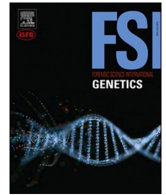
Forensic Science International: Genetics 81 (2026) 103391

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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



Microhaplotypes in forensic genetics: From exploration to application in degraded DNA specimens

Chiara Turchi [a], Filomena Melchionda [a,*], Fabiano Gentile [b], Alberto Marino [b], Domenico Colloca [b], Mauro Pesaresi [a], Andrew J. Pakstis [c], Kenneth K. Kidd [c]

[a] Section of Legal Medicine, Department of Biomedical Sciences and Public Health, Polytechnic University of Marche, Via Tronto, Ancona, AN 60126, Italy
[b] Reparto Carabinieri Investigazioni Scientifiche di Parma, Sezione di Biologia, Strada della Fonderie 10, Parma 43100, Italy
[c] Yale University School of Medicine, Department of Genetics, 333 Cedar Street, New Haven, CT 06520, United States

ARTICLE INFO

Keywords:
Microhaplotypes
Degraded DNA
Casework samples
Massive Parallel Sequencing
Human identification

ABSTRACT

Microhaplotypes have emerged as powerful forensic markers over the past decade. This paper sets out the development of a MPS panel of microhaps and its potential for application to identification, analysis of degraded DNA, ancestry inference, and identification of close biological relationships. To make it more effective when dealing with fragmented DNA, the MPS assay was designed to ensure a reduced amplicon size of less than 140 bp. After MPS assay validation, a panel of 76 microhaps, comprised of 299 different SNPs and spread across the autosomal human genome, was established. A total of 102 Italian individuals were analyzed to estimate the genotype and haplotype frequencies. The effective number of alleles at each locus (A_e) for the Italian population ranges from 1.926 to 6.187, with 59 MHs that have values greater than 3.0. The matching probability (PI) ranges from 0.055 to 0.345 and the cumulative PI value is 11.763E-66. Complete and reliable profiles were obtained with as little as 0.05 ng. The MHs panel was then validated on real forensic specimens chosen on the basis of their DNA content and degradation level. The majority of the casework samples analyzed showed complete or nearly complete MH profiles even in degraded samples. To assess the informative power of MH profiles in forensic casework, probabilistic genotyping on partial MH profiles has been used. The resulting likelihood ratio values range from 7.84E+09 to 2.70E+34, thus defining an extremely strong support for the hypothesis that the genetic profile in a casework sample comes from the reference sample. Pairwise kinship simulations using allele frequencies from Italian population samples showed that full- and half-sibling relationships can be readily distinguished from unrelated individuals. For evaluation of the 76 MH panel's utility for ancestry informativeness, PCA and STRUCTURE analyses are also presented comparing the newly collected sample from Ancona Italy with the 26 populations of the 1000 Genomes Project.

The results of the analysis confirmed the effectiveness of these short microhaplotypes in typing, with high sensitivity, samples with highly degraded DNA typically encountered in forensic cases.

1. Introduction

In forensic genetic casework, the analysis of degraded DNA samples is still a challenging issue, especially when genomic fragmentation is so extensive that PCR fails due to loss of annealing sites. In such circumstances, the amplification efficiency of short tandem repeat (STR) loci could decrease, especially for high molecular weight markers, resulting in partial profiles and in a loss of test informativeness. Forensic researchers have attempted to minimize the impact of degradation by shortening the PCR amplicon size of STRs and by using novel low molecular weight genetic markers. Sets of individual single nucleotide polymorphisms (SNPs) and insertion/deletion SNPs (Indels) [1–3] were proposed as markers for analysis of degraded DNA. However, these markers are less polymorphic than STRs and therefore a large number of SNPs/Indels are required to reach the discriminant power of STRs. Moreover, single SNPs/Indels were not very informative in the analysis

* Corresponding author.
E-mail addresses: c.turchi@univpm.it (C. Turchi), f.melchionda@staff.univpm.it (F. Melchionda), fabiano.gentile@carabinieri.it (F. Gentile), alberto.marino@carabinieri.it (A. Marino), domenico.colloca@carabinieri.it (D. Colloca), m.pesaresi@univpm.it (M. Pesaresi), andrew.pakstis@yale.edu (A.J. Pakstis), kenneth.kidd@yale.edu (K.K. Kidd).

https://doi.org/10.1016/j.fsigen.2025.103391
Received 10 July 2025; Received in revised form 1 November 2025; Accepted 12 November 2025
Available online 17 November 2025
1872-4973/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

of DNA mixtures [4].

The forensic genetic community now agrees that microhaplotypes (microhaps or MHs) are the genetic markers that combine the requirements for useful degraded DNA analysis. Kidd et al. [5] first defined microhaplotypes as small genetic markers (<300 nucleotides) consisting of at least two closely linked SNPs that present three or more allelic combinations (i.e. haplotypes). The multi-allelic nature of microhaplotypes makes them similar to STRs, but in general MHs have significant advantages compared to STRs [6]. Their short amplicon size allows analysis of degraded or low-quality DNA samples, and same-size alleles within any one MH locus avoided alleles drop-out. Furthermore, the lower mutation rate of MHs provides increased allelic stability across generations, improving the reliability of forensic and kinship inference. Finally, the absence of stutter helps to overcome common analytical challenges associated with STR profiling, particularly in complex or mixed DNA samples. Moreover, the effectiveness of the MHs for individual identification tends to correlate with the informativeness of the markers for ancestry inference, with the exception of markers with few alleles.

The introduction of massively parallel sequencing (MPS) technologies greatly enhances the utility of microhaps for forensic genetic applications. This is because MPS enables the direct resolution of the genetic phase, i.e. the cis/trans relationships among individual SNP alleles [7]. The usefulness and varied applications of microhaplotypes in forensic genetics are no longer under question. Indeed, in the last ten years several studies reported that MHs are highly informative for individual identification [8–14], testing biokinship [15–19], typing degraded DNA [9,20,21], inference of ancestry [22,23] and in DNA mixture interpretation [24,25]. It should be noted that some of these studies evaluated different forensic applications of MHs simultaneously [9,10,12,15,24,25].

To assess the feasibility of microhaplotypes for analysis of degraded DNA, several MH panels with amplicon sizes less than 200 bp have been proposed [9,20,21,24,26–30]. However, only a few of these studies showed the direct applicability of microhap loci with degraded DNA samples.

De la Puente et al. [9] developed an MPS panel comprised of 118 MHs with an amplicon size ranging between 125-175 nucleotides, that showed a high level of sensitivity when sequencing very degraded DNA. To assess the potential of the MHs panel for individual identification and mixture deconvolution, parameters such as allele frequency, gene diversity, and random match probability (RMP) were calculated from 1000 Genomes populations data (1KG), https://www.internationalgenome.org). Furthermore, to evaluate the informativity of the whole MHs panel in the resolution of parental relationships, simulations of different pedigrees were performed using 1KG European haplotype frequencies. Forensic sensitivity assessment was performed on a 2800 M DNA dilution series and a set of artificially degraded DNA samples. The results showed that complete and reliable profiles were obtained when using as little as 0.125 ng as input DNA. In addition, given the small size of the MH loci, a successful amplification with ~95 % of the correct genotypes was obtained in the sonicated samples, which had previously been tested with STR markers and had negative results. In a subsequent study [31], the same group tested the capacity of this MHs panel to infer the bio-geographical ancestry analyzing DNA data available in the 1KG and HGDP-CEPH Human Genome Diversity Panel reference datasets. Overall, the proposed MHs panel was sufficiently informative for forensic purposes, but no population data were genotyped in this study and therefore parameters such as allele frequency, gene diversity, and RMP were not inferred from real samples.

In our previous study [20], an MPS panel comprised of 29 microhap loci with amplicon sizes below 180 bp was designed with the aim to investigate its effectiveness with low amounts of degraded samples. Therefore, real forensic samples together with artificially degraded DNAs and a set of 2800 M DNA dilutions, for sensitivity test, were analysed. To investigate the performance of this MPS panel, relative Depth of Coverage (rDoC) was evaluated both for degraded and non-degraded DNA samples. The results showed that all 29 MH loci had been uniformly amplified and sequenced in both sample groups. The rDoC values were similar between the longer and the smaller amplicons; so, the degradation did not have an impact on the panel. The lowest DNA input for full profiles was 0.1 ng, even in highly degraded samples. The analysis of artificially degraded DNA and forensic casework samples confirmed that the high DNA degradation level did not prevent the correct genotyping, but the critical parameter was the quantity of input DNA that affected the results both in terms of coverage and typing. In a later publication [15], we evaluated the informativity of this 29 MHs panel in complex scenarios of kinship tests. One hundred thousand simulations for each pedigree (full-siblings, half-siblings and first cousins) were performed using the haplotype frequency estimates published in our previous article [8]. The results showed that a certain degree of overlap of likelihood ratio (LR) distribution was visible for half-siblings and first cousins' scenarios, indicating that these microhaps were unable to distinguish between related and unrelated hypotheses. Only the full-sibling relationship was readily distinguished from the unrelated condition. Therefore, this 29 MHs panel is very useful for genotyping samples with degraded DNA but not sufficiently informative in the resolution of complex parental kinship due to the small number of evaluated microhaps loci.

Zhu et al. developed a 146-plex microhaplotype MPS panel with an average fragment length of 202 bp and a maximum length of 374 bp. In this study, a single-primer extension based MPS library preparation method in combination with a flexible allele-calling strategy was used to improve the detection of MH markers for degraded DNA [21]. The forensic efficacy of the panel was evaluated using 12 low-template samples, 10 mock-degraded DNA with various degrees of degradation, and 8 forensic casework samples. Also, sensitivity experiments were performed with standard DNA NA12878 (Coriell Institute) dilutions. 103 Chinese Han individuals in Beijing (CHB) population data from the 1000 Genomes Project Phase III were used to calculate forensic parameters such as discrimination power, observed heterozygosity, A_e and linkage disequilibrium (LD). Utilising a flexible allelic calling strategy, it was observed that allelic drop-out occurred starting from 0.5 ng and locus drop-out occurred as for 0.031 ng. To evaluate the ability of the proposed assay to analyze degraded DNA, ten mock degraded samples were tested and 36 out of 146 loci could still be detected in the most degraded samples, with RMP of 10^{-5}. Overall, the results showed that the MHs system combined with the flexible allele-calling method proposed by Zhu et al. could improve the analysis of degraded DNA and gave good results in low-template DNA analysis.

In view of these studies, we decided to design a new microhap panel with smaller amplicon sizes and to genotype an Italian population sample to infer haplotype frequencies. Reliable frequency datasets are essential for weight-of-evidence calculations, making analysis of real population samples crucial. This is particularly the case when attempting to identify rare or specific haplotypes within specific populations.

For these purposes, we present a study on selecting, developing and validating a panel of 76 MHs, with amplicons length below 140 bp to allow degraded DNA typing. MHs were selected on the basis of their effectiveness in individual identification, which turns out to be informative also for kinship analysis and mixture deconvolution. Markers' ability to characterize ancestry across major continental regions was also investigated.

In order to evaluate the selected microhaps for forensic genetics, the MH panel was validated on real forensic specimens from casework, chosen on the basis of their DNA content and degradation level. Haplotype frequencies were estimated by analysing our Italian population sample.

2. Material and methods

2.1. DNA samples

Written informed consent was obtained from all living donors and the study was reviewed by the Comitato Etico per la Ricerca di Ateneo of the Marche Polytechnic University of Ancona, Italy (Prot. n. 0086284, 10/05/2023).

Blood samples were collected from 102 Italians in the general population. Human genomic DNA was extracted using the QIAamp® DNA Micro Handbook (Qiagen, Hilden, Germany) according to the manufacturer's protocol [32]. In addition, twenty-five casework samples including bloodstains (n = 6), bones (n = 9), touch DNA on forensic evidence (n = 6), fingernails (n = 2), cadaveric blood (n = 1), and cadaveric muscle (n = 1) were typed; alongside eight reference samples: cadaveric blood (n = 3), bloodstain (n = 2), buccal swab (n = 1), cadaveric tissue (n = 1) and tooth (n = 1). Casework samples were extracted using the DNA IQ™ Casework Pro Kit for Maxwell® 16 instrument (Promega, Madison, WI, USA), as described in the manufacturer's manual [33]. Bones and tooth were extracted following the Bone DNA Extraction kit (Promega) protocol [34] on the Maxwell® 16 instrument. DNA quantity and quality of forensic specimens were assessed by duplicated tests on the Applied Biosystems® 7500 real-time PCR System (Thermo Fisher Scientific, Waltham, MA, USA) using the Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific). The limit of quantification (LOQ) of the assay was at 0.005 ng/µl and, as an indicator of DNA degradation, the ratio of the quantification results of small DNA fragments (80 bp) with that of large DNA fragments (214 bp) was evaluated [35]. The 2800 M Control DNA (Promega) was used at serial dilutions (1 ng, 0.5 ng, 0.1 ng, 0.05 ng, 0.025 ng) for sensitivity evaluation.

2.2. Selection of loci

Microhaps were selected from the literature [9,20] and from the 1000 Genomes Project database (1KGP, https://www.internationalgenome.org), using the following criteria: 1- distance between the two outermost SNPs under 90 bp, 2- average A_e values > 3 among the populations (1KGP), 3- exclude Indels.

Nomenclature of the selected microhap loci was done in accordance with MicroHapDB (v 0.10) database [36] usage. All MHs were submitted to MicroHapDB and checked for SNPs compositions and name matching. When additional SNPs were present in our MHs, we decided to keep the name of the pre-existing locus in the MicroHapDB and add the rs ID number of the extra SNP (e.g. mh10USC-10pA+rs10795792). ALlele FREquency Database (ALFRED, https://alfred.med.yale.edu/) was also queried for microhaps not present in the MicrohpDB and in case of negative feedback, a novel name according to the nomenclature [37] was assigned to the microhap locus. Description of the 77 microhaplotypes panel evaluated in this study is reported in Supplementary Table S1.

2.3. Primer design for massive parallel sequencing

PCR primers for MPS libraries were designed on the Ion AmpliSeq Designer tool (Thermo Fisher Scientific, https://ampliseq.com/), under conditions such that each microhap would be amplified in a single PCR fragment and keeping the amplicon size below 140 bp to also test the panel's usefulness with degraded DNA. One primer pool for multiplex PCR reactions was designed, with insert size (i.e. the DNA region between the two PCR primer's 3′ ends) ranging from 62 to 105 bp.

2.4. MPS libraries preparation and sequencing

Preparation of DNA libraries was manually performed using the Precision ID Library kit (Thermo Fisher Scientific) kit according to the manufacturer's protocol (MAN0017767, rev C.0) but in half-volume reactions. The library-PCR reaction consisted of 2 µl of 5X Ion AmpliSeq™ HiFi Mix, 5 µl of 2X Ion AmpliSeq™ Primer Pool, variable volume of DNA and nuclease-free water, to reach the final volume reaction of 10 µl.

For population study, 1 ng of input DNA was amplified according to manufacturer's protocol; conversely, for casework samples amplification was performed with DNA input ranging from about 0.078 ng to 1 ng.

Sensitivity evaluation of the assay was performed by testing in duplicate dilutions of 2800 M Control DNA (Promega) for DNA input concentrations of 1 ng, 0.5 ng, 0.1 ng, 0.05 ng and 0.025 ng.

Thermal cycling was performed on the Veriti™ 96-Well Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific) using the following conditions: enzyme activation for 2 min at 99 °C, amplification for 22 or 25 PCR cycles of 15 s at 99 °C and 4 min at 60 °C. Negative amplification controls (no template) were sequenced for each PCR cycle. After PCR reaction, primer sequences were partially digested by adding 1 µl FuPa Reagent and incubated for 10 min at 50 °C, for 10 min at 55 °C, for 20 min at 60 °C. For ligation of microhaps libraries with adaptors, 2 µl Switch Solution, 1 µl DNA Ligase, 1 µl diluted Ion Xpress™ Barcode Adapters were added into 11 µl digested PCR reaction, and the mix was then incubated for 30 min at 22 °C, 10 min at 72 °C. After ligation, each library was purified with Agencourt™ AMPure™ XP Reagent (Beckman Coulter) and finally quantified using Ion Library TaqMan® Quantification kit (Thermo Fisher Scientific) following manufacturer's protocols. Libraries were sequenced on Ion GeneStudio™ S5 System (Thermo Fisher Scientific), performing the libraries enrichment and chip loading steps using the Ion Chef ™ Instrument (Thermo Fisher Scientific) and the Ion S5™ Precision ID Chef & Sequencing Kit (Thermo Fisher Scientific). An appropriate volume of each manually prepared library, adjusted to 40 pM concentrations, was combined to create an approximately equimolar pooled library for the Ion Chef ™ Instrument (Thermo Fisher Scientific) and it was loaded on Ion 520™ (Thermo Fisher Scientific).

2.5. MPS data analysis

The alignment of reads against human reference genome (GRCh37/hg19) was performed on the Torrent Suite v.5.12.3. Coverage analysis was carried out by the Coverage Analysis (v5.12.0.0) plugin, which provides information about mapped reads, on-target percentage and mean depth of coverage downloadable for each sample library. Haplotypes were called using the HID_Microhaplotype_Research_PluginV1.5 (Thermo Fisher Scientific) and was also used to analyze the sequencing reads of each library and output display. The plugin was run with default settings as: "min total read coverage per position" = 20, "min # (number) of allele count to include in report" = 5, "min allele frequency (heterozygous)" = 10; "min allele frequency (homozygous)" = 90. The plugin reported the detailed information on genotype, coverage, allele sequence for each MH locus in a csv file. Also, the report containing the BAM/BAI files and another version of csv, named "norare", which contained a condensed version of the genotype of the sample. The software Integrative Genomics Viewer (IGV, v.2.8.0, Broad Institute and UC San Diego) [38] was used to check the compliance of target regions sequenced and to verify the MH calls flagged by the plugin.

The relative depth of coverage (rDoC) across all target sequences, calculated as the ratio of depth of coverage (DoC) at single amplicon to total DoC of the sample, was obtained in order to evaluate the performance of the designed panel.

2.6. Biostatistic analysis

Haplotype and genotype frequencies, Hardy-Weinberg equilibrium, observed heterozygosity, and gene diversity (expected heterozygosity) were calculated using PowerMarker v3.25 software [39]. Matching probability (PI) for each MH was calculated as the sum of the squares of

the genotype frequencies and then the product of these values across all MH included in the panel was obtained. The effective number of alleles (A_e) at a MH locus was calculated according to the formula: $A_e = 1/\sum p_i^2$ (pi is the allele/haplotype frequency of i_{th} allele/haplotype in one population).

In order to assess the statistical power of the microhaplotype panel for kinship analysis, simulations of complex kinship testing scenarios were performed with Familias v 3.2.6 software [40]. Pedigrees involving pairwise relationships, including full siblings, half siblings and first cousins, were tested by using the haplotype frequencies estimated in this population study. The number of simulations for each pedigree was set to one hundred thousand and the resulting distributions of likelihood ratios (LRs) for each kinship hypothesis vs. the values for the unrelated hypothesis were plotted by using R v4.0.3 (http://www.rproject.org/2019).

The statistical power of the microhaplotype panel in forensics was evaluated also against data in the 1000 Genomes (1KG) Consortium. The defining SNP genotypes for the selected MHs were retrieved from the corresponding VCF files of the Ensembl browser (release 112, https://www.ensembl.org) for the 2504 individuals in the 26 populations in the 1KG consortium. Microhaplotypes were then generated with the PHASE v2.1.1 program [41,42] and from these, the individual population MH-allele frequencies, the A_e values and Rosenberg's Informativeness (I_n) [43] were calculated for the 26 populations.

For forensic samples that showed a partial MHs profile, the likelihood ratio (LR) value was calculated using the LRmix Studio software (v.2.1.5) [44]. The data generated from 102 Italian individuals were used as reference population data in the calculation of the LR values. Drop-out probability was estimated by using drop-out estimator plus sensitivity analysis of LRmix Studio software.

PAST software (v. 4.15) [45] was used to generate two-dimensional plots of the first three PC's in the Principal Component Analysis (PCA).

Finally, population structure was analysed by STRUCTURE software (v.2.3.4) [46] using as input the numeric genotype profiles for the microhaps based on the individuals in the 26 population of 1KG consortium. Parameters used are: standard admixture model, 10000 burn-ins, 10000 iterations, 20 runs at each K-level (K=2–9).

3. Results and discussion

3.1. Microhaplotypes panel design and features

A total of 105 microhaplotypes were initially selected and then submitted to primer design by using the Ion AmpliSeq Designer tool (Thermo Fisher Scientific).

All DNA markers included in the designed panel were screened for the presence of additional SNPs with MAF > 0.1 in 1KG (http://grch37.ensembl.org/index.html) and Indels, which could contribute partially to the MH definition. The final microhaps panel comprised of 77 microhaps, which included 304 SNPs mapping in all autosomes except chromosome 15. All information of the MHs loci typed in the present study, including chromosome position and reference SNP ID number ("rs#") related to dbSNP (https://www.ncbi.nlm.nih.gov/snp/) were summarized in Supplementary Table S1. Moreover, Supplementary Figure S1 shows the ranking of the 77 MHs panel by Average A_e for the 26 populations of the 1KG Consortium. Our results showed that there is no strong correlation between the number of SNPs and the molecular extent (bp) of microhaps (Supplementary Figure S2). This indicates that SNP density is independent of sequence length, so longer microhaplotypes are not necessarily more genetically informative than shorter ones. More detailed information on this section can be found in the Supplementary Materials, Supplementary Table S2 and Supplementary Figure S3.

3.2. MPS panel performance

The analysis of MPS data of population samples displayed a good uniformity of depth of coverage (> 90 %) for the majority of loci. High average number of sequence reads (> 300 reads) were obtained for most MHs; however, five markers showed low read mean values (< 300 reads). To assess the performance of the designed panel the relative depth of coverage (rDoC), was evaluated (Fig. 1). The locus that showed the lowest rDoC value was mh06USC-6pB+rs3001007, with an average number of sequence reads about 17, and therefore this locus was excluded from the final panel. This adjustment leading to a final core panel of 76 MHs loci and 299 SNPs (Supplementary Table S3). De la Puente [9] also reported the locus mh06USC-6pB+rs3001007 characterized as underperforming in terms of sequence coverage. Poor primer performance likely explains this result and upward adjustments of individual primer ratios or new primer design could be attempted for better coverage balance. For now, we are not able to check if the PCR primers used in this panel are the same or similar to those used by De la Puente [9]. This possibility appears to be rather remote, given that a novel primer design was undertaken in the present study. Therefore, it is also plausible that the genomic region in which the 6pB+rs3001007 marker maps is somehow hindering to PCR.

The remaining mh02CT-001, mh12USC-12qA+rs7954318, mh18USC-18qC and mh09USC-9qB loci, although they are underperforming in terms of sequence coverage, were not discarded but have been carefully monitored in the genotyping phase as they displayed a higher probability of allele drop-out or locus drop-out, as explained in the section below. In general, an optimization of primer concentration could be provided for these loci in order to improve the uniformity of depth of coverage across the individual MHs in the panel. These MHs should be manually checked when typed to avoid errors in data interpretation.

Moreover, the observed locus balance did not correlate with the amplicon size, but it rather reflected the efficiency of the primer sets designed for the multiplex PCR.

To evaluate if the DNA degradation status can affect the performance of the panel, the rDoC for casework samples was also evaluated (Fig. 2).

The rDoC distribution in casework samples was consistent with the reference samples in all loci, except one microhaplotype (mh16USC-16pA+rs12325410) that showed an increase of rDoC value in casework samples. To investigate the reasons for this irregular value of rDoC shown by the microhap mh16USC-16pA+rs12325410, an in-depth analysis data was carried out with the IGV software [38] comparing a reference sample and a casework sample (Supplementary Figure S4). We noted the occurrence of short reads aligned in the 3' end of the target region in both samples; however, this phenomenon is more accentuated in the casework samples when a low DNA input was used for library amplification. We considered that these short reads probably represent primer dimers which were not filtered by the Torrent Suite Software (v.5.12.3, Thermo Fisher Scientific) and whose alignment partially overlaps with the nucleotides at the end of the target region, defined by the panel-designed bed file. These dimers are more present in samples amplified with a low amount of DNA; however, they do not affect the correct genotyping of the microhaplotype as they are filtered by HID_Microhaplotype-Research_Plugin (v1.5, Thermo Fisher Scientific).

3.3. Sensitivity study

A decreasing amount of 2800 M Control DNA (Promega), ranging from 1 ng to 0.025 ng, was used as template DNA to test the sensitivity of the designed panel. Concordant results were obtained between the 2800 M's library replicates, both in terms of library concentration and reading depth. As expected, a reduction in library concentration was observed with the decreasing DNA input, but the results were always higher than the 40 pM required for clonal amplification on the Ion Chef™ System.

Full and reliable profiles were obtained with an amount of DNA as small as 0.050 ng. When using 0.025 ng of input DNA, allelic drop-out events were observed. In the first replicate one allelic drop-out event

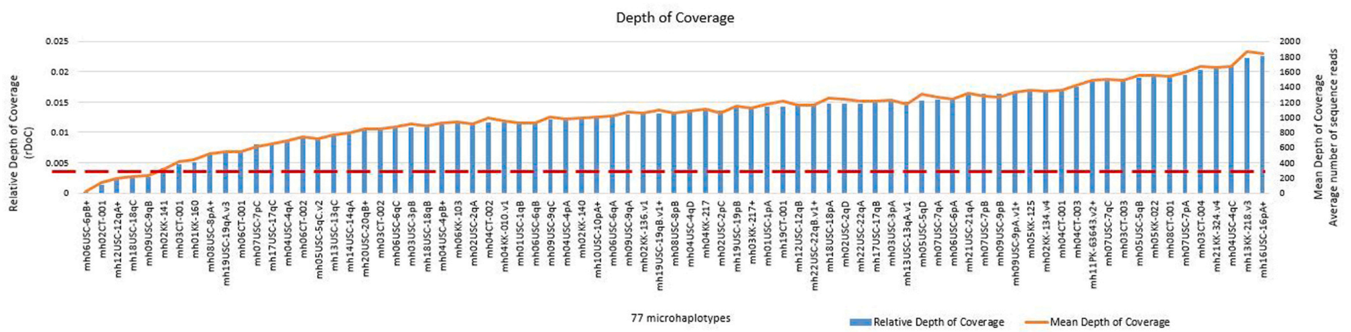


Fig. 1. Assessment of depth of coverage in the 77 MHs panel, with the rDoC values in the right axis and the average number of sequence reads in the left one. Microhaplotypes with the "+ " symbol are those with extra SNPs. The red dashed line represents the threshold of 300 reads.

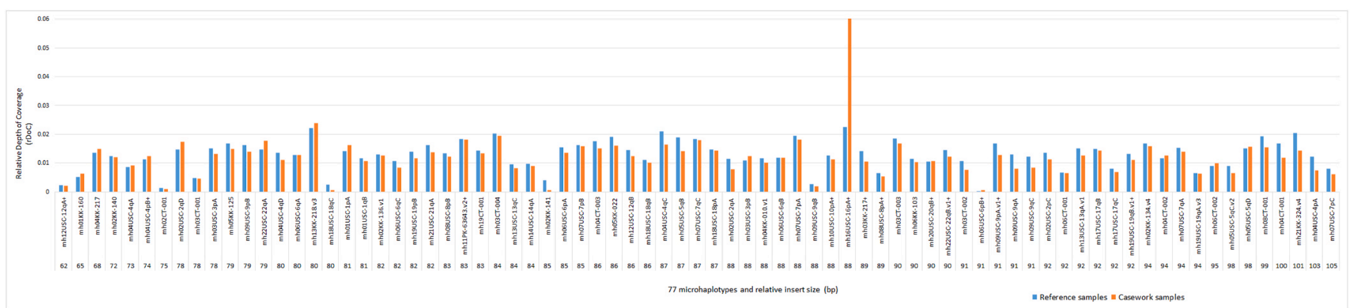


Fig. 2. Comparison of average rDoC distribution observed in reference samples and casework samples. X-axis: mean rDoC value; Y-axis: 77 microhaplotypes sorted by increasing insert size (i.e. the DNA region between the two PCR primer's 3′ ends). Microhaplotypes with the "+" symbol are those with extra SNPs.

(mh01USC-1qB) left 98.6 % correct genotyping; in the second replicate, two allelic drop-out events (mh02USC-2qA and mh18USC-18qC) left 97 % of the loci correctly genotyped. Full and reproducible profiles were obtained with 0.050 ng DNA, while samples with lower input should be replicated to confirm results.

3.4. MHs panel individualization informativeness

The haplotype (i.e., allele) frequencies calculated from MPS data from 102 unrelated Italian samples, as well as the genotype frequencies, together with other forensically important statistics are available in Table 1, Supplementary Figure S5, Supplementary Table S4 and S5.

The sample size analysed in this study may represent a limitation, anyway no significant differences were observed relative to haplotypes frequency data from 1000 Genomes EUR.

In the Supplementary Figure S5, the majority of microhaps exhibited 4 and 5 alleles (20 and 27 MHs, respectively). The frequencies of the 408 alleles ranged from 0.005 to 0.666. Starting from haplotype frequencies, we calculated the effective number of alleles at each locus (A_e). Overall, the A_e values for the 76 microhaps range from 1.926 to 6.187, with 59 MHs that have values greater than 3.0. The distribution of 76 microhaps by their A_e values is illustrated in Supplementary Figure S6.

Genotypes for all 76 MHs were tested for deviation from Hardy-Weinberg Equilibrium (HWE test) and ten MHs were significant at the 0.05 level but none remained significant after a Bonferroni correction test. To exclude variant calling errors all samples were checked on IGV software for genotype consistency at these ten MH loci and no genotyping error was detected after IGV inspection. The gene diversity values for the 76 microhaps range from 0.481 to 0.838, with a mean value of 0.707 and a median equal to 0.718. The matching probability (PI) ranges from 0.055 to 0.345 and the cumulative PI value is 11.763E-66.

3.5. Casework samples analysis

Thirty-one biological specimens from real forensic casework, which previously provided partial STR typing results, were genotyped to evaluate the performance of the 76 MHs panel on challenging samples.

DNA quantification using the Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific) shows a very low level of DNA, i.e. below the LOQ of 0.005 ng/μl, in six samples (four touch DNA samples, one cadaveric muscle, and one skeletal remain). For these samples, the maximum allowed volume (3 μl) of DNA input was used for libraries amplification. DNA input ranging from 0.077 ng to 1 ng was used for library preparation of the remaining samples (Supplementary Table S6).

DNA fragmentation was also evaluated in these samples by estimating the degradation index (DI) from the Quantifiler™ Trio DNA Quantification Kit data (Supplementary Table S6). The results indicated that five samples show a significant degradation (DI > 10), while most of these samples have moderate DNA degradation (DI from 1 to 10), and two samples don't seem to be degraded (DI < 1).

MPS library quantification using Ion Library TaqMan® Quantification kit (Thermo Fisher Scientific) showed that all casework samples displayed values above 40 pM, i.e. the library concentration required for the sequencing assay (Fig. 3), except one sample which displayed a value of 30.54 pM. Nevertheless, this sample (Sample 17 in Supplementary Table S6) was submitted to MPS run and resulted in a complete MH profile, with high mean depth values and therefore a library underestimation for this sample has been assumed. All libraries were loaded on a single sequencing chip.

The genotyping results of the MHs panel showed that the degree of DNA degradation did not influence the library amplification and target sequencing (Supplementary Table S6). Instead, it seems that the critical factor was DNA input, not degradation level. A more detailed analysis of the five samples with the highest degradation index indicates that three of these samples (Sample 7, Sample 20, and Sample 22) showed a complete MH profile. Conversely, the remaining two samples (Sample 1

Table 1
The table lists the summary statistics of the 102 Ancona samples on 76 microhaps evaluated in this study. The chromosome number is reported in the first column, followed by microhaplotype name, the number of SNPs included in each microhaps locus. The statistical values reported in the remaining columns are: number of genotypes and haplotypes observed, Major Allele Frequency (MAF), Effective number of Alleles (Ae), Gene Diversity and Matching Probability (PI).

#Chr	Microhaplotype name	#SNPs	#Genotype	#Haplotype	MAF	Ae	Gene Diversity	PI
1	mh01USC-1pA	4	9	4	0.338	3.244	0.692	0.162
1	mh01KK-160	4	10	5	0.510	2.740	0.635	0.203
1	mh01USC-1qB	4	14	6	0.422	3.582	0.721	0.140
2	mh02KK-140	4	15	5	0.270	4.588	0.782	0.083
2	mh02KK-141	4	10	5	0.387	3.311	0.698	0.158
2	mh02CT-001	4	16	7	0.348	4.014	0.751	0.113
2	mh02USC-2pC	4	15	5	0.240	4.788	0.791	0.087
2	mh02USC-2qA	4	10	4	0.314	3.905	0.744	0.125
2	mh02KK-134.v4	3	27	8	0.221	6.187	0.838	0.055
2	mh02USC-2qD	3	10	4	0.314	3.848	0.740	0.118
2	mh02KK-136.v1	3	16	6	0.299	4.448	0.775	0.091
3	mh03CT-001	3	8	4	0.475	2.759	0.638	0.202
3	mh03USC-3pA	4	15	5	0.377	3.763	0.734	0.139
3	mh03KK-217 +rs1485138	3	7	5	0.588	2.088	0.521	0.315
3	mh03CT-002	4	10	4	0.397	3.392	0.705	0.144
3	mh03USC-3pB	5	11	5	0.505	2.918	0.657	0.176
3	mh03CT-003	4	14	5	0.515	2.952	0.661	0.159
3	mh03CT-004	4	12	6	0.559	2.480	0.597	0.225
4	mh04KK-010.v1	2	10	4	0.436	3.129	0.680	0.165
4	mh04CT-001	3	6	3	0.667	1.926	0.481	0.345
4	mh04USC-4pA	4	16	7	0.397	3.897	0.743	0.111
4	mh04USC-4pB+rs16992506	4	10	4	0.333	3.597	0.722	0.149
4	mh04USC-4qA	3	10	4	0.309	3.706	0.730	0.118
4	mh04USC-4qC	4	6	3	0.417	2.717	0.632	0.230
4	mh04KK-217	3	13	6	0.456	3.296	0.697	0.142
4	mh04USC-4qD	4	12	5	0.333	3.690	0.729	0.126
4	mh04CT-002	5	21	10	0.328	4.227	0.763	0.096
4	mh04CT-003	4	12	6	0.353	3.664	0.727	0.131
5	mh05KK-125	4	10	5	0.373	3.407	0.707	0.148
5	mh05KK-022	2	6	3	0.368	2.982	0.665	0.180
5	mh05USC-5qB	4	14	5	0.382	3.902	0.744	0.106
5	mh05USC-5qC.v2	4	17	7	0.289	4.334	0.769	0.110
5	mh05USC-5qD	3	10	4	0.358	3.465	0.711	0.132
6	mh06USC-6pA	3	9	4	0.338	3.535	0.717	0.136
6	mh06KK-103	4	13	6	0.451	3.131	0.681	0.158
6	mh06CT-001	4	13	6	0.574	2.599	0.615	0.184
6	mh06CT-002	4	14	5	0.309	4.078	0.755	0.112
6	mh06USC-6qA	3	9	4	0.353	3.250	0.692	0.168
6	mh06USC-6qB	4	11	5	0.382	3.206	0.688	0.156
6	mh06USC-6qC	3	9	4	0.397	3.206	0.688	0.146
7	mh07USC-7pA	4	11	5	0.314	3.866	0.741	0.118
7	mh07USC-7pB	4	27	8	0.284	5.610	0.822	0.057
7	mh07USC-7pC	5	11	6	0.627	2.189	0.543	0.264
7	mh07USC-7qA	3	10	4	0.363	3.366	0.703	0.152
7	mh07USC-7qC	5	14	5	0.368	4.065	0.754	0.113
8	mh08USC-8pA+rs73660013	5	14	7	0.382	3.855	0.741	0.116
8	mh08CT-001	5	15	8	0.392	3.440	0.709	0.146
8	mh08USC-8pB	5	11	5	0.613	2.293	0.564	0.244
9	mh09USC-9pA.v1 +rs1535837	6	25	10	0.348	4.539	0.780	0.080
9	mh09USC-9pB	4	6	3	0.461	2.777	0.640	0.220
9	mh09USC-9qA	3	10	4	0.485	3.079	0.675	0.161
9	mh09USC-9qB	4	9	5	0.412	2.920	0.658	0.183
9	mh09USC-9qC	5	17	7	0.490	3.183	0.686	0.136
10	mh10USC-10pA+rs10795792	4	10	4	0.456	2.953	0.661	0.195
11	mh11PK-63643.v2 +rs2420397	5	14	5	0.319	3.950	0.747	0.110
12	mh12USC-12qA+rs7954318	4	10	4	0.347	3.649	0.726	0.128
12	mh12USC-12qB	4	22	8	0.348	4.618	0.783	0.082
13	mh13USC-13qA.v1	4	23	9	0.235	4.941	0.798	0.077
13	mh13KK-218.v3	3	23	8	0.299	5.272	0.810	0.063
13	mh13USC-13qC	4	11	5	0.441	3.290	0.696	0.147
14	mh14USC-14qA	3	11	5	0.397	3.550	0.718	0.143
16	mh16USC-16pA+rs12325410	5	13	5	0.520	2.866	0.651	0.181
17	mh17USC-17qB	3	13	5	0.436	3.369	0.703	0.138
17	mh17USC-17qC	4	9	4	0.368	3.214	0.689	0.188
18	mh18USC-18pA	5	13	5	0.456	3.384	0.704	0.128
18	mh18USC-18qB	3	9	4	0.397	3.208	0.688	0.162
18	mh18USC-18qC	4	10	4	0.363	3.590	0.721	0.127
19	mh19USC-19pB	5	12	5	0.338	3.702	0.730	0.117
19	mh19CT-001	4	11	5	0.475	2.598	0.615	0.226
19	mh19USC-19qA.v3	5	18	7	0.392	3.855	0.741	0.107
19	mh19USC-19qB.v1 +rs10404319	4	16	6	0.309	4.407	0.773	0.094
20	mh20USC-20qB+rs62197861	5	10	4	0.314	3.746	0.733	0.127

(*continued on next page*)

Table 1 (*continued*)

#Chr	Microhaplotype name	#SNPs	#Genotype	#Haplotype	MAF	Ae	Gene Diversity	PI
21	mh21USC-21qA	4	12	5	0.319	3.826	0.739	0.115
21	mh21KK-324.v4	3	15	7	0.309	3.957	0.747	0.123
22	mh22USC-22qA	4	11	5	0.324	3.917	0.126	0.126
22	mh22USC-22qB.v1 +rs4925401	6	23	9	0.353	4.627	0.082	0.082

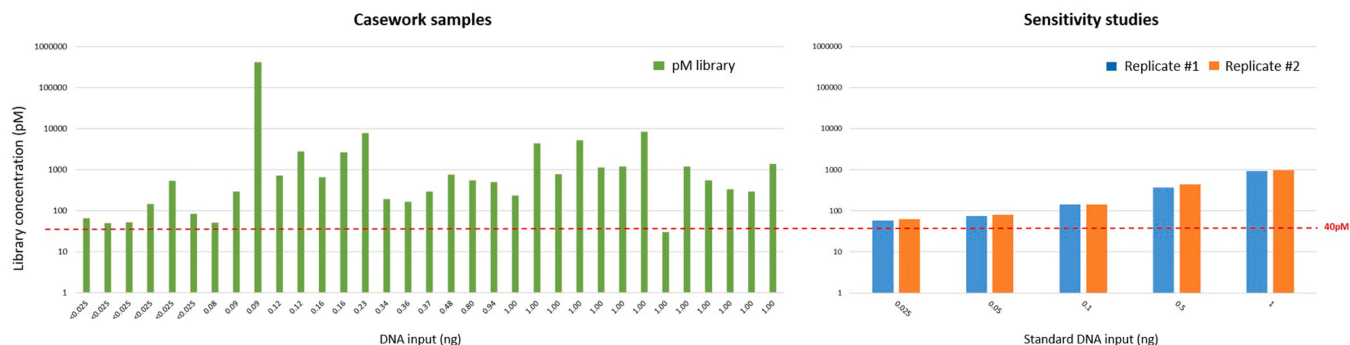


Fig. 3. The plots show the library concentration values obtained for the casework samples and for the 2800 M dilutions (sensitivity study). The red line marks the library concentration required and used for clonal amplification (40 pM) on the Ion Chef™ system. The x-axis indicates the input DNA (ng) used for library amplification, while the y-axis indicates the library concentration.

and Sample 21) reported a partial MHs typing, but they also had a lower amount of DNA. The same observations also apply to samples with moderate degradation. Therefore, we can confirm that, as observed in the rDoC analysis (Fig. 2), the performance of the designed MHs panel is not influenced by the level of DNA degradation, but rather by the amount of input DNA used for library preparation.

The majority of the casework samples analyzed showed complete or nearly complete MH profiles even in degraded samples, with 100 % and 92–97 % of MHs typed, respectively. Note that, in samples with almost full profiles (92–97 % of MHs typed) the missing MHs loci are those underperforming in rDoC analysis. Partial MH profiles (< 84 % loci genotyped) and allelic drop-out were observed in samples showing the amount of DNA below the LOQ and different degrees of degradation (Supplementary Table S6).

It would be worthwhile to ascertain the forensic informativeness of these partial profiles. The interpretation of complex DNA profiles, i.e. profiles characterized by allelic drop-out, allelic drop-in and degradation or inhibition, remains a significant challenge in forensic analysis. In this case, a likelihood ratio (LR) approach is required for reliable interpretation of these types of profiles [47]. In this study, the probabilistic genotyping on partial MHs profiles with less than 71 % of MHs typed, has been used to calculate LRs and assess the informative power of MH profiles in forensic casework. Likelihood ratio values were calculated using the statistical software LRmix Studio (v.2.1.5) by using allele frequencies estimated from Italian population samples described above. This software enables the estimation of a range of plausible values for the probability of drop-out and to include them in the calculation of the likelihood ratio. The LR value was calculated in order to assess the hypothesis that the observed profile is the same as the reference sample. The resulting LR values range from 7.84E+09 to 2.70E+34 (Table 2), thus defining extremely strong support for the hypothesis that the MH profiles in casework samples come from the reference sample [48]. It is important to note that the probabilistic genotyping analysis was conducted merely in an exploratory mode to assess the effective use of the observed MH partial profiles.

Table 2
Forensic informativeness of some partial MH profiles observed in this study. The LR values were estimated using the maximum drop out values identified by using LRmix Studio software.

Sample	% Allelic drop out	% loci genotyped	LR
Sample 1	47.83 %	61 %	9.54E+ 23
Sample 8	42.11 %	50 %	4.67E+ 16
Sample 9	65.38 %	34 %	7.84E+ 09
Sample 12	35.19 %	71 %	2.70E+ 34
Sample 18	31.48 %	71 %	7.28E+ 21

3.6. *Kinship analysis*

Kinship simulations of complex kinship testing scenarios were performed with Familias v 3.2.6 software [40] by using 76 MH allele frequencies estimated from Italian population samples. The outcomes of the kinship analysis simulations for full siblings, half siblings and first cousins are represented in Fig. 4.

Considering the 76 MHs panel, the distribution of LR of full siblings and unrelated individuals do not overlap, with average log10 (LRs) for related individuals of 14.15 and for unrelated individuals of −12.73. Even for half siblings' scenario the LR distribution clearly separate related and unrelated individuals, with average log10 LRs of 3.57 for related pairs and −3.46 unrelated pairs. For first cousin simulation the LR distribution plot displayed some degree of overlap, with average log10 LRs of 0.94 for related pairs and −0.86 unrelated pairs. The extent of overlap in the LR distributions of first cousins clearly suggests the need to analyze additional markers to resolve this degree of kinship. This panel provides greater discriminatory power than our previous 29 MH panel [20], allowing the distinction of both full and half siblings.

Overall, our results are comparable to those obtained by other research groups. In the study developed by Tomas et al. [18], 125 samples from Danish individuals and 167 samples from Greenlanders were analyzed using the Ion AmpliSeq™ MH-74 Plex Research Panel (Thermo Fisher Scientific). Simulation and LRs calculations for three degrees of kinship scenarios: full siblings (first degree), half-siblings (second degree), and first cousins (fourth degree) vs unrelated individuals were tested. The distribution of Log10(LRs) of related and unrelated individuals did not overlap for full and half siblings, which indicated that the MH panel could differentiate these kinship relationships. However, the degree of overlap in the Log10(LRs) distributions of first cousins suggested the need to analyze additional markers. Gu et al. [14] proposed a novel panel of 33 microhaplotype loci which could provide new insights into forensic complex kinships testing. In the

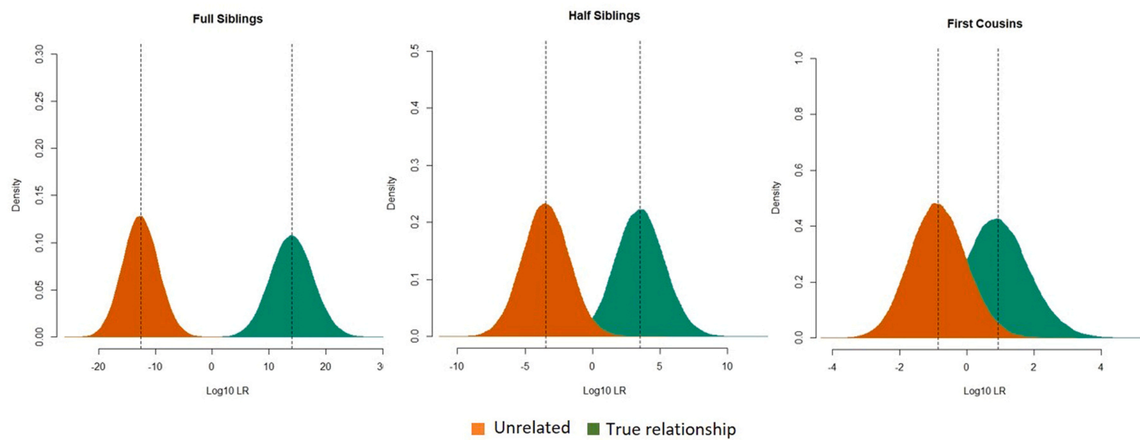


Fig. 4. Log10 (LRs) distributions obtained from pedigree tests for the supposed true relationships plotted against unrelated conditions.

simulated kinship analysis, the 33-plex panel has effectively distinguished 1st degree (parent-child, half-siblings) and 2nd degree (grandfather-grandson and aunt-nephew) parental relationships and unrelated individuals. Similarly, Du et al. [17] developed a 188-plex panel for the identification of 2nd and 3rd degree kinship. The panel was tested on 30 extended families consisting of 179 2nd degree and 121 3rd degree relatives. They found that 188 microhaplotypes could be used to solve the 2nd degree identification, whereas these loci were not enough to assess 3rd degree kinship. Bai et al. [24] developed a 185 MHs panel and demonstrated its application for kinship analysis testing different scenarios of relationship inference. Based on their results, the 185 MHs panel was suitable for parent-offspring, full-sibling and 2nd degree relatives testing. For more distant relatives (3rd degree relatives) a considerable overlap of the distributions of Log10 (LRs) was found; however, there was a clear improvement when data from one well-chosen extra relative were available. In our previous study [20], kinship simulations were performed on two microhaplotype panels composed of 87 and 29 MHs, respectively. In agreement with the other studies conducted, the 87 MHs panel readily distinguished full siblings and half siblings relationships relative to the unrelated condition. However, a certain degree of overlap was visible for first cousin relationship. In the analysis of 29 MHs panel, LR distributions overlapped for the half siblings and first cousins were reported; only for full siblings relationship was the LR distribution readily distinguished. We reported that the 87 MHs panel results more informative in resolution of kinship scenario respect the 29 MH panel and, for the latter, the lower informative power is due to the smaller number of microhaps present in the panel. Similar results were obtained by de la Puente [9] and colleagues in which the results of kinship analysis simulations for full siblings, half siblings, first cousins and second cousins showed a no overlap in the distribution of likelihood ratio (LRs) for half and full siblings.

Overall, the kinship analysis on our 76 MH panel is consistent with the current literature; the degree of overlap in the LR distribution of first cousins suggests the need for additional markers for better discrimination among relatives.

3.7. Comparative analysis with 26 populations in the 1000 Genomes project

The informativeness data of the 76 MHs in the Italian population typed in this study (ANC) has been compared with the 26 populations in the 1000 Genomes (1KG) Consortium. Comparative statistics for each of the 76 MH based on 27 populations can be found in Supplementary Table S7; the table includes effective allele number (Ae) and Rosenberg's Informativeness (In). Cumulative match probabilities and average Ae values for the 76 MH for each of the 27 populations can be found in Supplementary Table S8.

The A_e values have been explored in comparative analyses with the molecular extent (bp) of the microhaps, computed by subtracting nucleotide positions of initial from final MH defining SNPs and then adding one. Our results showed that there is no strong correlation between A_e and molecular extent of the microhaps (Supplementary Figure S7).

Fig. 5 shows the average A_e values (linear scale) and the random match probability (RMP) values (log scale) for the 27 populations. The 76 MHs panel has high A_e in all populations with lower values for East Asia and Peru (PEL). The random match probability is smallest for the admixed populations, two African and three American. Even the largest values, around 10^{-57}, define good individualization irrespective of population. Principal Components Analysis of the 27 populations shows that ANC clusters tightly with the other European populations (Fig. 6).

The group differences showed in Fig. 5, though mostly small, are supported by STRUCTURE analyses which define five clearly differentiated clusters (Supplemental Figure S8). STRUCTURE analysis was carried out on the 76 MH panel for 27 populations (2606 individuals). The highest likelihood run results for $K = 4$ through 7 are shown in the supplemental Figure S7. The standard admixture model with 10 K burn-ins and 10 K iterations was applied; $K = 2$–9 was tested with 20 runs per K-level. The 5-cluster ($K = 5$) runs produced the best result with ~73% to 99% of the individuals in the 21 least admixed populations studied being assigned to four geographical clusters (Sub-Saharan Africa, Europe, So.Cen. Asia, and East Asia) with estimated cluster membership values > 0.9. The 5th cluster (native Americans from Lima Peru) assigned 67 % of the individuals with estimated cluster membership values $> .9$. The five highly admixed populations (ASW, ACB, MXL, CLM, PUR) usually had very low percentages of the individuals assigned to a single ancestral cluster with est. cluster membership values > 0.9.

4. Conclusions

This study explores the application of microhaplotypes in the context of forensic genetic analysis, with a particular emphasis on the utilisation of degraded DNA specimens.

In this study, the sequencing performance and forensic parameters of a new panel of 76 microhaps were investigated using MPS technology on our Italian samples. This MH panel is highly polymorphic and suitable for forensic applications. Based on the 27 populations studied representing five geographical regions of the world, the overall average Ae for the 76 microhaps is 3.48 and ranges from 2.46 to 4.90. 72 % of the individual population Ae values computable (1483 out of 2052 =27 populations x 76 MH) have values ≥ 3.0 with a large fraction of the total (21.5 %) also being > 4.0. The proposed panel effectively distinguished

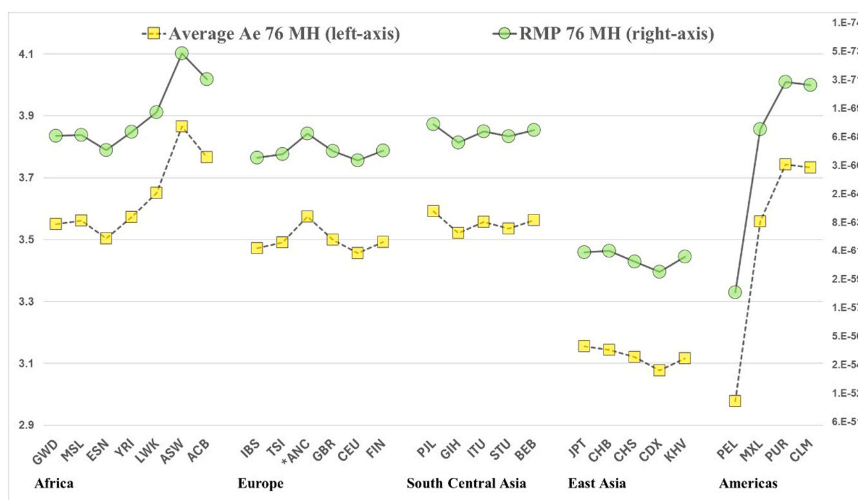


Fig. 5. The average Ae values (linear scale) and the RMP (log scale) in each of the 27 populations.

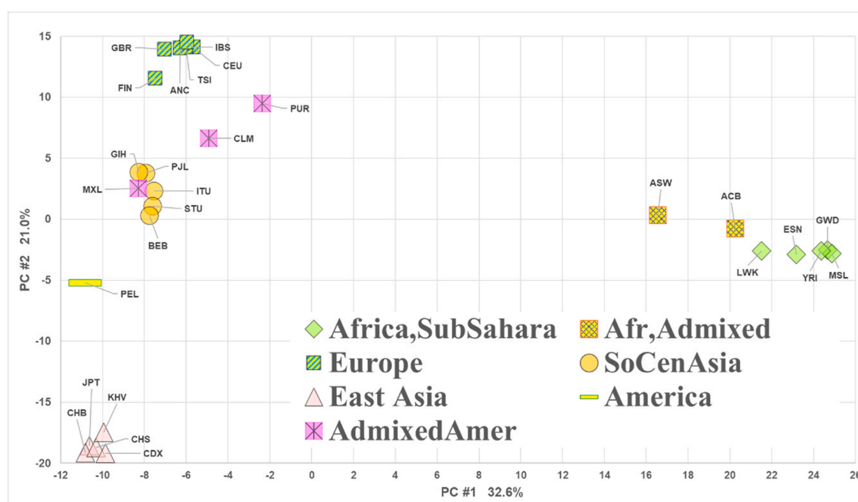


Fig. 6. PCA results for the 27 populations based on the population frequencies for the final 76 MH.

first- and second-degree relationships in simulations, demonstrating its utility for kinship analysis.

The sensitivity test showed that full and reliable profiles could be obtained with as little as 0.05 ng of input DNA. Most casework samples analyzed showed complete or nearly complete MH profiles, even when degraded. Partial MH profiles and allelic drop-out were observed only in samples with DNA quantity below the LOQ of the assays. These results confirmed that the DNA input amount, rather than the DNA degradation level, is the critical parameter affecting correct genotyping. In addition, the informativeness of the partial profiles was assessed using a probabilistic genotyping approach and an LR value exceeding the threshold of 10^6 was obtained, even for the profile in which only 26 MH were typed out of a total of 76 MH.

Overall, the study demonstrated the microhap panel's ability to process low-level and/or degraded DNA that is typically found in forensic cases.

CRediT authorship contribution statement

Filomena Melchionda: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Chiara Turchi:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Kenneth K. Kidd:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Andrew J. Pakstis:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Mauro Pesaresi:** Writing – review & editing, Resources. **Domenico Colloca:** Writing – review & editing, Resources, Investigation. **Alberto Marino:** Writing – review & editing, Resources, Investigation. **Fabiano Gentile:** Writing – review & editing, Resources, Investigation.

Financial support

This research received financial support from "Fondo per la Ricerca Scientifica di Ateneo", Polytechnic University of Marche.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the

online version at doi:10.1016/j.fsigen.2025.103391.

References

[1] C. Børsting, H.S. Mogensen, N. Morling, Forensic genetic SNP typing of low-template DNA and highly degraded DNA from crime case samples, Forensic Sci. Int. Genet. 7 (2013) 345–352, https://doi.org/10.1016/j.fsigen.2013.02.004.

[2] K.B. Gettings, K.M. Kiesler, P.M. Vallone, Performance of a next generation sequencing SNP assay on degraded DNA, Forensic Sci. Int. Genet. 19 (2015) 1–9, https://doi.org/10.1016/j.fsigen.2015.04.010.

[3] J. Liu, W. Du, L. Jiang, C. Liu, L. Chen, Y. Zheng, Y. Hou, C. Liu, Z. Wang, Development and validation of a forensic multiplex Indel assay: the AGCU InDel 60 kit, Electrophoresis 43 (18-19) (2022) 1871–1881, https://doi.org/10.1002/elps.202100376.

[4] F. Oldoni, D. Podini, Forensic molecular biomarkers for mixture analysis, Forensic Sci. Int. Genet. 41 (2019) 107–119, https://doi.org/10.1016/j.fsigen.2019.04.003.

[5] K.K. Kidd, A.J. Pakstis, W.C. Speed, R. Lagace, J. Chang, S. Wootton, et al., Microhaplotype loci are a powerful new type of forensic marker, Forensic Sci. Int. Genet. Suppl. Ser. 4 (1) (2013) e123–e124, https://doi.org/10.1016/j.fsigss.2013.10.063.

[6] K.K. Kidd, W.C. Speed, Criteria for selecting microhaplotypes: mixture detection and deconvolution, Invest. Genet. 6 (1) (2015), https://doi.org/10.1186/s13323-014-0018-3.

[7] 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. 2 (2017) 29–37, https://doi.org/10.1016/j.fsigen.2017.03.014.

[8] 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, https://doi.org/10.1016/j.fsigen.2019.04.009.

[9] M. de la Puente, C. Phillips, C. Xavier, et al., Building a custom large-scale panel of novel microhaplotypes for forensic identification using MiSeq and Ion S5 massively parallel sequencing systems, Forensic Sci. Int. Genet. 45 (2020) 102213, https://doi.org/10.1016/j.fsigen.2019.102213.

[10] J.B. Pang, M. Rao, Q.F. Chen, A.Q. Ji, C. Zhang, K.L. Kang, H. Wu, J. Ye, S.J. Nie, L. Wang, A 124-plex Microhaplotype Panel Based on Next-generation Sequencing Developed for Forensic Applications, Sci. Rep. 10 (2020), https://doi.org/10.1038/s41598-020-58980-x.

[11] N. Gandotra, W.C. Speed, W. Qin, Y. Tang, A.J. Pakstis, K.K. Kidd, C. Scharfe, Validation of novel forensic DNA markers using multiplex microhaplotype sequencing, Forensic Sci.Int. Genet. 47 (2020) 102275, https://doi.org/10.1016/j.fsigen.2020.102275.

[12] Y.S. Feng, C. Zhang, Q.F. Chen, Y. Wang, K.L. Kang, J. Zhao, A.Q. Ji, J. Ye, L. Wang, Evaluation of the MHSeqTyper47 kit for forensically challenging DNA samples, Forensic Sci. Int. Genet. 61 (2022) 102763, https://doi.org/10.1016/j.fsigen.2022.102763.

[13] R. Zhang, J. Xue, M. Tan, D. Chen, Y. Xiao, G. Liu, Y. Zheng, Q. Wu, M. Liao, Meili Lv, S. Qu, W. Liang, An MPS-based 50plex microhaplotype assay for forensic DNA analysis, –865, Genes 14 (2023) 865, https://doi.org/10.3390/genes14040865.

[14] C. Gu, W. Huo, X. Huang, L. Chen, S. Tian, Q. Ran, Z. Ren, Q. Wang, M. Yang, J. Ji, Y. Liu, M. Zhong, K. Wang, D. Song, J. Huang, H. Zhang, X. Jin, Developmental and validation of a novel small and high-efficient panel of microhaplotypes for forensic genetics by the next generation sequencing, BMC Genom. 25 (2024), https://doi.org/10.1186/s12864-024-10880-4.

[15] C. Turchi, F. Melchionda, M. Pesaresi, E. Ciarimboli, C. Bini, P. Fattorini, A. Tagliabracci, Exploring the usefulness of microhaplotypes in forensic identification using massive parallel sequencing technology, Minerva Med. 140 (2020), https://doi.org/10.23736/s0026-4849.20.01790-3.

[16] F. Oldoni, C. Della Rocca, D. Podini, Investigation of 74 microhaplotypes for kinship testing in US populations, Forensic Sci. Int. Genet. Suppl. Ser. 8 (2022) 40–41, https://doi.org/10.1016/j.fsigss.2022.09.015.

[17] Q. Du, G. Ma, C. Lu, Q. Wang, L. Fu, B. Cong, S. Li, Development and evaluation of a novel panel containing 188 microhaplotypes for 2nd-degree kinship testing in the Hebei Han population, Forensic Sci. Int. Genet. 65 (2023) 102855, https://doi.org/10.1016/j.fsigen.2023.102855.

[18] C. Tomas, P. Rodrigues, C.G. Jønck, Zohal Barekzay, Halimureti Simayijiang, V. Pereira, Claus Børsting, Performance of a 74-microhaplotype assay in kinship analyses, Genes 15 (2024) 224, https://doi.org/10.3390/genes15020224.

[19] S. Gao, Q. Wang, Y. Gao, X. Feng, K. Pang, H. Li, F. Zheng, J. Lu, B. Li, J. Liu, M. Yang, K. Li, H. Ismayiljan, H. Yang, J. Yan, X. Guo, Y. Yin, Development and validation of a multiplex panel with 232 microhaplotypes and software for forensic kinship analysis, Forensic Sci. Int. Genet. 76 (2024) 103212, https://doi.org/10.1016/j.fsigen.2024.103212.

[20] C. Turchi, F. Melchionda, M. Pesaresi, P. Fattorini, A. Tagliabracci, Performance of a massive parallel sequencing microhaplotypes assay on degraded DNA, Forensic Sci. Int. Genet. Suppl. Ser. 7 (1) (2019) 782–783, https://doi.org/10.1016/j.fsigss.2019.10.176.

[21] Q. Zhu, P. Shu, H. Wang, Y. Wei, Y. Wang, Y. Wang, J. Zhang, Feasibility study of a single-primer extension-based microhaplotype NGS system, Electrophoresis (2024), https://doi.org/10.1002/elps.202400012.

[22] J. Liu, Y. Shu-qi, Y. Wen, L. Jiang, S. Li, M. Zhao, X. Chen, Z. Wang, Massively parallel sequencing of 74 microhaplotypes and forensic characteristics in three Chinese Sino-Tibetan populations, Forensic Sci. Int. Genet. 66 (2023) 102905, https://doi.org/10.1016/j.fsigen.2023.102905.

[23] R. Tao, Q. Yang, R. Xia, X. Zhang, A. Chen, C. Li, S. Zhang, A sequence-based 163plex microhaplotype assay for forensic DNA analysis, Front. Genet. 13 (2022), https://doi.org/10.3389/fgene.2022.988223.

[24] Z. Bai, N. Zhang, J. Liu, H. Ding, Y. Zhang, T. Wang, J. Gao, X. Ou, Identification of missing persons through kinship analysis by microhaplotype sequencing of single-source DNA and two-person DNA mixtures, Forensic Sci. Int. Genet. 58 (2022) 102689, https://doi.org/10.1016/j.fsigen.2022.102689.

[25] F. Oldoni, D. Bader, C. Fantinato, S.C. Wootton, R. Lagacé, K.K. Kidd, D. Podini, A sequence-based 74plex microhaplotype assay for analysis of forensic DNA mixtures, Forensic Sci. Int. Genet. 49 (2020) 102367, https://doi.org/10.1016/j.fsigen.2020.102367.

[26] K.J. van der Gaag, R.H. de Leeuw, J.F.J. Laros, J.T. den Dunnen, P. de Knijff, Short hypervariable microhaplotypes: a novel set of very short high discriminating power loci without stutter artefacts, Forensic Sci. Int. Genet. 35 (2018) 169–175, https://doi.org/10.1016/j.fsigen.2018.05.008.

[27] P. Chen, W. Zhu, F. Tong, Y. Pu, Y. Yu, S. Huang, Z. Li, L. Zhang, W. Liang, F. Chen, Identifying novel microhaplotypes for ancestry inference, Int. J. Leg. Med. 133 (2018) 983–988, https://doi.org/10.1007/s00414-018-1881-x.

[28] P. Chen, C. Deng, Z. Li, Y. Pu, J. Yang, Y. Yu, K. Li, D. Li, W. Liang, L. Zhang, F. Chen, A microhaplotypes panel for massively parallel sequencing analysis of DNA mixtures, Forensic Sci. Int. Genet. 40 (2019) 140–149, https://doi.org/10.1016/j.fsigen.2019.02.018.

[29] D. Wen, H. Xing, Y. Liu, J. Li, W. Qu, W. He, C. Wang, R. Xu, Y. Liu, H. Jia, L. Zha, The application of short and highly polymorphic microhaplotype loci in paternity testing and sibling testing of temperature-dependent degraded samples, Front. Genet. 13 (2022), https://doi.org/10.3389/fgene.2022.983811.

[30] Y. Tan, H. Tian, Y. Xiao, B. Xu, H. Wang, M. Yang, S. Liu, Screening a new set of microhaplotypes in exonic regions for sample identity testing and paternity testing during whole exome sequencing analysis, Int. J. Leg. Med. 139 (2024) 77–85, https://doi.org/10.1007/s00414-024-03326-9.

[31] M. de la Puente, J. Ruiz-Ramírez, A. Ambroa-Conde, C. Xavier, J. Amigo, M.Á. Casares de Cal, A. Gómez-Tato, Á. Carracedo, W. Parson, C. Phillips, M. V. Lareu, Broadening the applicability of a custom multi-platform panel of microhaplotypes: bio-geographical ancestry inference and expanded reference data, Front Genet. 11 (2020) 581041, https://doi.org/10.3389/fgene.2020.581041.

[32] QIAamp® DNA Micro Handbook, Third Edition, December 2014, Qiagen.

[33] DNA IQ™ Casework Pro Kit for Maxwell® 16, Promega, Technical manual TM332, Revised 12/16.

[34] Bone DNA extraction kit, Promega, Technical manual TM691, Revised 8/22.

[35] Quantifiler™ HP and Trio DNA Quantification Kits user guide (man.4485354, rev. H), Thermo Fisher Scientific.

[36] D.S. Standage, R.N. Mitchell, MicroHapDB: a portable and extensible database of all published microhaplotype marker and frequency data, Front. Genet. 11 (2020) 781, https://doi.org/10.3389/fgene.2020.00781.

[37] K.K. Kidd, Proposed nomenclature for microhaplotypes, Hum. Genom. 10 (2016) 16, https://doi.org/10.1186/s40246-016-0078-ydd.

[38] H. Thorvaldsdóttir, J.T. Robinson, Jill P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration, Brief. Bioinforma. 14 (2013) 178–192, https://doi.org/10.1093/bib/bbs017.

[39] K. Liu, S.V. Muse, PowerMarker: an integrated analysis environment for genetic marker analysis, Bioinformatics 21 (9) (2005) 2128–2129, https://doi.org/10.1093/bioinformatics/bti282.

[40] D. Kling, A.O. Tillmar, T. Egeland, Familias 3 – extensions and new functionality, Forensic Sci. Int. Genet. 13 (2014) 121–127, https://doi.org/10.1016/j.fsigen.2014.07.004.

[41] M. Stephens, N. Smith, P.A. Donnelly, New statistical method for haplotype reconstruction from population data, Am. J. Hum. Genet. 68 (2001) 978–989, https://doi.org/10.1086/319501.

[42] M. Stephens, P. Scheet, Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation, Am. J. Hum. Genet. 76 (2005) 449–462, https://doi.org/10.1086/428594.

[43] N.A. Rosenberg, L.M. Lei, R. Ward, J.K. Pritchard, Informativeness of genetic markers for inference of ancestry, Am. J. Hum. Genet. 73 (2003) 1402–1422, https://doi.org/10.1086/380416.

[44] M.M. Foley, Likelihood ratio calculation Using LRmix studio, Methods Mol. Biol. (Clifton N. J.) 2685 (2023) 307–328, https://doi.org/10.1007/978-1-0716-3295-6_19.

[45] Ø. Hammer, D.A.T. Harper, P.D. Ryan, PAST: paleontological statistics software package for education and data analysis, Palaeontol. Electron. 4 (1) (2001) 9. ⟨http://palaeo-electronica.org/2001_1/past/issue1_01.htm⟩.

[46] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data, Genetics 155 (2) (2000) 945–959, https://doi.org/10.1093/genetics/155.2.945.

[47] P. Gill, L. Gusmão, H. Haned, W.R. Mayr, N. Morling, W. Parson, L. Prieto, M. Prinz, H. Schneider, P.M. Schneider, B.S. Weir, DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods, Forensic Sci. Int. Genet. 6 (6) (2012) 679–688, https://doi.org/10.1016/j.fsigen.2012.06.002.

[48] ENFSI Guideline for evaluative reporting in forensic science. Strengthening the Evaluation of Forensic Results across Europe (STEOFRAE) – Version 3, March 2015.