author.
E-mail address: c.turchi@univpm.it (C. Turchi).

that the SNPs included in the MH locus were amplified in a single amplicon. One primer pool for multiplex PCR reactions was designed, with amplicons range between 126 bp and 174 bp. Different types of samples were analyzed: blood samples (reference samples), bone remains, FFPE tissues together with a set of artificially degraded DNAs. Sensitivity study was performed by using the 2800 M DNA (Promega, WI, USA) as control, diluted to concentrations of 5 ng/ul, 1 ng/ul, 500 pg/ul, 100 pg/ul, 50 pg/ul and 25 pg/ul.

After assessing the DNA quantity and DNA degradation index (DI) by Quantifiler™ Trio DNA Quantification Kit in selected samples, libraries for massive parallel sequencing analysis were prepared with Precision ID Library kit according to the user guide (MAN0015830). Amplifications were performed with DNA input ranging from 5 ng to 25 pg and testing different number of PCR cycles (21 and 25). Each library was submitted to emulsion PCR by using the Ion PGM™ Hi-Q™ View OT2 Kit. The template-positive Ion PGM™ Hi-Q™ View ISPs were enriched on Ion OneTouch™ ES Instrument and sequenced on Ion PGM™ System by using Ion PGM™ Hi-Q™ View sequencing Kit, 318 chip v2 types and 200 base read mode. A total of 28 barcoded libraries were sequenced on one chip. Raw data was processed by the Torrent Suite version 5.0.4 and the reads aligned against human reference genome (GRCh37/hg19). Variant calling of the 74 SNPs and microhaplotypes allele notation was performed using SAMtools 1.3.1 (http://samtools.sourceforge.net/) and GenomeAnalysisTK-3.7. A minimum coverage of 50x was used as thresholds for reliable genotyping. For ambiguous variant positions binary alignment map (BAM) and binary alignment index (BAI) files were visualized using the Integrative Genome Viewer (IGV) packing 2.3.68 to manually assign the correct variants.

3. Results

The average coverage of sequenced libraries by MPS ranged between 94.56 and 8630 (median: 3716.5; mean: 3477.63). We also observed a good uniformity of coverage (median: 94.73%; mean: 94.63%).

The Depth of Coverage (DoC) across all loci was calculated considering separately reference (not degraded DNA) and degraded samples. Results shown a uniform DoC distribution across all 29 loci in both sample types, in spite of amplicons size.

Genotyping results of the 29 MH panel show that all reference samples display full genotypes concordant with the expected profiles (previously determined by using the former 87 MH panel).

Genotyping results of sensitivity study shown that full and reliable genotypes could be obtained when the amount of DNA template range from 5 ng to 0.1 ng. When 50 pg of DNA was amplified, 93% and 90% of correct genotypes were observed by using 21 and 25 PCR cycles, respectively. With 25 pg of input DNA 52% and 38% of correct genotypes were observed by using 21 and 25 PCR cycles, respectively.

The two artificially degraded DNA samples gave different genotyping results. One sample with very high degraded DNA (DI = nc, not calculable) were amplified with approximately 1 ng of input DNA using both 21 and 25 cycles. Full genotypes concordant with the expected profiles were obtained in both with 21 and 25 number of PCR cycles. The other artificially degraded DNA sample displayed very low DNA content (not quantifiable as showed DNA quantity below the limit of quantification, loq) and DI not calculable. Partial profiles were obtained, with only 39% and 33% of loci correctly genotyped, both at 21 and 25 PCR cycles, respectively.

The forensic specimens tested consisted of two FFPE tissue samples and two bone remains, with different DNA contents and DI. Full profiles were observed for all samples except one, where partial profile was obtained.

4. Discussion

The principal aims of this study were to validate the 29-MH panel and to investigate its effectiveness with low amounts of degraded samples. We genotyped real forensic samples together with a set of artificially degraded DNAs. Also, a sensitivity test was assessed by a set of 2800 M DNA dilutions. This study investigated the concordance typing results observed in the sensitivity test and evaluated the performance of the 29-MH panel with respect to the DNA quantity and the DNA degradation index, in order to optimize the analytical conditions of such challenging samples.

The MPS results showed a very good performance of the designed panel. All the 29 loci have been uniformly amplified and sequenced in all types of samples, without substantial differences between degraded and not degraded DNAs. Interestingly, the amplicon size didn't affect the depth of coverage, as longer amplicons displayed similar DoC values respect to smaller ones. Therefore, the 29 Mh panel results suitable also for degraded DNA typing. This observation was also corroborated by genotyping results. First, as resulted by testing reference samples, the 29 MH panel showed high genotyping accuracy. Then, as assessed by sensitivity test, full and reproducible profiles, even with highly degraded samples, could be obtained with 0.1 ng of input DNA. It is interesting to note that the increment of the number of PCR cycles does not results in increase of sensitivity or in an improvement in typing results in samples with low amounts of DNA. Conversely occurrence of drop-in and drop-out events were observed at 25 number of PCR cycles. Same results were observed both in artificially degraded DNA both in forensic samples, which confirm that the high DNA degradation level did not influence the correct genotyping. The critical parameter that affect the result, both in terms of coverage both in terms of correct genotyping is the quantity of input DNA.

5. Conclusion

The introduction of MPS technologies within the forensic community allowed exploring a new type of marker useful for the analysis of degraded DNA samples, known as microhaplotype. The performance of 29 microhaps panel was evaluated with respect to DNA quantity, DNA degradation index and typing results. The results shown that, full profiles can be obtained even in highly degraded samples when the amount of template range from 0.1 to 5 ng. Finally, the increment of the number of PCR cycles does not seem to provide an improvement in genotyping results of low amounts of degraded samples as, in front of higher number of typed loci, higher frequencies of drop-in and drop-out events leading to mistyping are found at 25 cycles.

In conclusion, the combined matching probability ($PI = 1.3 \times 10^{-21}$) of the panel and the tests carried out until now shown that the 29 MH panel could be a powerful tool for individual identification and that it results to be sensitive and reliable in degraded DNA typing. However, further validation studies are needed for better assess its effectiveness and usefulness in forensic caseworks.

Declaration of Competing Interest

None.

References

[1] K.K. Kidd, W.C. Speed, A.J. Pakstis, et al., Evaluating 130 microhaplotypes across a global set of 83 populations, Forensic Sci. Int. Genet. 29 (2017) 29–37.
[2] F. Oldoni, K.K. Kidd, D. Podini, Microhaplotypes in forensic genetics, Forensic Sci. Int. Genet. 38 (2019) 54–69.
[3] C. Turchi, M. Pesaresi, A. Tagliabracci, A microhaploltypes panel for forensic genetics using massive parallel sequencing, Forensic Sci. Int. Genet. Suppl. Ser. 6 (2017) e117–e118.
[4] C. Turchi, F. Melchionda, M. Pesaresi, et al., Evaluation of a microhaplotypes panel for forensic genetics using massive parallel sequencing technology, Forensic Sci. Int. Genet. 41 (2019) 120–127.