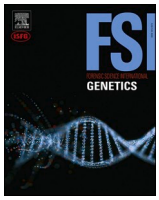
Forensic Science International: Genetics 67 (2023) 102937

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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



Adapting an established Ampliseq microhaplotype panel to nanopore sequencing through direct PCR

L. Casanova-Adán, A. Mosquera-Miguel, J. González-Bao, A. Ambroa-Conde, J. Ruiz-Ramírez, A. Cabrejas-Olalla, E. González-Martín, A. Freire-Aradas, A. Rodríguez-López, C. Phillips, MV Lareu, M. de la Puente *

Forensic Genetics Unit, Institute of Forensic Sciences, Universidade de Santiago de Compostela, Spain

ARTICLE INFO

Keywords:
Nanopore sequencing
MinION
Microhaplotypes
Kinship testing
Direct PCR
Field laboratories

ABSTRACT

We have adapted an established Ampliseq microhaplotype panel for nanopore sequencing with the Oxford Nanopore Technologies (ONT) system, as a cost-effective and highly scalable solution for forensic genetics applications. For this purpose, we designed a protocol combining direct PCR amplification from unextracted DNA with ONT library construction and sequencing using the MinION device and workflow. The analysis of reference samples at input amounts of 5–10 ng of DNA demonstrates stable coverage patterns, allele balance, and strand bias, reaching profile completeness and concordance rates of ~95%. Similar levels were achieved when using direct-PCR from blood, buccal and semen swabs. Dilution series results indicate sensitivity is maintained down to 250 pg of input DNA, and informative profiles are produced down to 62.5 pg. Finally, we demonstrated the forensic utility of the nanopore workflow by analyzing two third degree pedigrees that showed low likelihood ratio values after the analysis of an extended panel of 38 STRs, achieving likelihood ratios 2–3 orders of magnitude higher when testing with the MinION–based haplotype data.

1. Introduction

The forensic community is currently exploring alternative sequencing methodologies beyond classic Sanger sequencing methods, with a specific focus on Massively Parallel Sequencing (MPS). MPS can directly determine the phase of the SNPs within an amplified sequence tract, which has prompted the investigation of Microhaplotypes (MHs) in forensics [1–4]. MHs represent powerful multipurpose forensic markers that combine the advantages of STRs as highly polymorphic markers, and SNPs as short amplicon polymorphisms. Applications of MHs include identification from degraded DNA [5–8], kinship testing [6,9,10], biogeographic ancestry inference [11–16] and mixture analysis [17–21].

Despite offering improved sensitivity and greater sequence detail, MPS technologies have not become fully established in routine forensic analysis. Most casework DNA analyses can be solved by typing STRs through capillary electrophoresis (CE) and these markers form the core of national criminal DNA databases. When a case requires extra DNA information that can be obtained through MPS, the lack of small sample number scalability of the two main platforms used in forensics (Ion S5 and MiSeq) raises the cost of the analysis, and this particularly affects low- and medium-throughput laboratories. Moreover, most routine laboratories do not have access to MPS platforms as they require considerable initial investment.

For these reasons, we set out to explore the capabilities for forensic genetic applications of the new methodology of nanopore sequencing from Oxford Nanopore Technologies (ONT). Nanopore sequencing represents a move from second generation to third generation sequencing, by exploiting the unique properties of nanopores. Two recent review publications [22,23] have detailed the possible benefits of nanopore sequencing: the technology shares the same advantage of high multiplexing capability as MPS and, in contrast to MPS, adds high scalability and low initial investment in equipment. In common with CE, ease of use is a key characteristic of nanopore sequencing, making it readily implementable in forensic laboratories regardless of the size and scope of their work. In contrast to MPS systems, nanopore technology is implemented in a small, highly portable pocket-size device: the MinION, which prompts its use in field laboratories. Moreover, nanopore-dedicated resources as well as access to a large community of active and experienced users are available to help troubleshooting or

* Corresponding author.
E-mail address: mdelcarmendela.puente@usc.es (M. de la Puente).

https://doi.org/10.1016/j.fsigen.2023.102937
Received 19 July 2023; Received in revised form 21 September 2023; Accepted 27 September 2023
Available online 28 September 2023

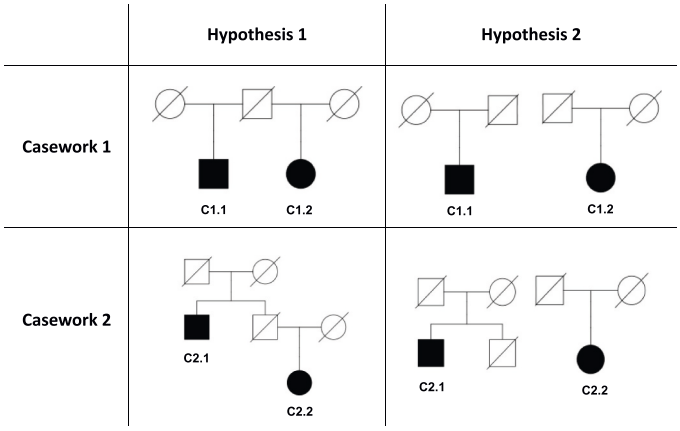


Fig. 1. Tested hypotheses for the two kinship caseworks considered in the study. Casework 1: half siblings vs unrelated. Casework 2: pairwise avuncular relationship vs unrelated.

self-training.

High sequence error rates and high DNA input requirements are generally identified as the major drawbacks of the use of nanopore sequencing in forensic analysis. However, continuous developments during the last decade of both nanopore structure and the bioinformatic interpretation tools have progressively reduced the sequence read error rate. Previous forensic studies using nanopore sequencing [24–28] concluded that genotyping accuracy is highly marker dependent, and especially low if poly-tracts are adjacent to the SNP position.

The protocol designed for this study combines a capture PCR based in a direct PCR kit, i.e., allowing amplification from unextracted DNA for the most common biological tissues (blood, saliva, semen), with ONT ligation library construction and sequencing. Implementing this direct PCR protocol reduces the DNA lost during the purification steps of the DNA extraction, reduces the cost per sample and the processing time. Thus, the main advantage of implementing direct PCR is that the final workflow equipment requirements are reduced to a thermal cycler, a laptop, and the MinION sequencer, indicating nanopore technology's applicability for forensic operations in the field.

2. Material and methods

2.1. DNA samples and controls

Sample swabs were collected from volunteer donors under informed consent, with a protocol approved by the Bioethical Committee of the Universidade de Santiago de Compostela (USC-30/2021 and USC-58/2022). When required, DNA extractions were performed using the PrepFiler Express Forensic DNA Extraction Kit (Applied Biosystems, AB) on the AutoMate Express Nucleic Acid Extraction System (AB), following the recommended protocol. DNA quantification was performed using the Quantifiler Trio DNA Quantification Kit (AB) on a 7500 Real-Time PCR System (AB).

A total of 15 reference DNA samples including six DNAs extracted

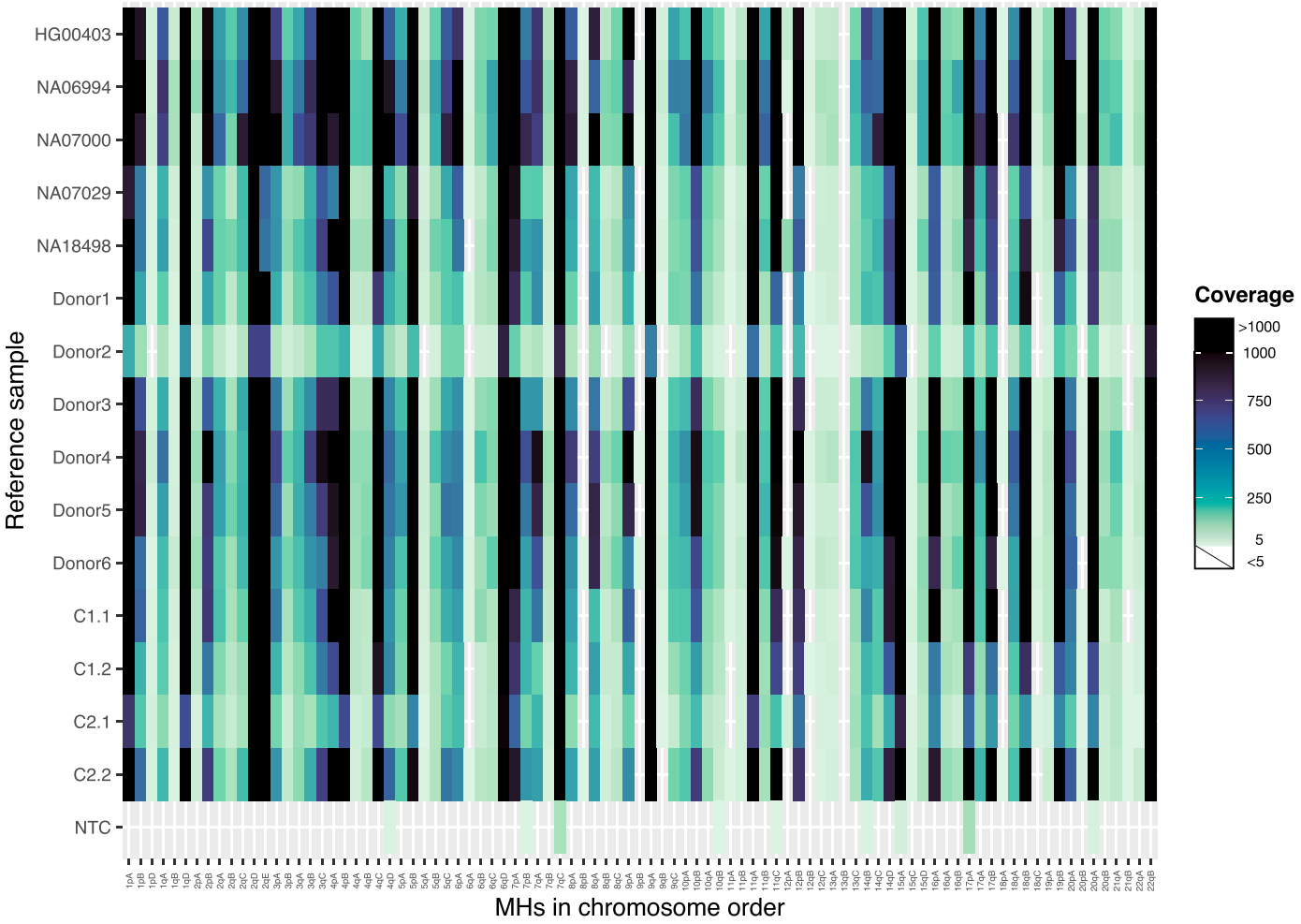


Fig. 2. Raster plot representations of sequence coverage (read depth) of the 91 autosomal MHs in chromosome order, for reference samples and the non-template control (NTC).

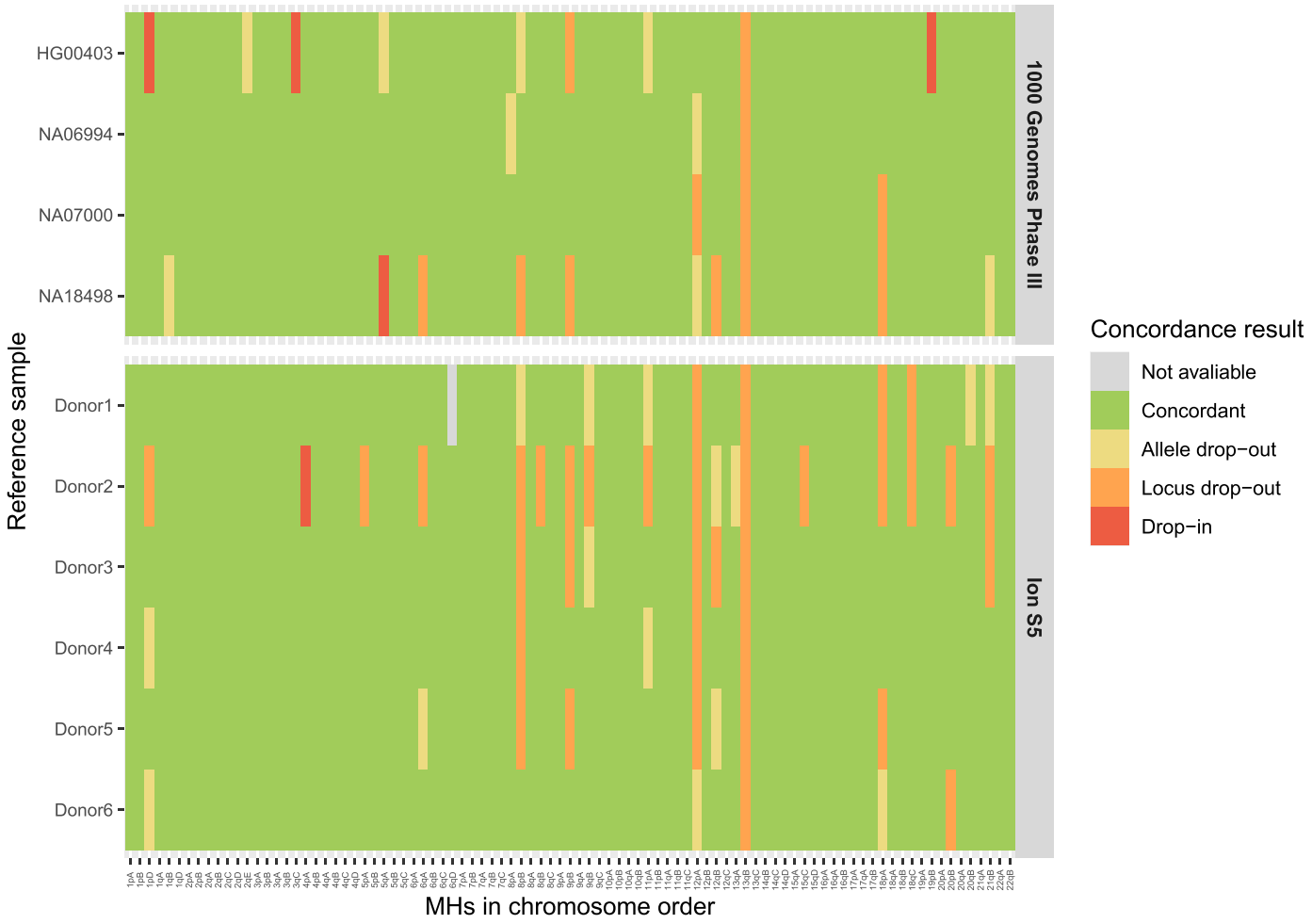


Fig. 3. Raster plot representations of genotype concordance results displayed as: concordant (green), allele drop-out (yellow), locus drop-out (orange), drop-in (red) and not available (grey, no reference genotype available for comparison), for the 91 autosomal MHs in chromosome order for the Coriell cell-line DNAs, compared to the 1000 Genomes Phase III data, and the six donor DNAs with previous genotypes obtained with the Ion S5 MPS system.

from donor buccal swabs, five Coriell cell-line DNA controls (HG00403, NA06994, NA07000, NA07029 and NA18498) and four DNAs as part of two kinship testing scenarios, half-siblings – C1.1 and C1.2 – and samples from an avuncular related pair – C2.1 and C2.2 – were analyzed at input amounts ranging from 5 to 10 ng. A single non-template control was included to assess levels of non-specific amplification.

A dilution series of DNA control 007 was prepared for input quantities of 5 ng, 1 ng, 0.5 ng, 250 pg, 125 pg and 62.5 pg, to evaluate sensitivity to low level input DNA.

Blood, semen, and buccal swabs were collected in duplicate from a single donor and analyzed to evaluate the direct PCR approach.

2.2. Optimization of a direct PCR library construction from Ampliseq primers for nanopore sequencing

PCR amplification was performed using the Platinum Direct PCR Universal Master Mix Kit (Invitrogen), following manufacturer's recommendations. This kit was selected as the most efficient for substituting the PCR reagents of the Ion AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific, TFS) amongst six different PCR kits assessed in prior testing. PCR reactions comprised 5 μL of the MHs panel Ampliseq primers as described in [6], 5 μL of Master Mix and up to 3 μL of DNA sample to a final volume of 13 μL. PCR cycling with GeneAmp PCR System 9700 or 2700 thermocyclers (AB) used steps: 10 min at 95 °C, 35 cycles of 30 s at 95 °C, 40 s at 62 °C and 1 min at 72 °C with a final extension of 20 min at 72 °C.

For amplification from donor body fluid swabs, cell lysis prior to direct PCR reactions was conducted by adding 20 μL of a freshly prepared mix comprising 0.6 μL of Platinum Proteinase K (Invitrogen) and 20 μL of Platinum Lysis solution (Invitrogen) to a fragment of the swab and incubating at 98° C for 1 min in a thermoblock, as detailed in the manufacturer's recommendations.

PCR product purification was carried out by adding 1 μL of ExoSAP-IT PCR Product Clean-Up Reagent (AB), and incubating at 37 °C for 15 min, then denaturation at 80 °C for 15 min in a GeneAmp PCR System 9700 or 2700 thermocycler (AB).

DNA yield was checked prior to library construction using the Qubit dsDNA HS Assay Kit and the Invitrogen Qubit 3 Fluorometer. The remaining 13 μL of purified PCR product were used for library construction using the SQK-LSK109 or SQK-LSK110 Ligation Sequencing Kit (ONT) and NEBNext Companion Module for Oxford Nanopore Technologies Ligation Sequencing Kit (New England Biolabs).

Library construction was performed following the Amplicons-by-ligation protocol for Flongle flow cells with the following modifications: (i) all reactions were adjusted to half volumes; (ii) for all the AMPure XP Reagent (Beckman Coulter) clean-ups, the ratio of magnetic beads was adjusted up to 3X to ensure the recovery of the fragments of interest (150–200 bp), incubation time was increased up to 15 min in a plate mixer and the elution step included an incubation at 37°C and for 15 min. Libraries were sequenced individually in FLO-FLG001 Flongle flow cells (R9.4.1) with Short Fragment Mode (20 bp).

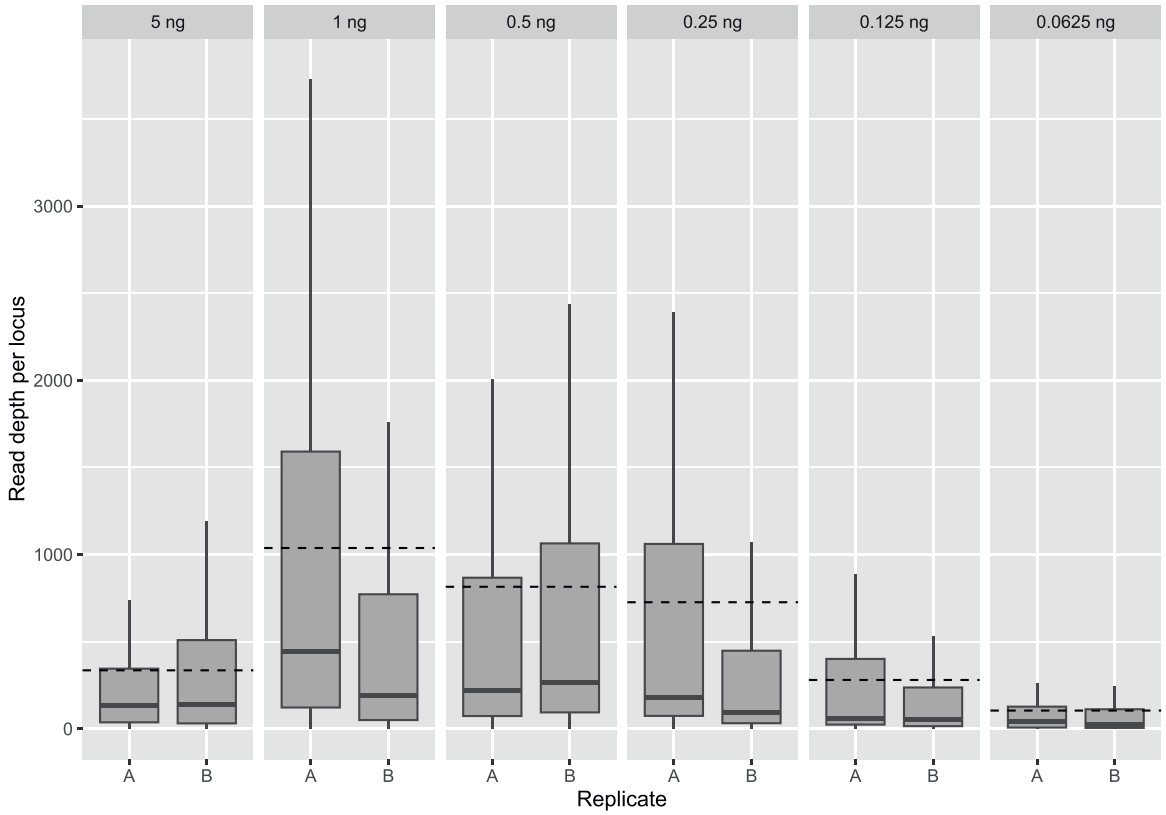


Fig. 4. Boxplots representing sequence read depth per locus (coverage) of replicates A and B of the sensitivity series at inputs of 5 ng, 1 ng, 0.5 ng, 0.25 ng, 0.125 ng and 0.0625 ng. Dashed lines represent mean overall values between the two replicates.

2.3. Data analysis

2.3.1. MH and SNP calling, sequence quality analysis and concordance

Basecalling from fa5 files was performed using MinKnow version 22.12.7 with basecalling algorithm FLO-MIN106 / FLO-FLG001 DNA (High Accuracy), to produce raw fastq files. A previously established MH calling pipeline [6] based on open software [29–32] was applied to generate MH haplotype calls from fastq files, using a minimum coverage per composite SNP allele threshold of 5 reads per allele and a minimum allele frequency of 0.2. All X-chromosome MHs and those autosomal MHs previously identified as underperforming (red and yellow pie chart sectors in Fig. 6 of [6]) were excluded from further analysis, leaving a total of 91 autosomal MHs. The produced microhaplotype calls and read depths were analyzed using R scripts [32].

Sequence quality was evaluated using the 15 reference samples and included an assessment at the MH, or composite SNP level. IGV v2.16.0 [33] was used for bam file visualization, with bam-readcount [34] software used to obtain metrics at the composite SNPs positions, and R scripts to calculate and represent the following parameters: (i) coverage as number of reads per SNP; (ii) allele balance as major allele frequency or reads of the most read allele / total coverage; (iii) read strand balance or strand bias as forward coverage / total coverage; and (iv) misincorporation rates as reads from no-called alleles / total coverage.

Genotype concordance was evaluated with two comparative evaluations: (i) MinION genotypes with those in public data repositories of 1000 Genomes Project Phase III [35] for HG00403, NA06994, NA07000 and NA18498 Coriell cell-line controls analyzed as reference samples; (ii) with genotypes previously obtained in the Ion S5 MPS system using protocols described in [6] for the six donor samples, the body fluid swabs analyzed by direct PCR and the 007 control DNA dilution series.

2.3.2. Kinship testing

The two kinship scenarios shown in Fig. 1, comprised half-siblings (Casework 1) and an avuncular related pair (Casework 2), which were selected for indicating low likelihood ratio (LR) values. Prior kinship analysis genotyped a total of 38 STRs following the recommended protocols, comprising: (i) 21 STRs of the GlobalFiler PCR Amplification Kit (AB); (ii) 9 non-overlapping STRs of the Investigator HDplex (Qiagen), and; (iii) an in-house panel of 8 pentameric STRs [36] (STR ATAAC02 excluded due to its close position to vWA and D12S391). These 38 STRs were applied to kinship testing with relevant population frequencies [36–38] using Familias 3.2.1 software [39].

MH profiles for the four kinship samples were generated using the previously described MinION pipeline, both pedigrees were tested using Familias 3.2.1 software [39] and applying 1000 Genomes European haplotype frequencies [35]. A total of 10,000 simulations of related-as-tested for a third-degree relationship and unrelated individuals were generated for the 38 STRs and the 91 autosomal MHs using Familias 3.2.1 [39], then plotted using R [32].

3. Results and discussion

3.1. Sequencing quality and concordance

3.1.1. Sequencing quality

First, Fig. 2 shows total coverage values per marker in the reference set samples and the non-template control. The Fig. 2 plot shows a stable pattern across samples, with certain MHs consistently exhibiting either high or low sequence read depth values. Mean MH coverage per sample (Supplementary Fig. S1) ranged from 112.87 to 826.07 reads.

Mean total coverage per marker values of the reference set (Supplementary Fig. S2) produced a typical slope profile attributable to the differential amplification in multiplex PCR, with a range from 2.86 to 5933.4 reads, with eleven MHs showing mean values lower than 20 reads: 18pA, 9pB, 8pB, 12qB, 6qA, 21qB, 11pA, 12pA, 9qB and 18qC. These markers do not completely match previously observed results on

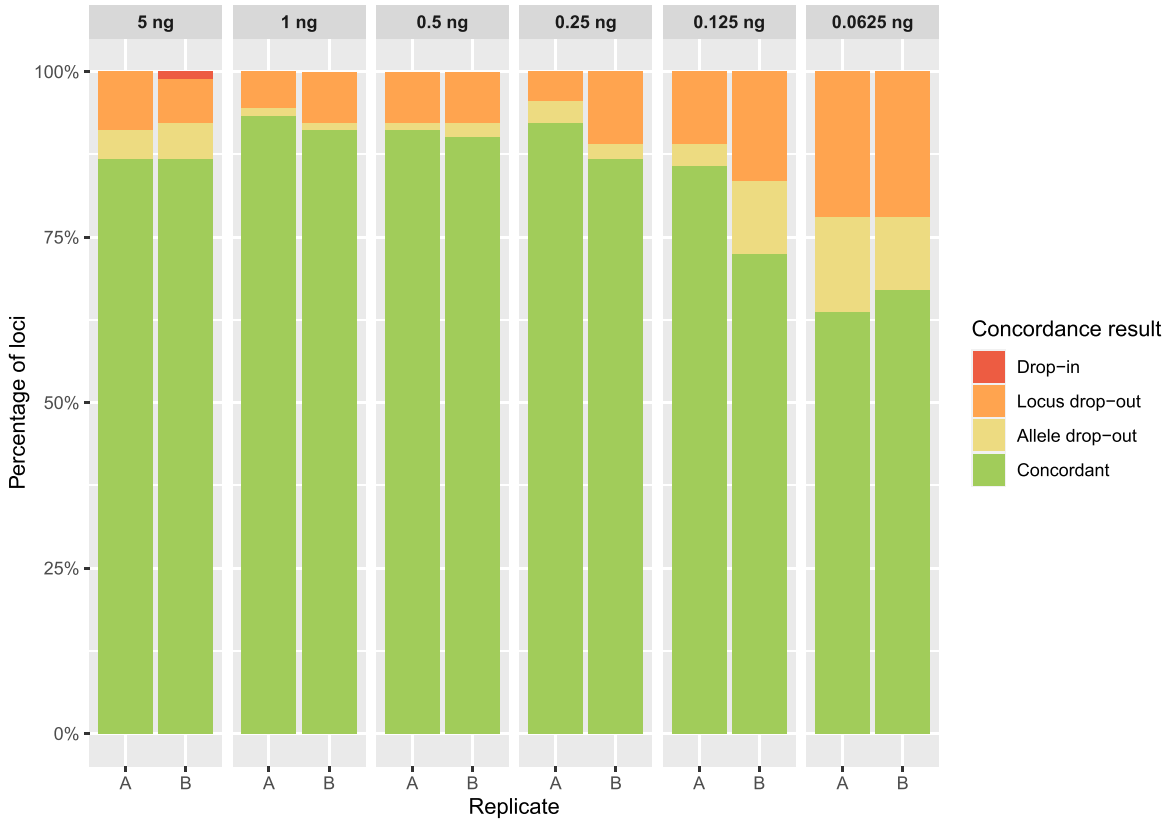


Fig. 5. Bar plot representations of genotype concordance results displayed as: concordant (green), allele drop-out (yellow), locus drop-out (orange) and drop-in (red) for replicates A and B of the sensitivity series at inputs of 5 ng, 1 ng, 0.5 ng, 0.25 ng, 0.125 ng and 0.0625 ng, compared to genotypes obtained with the Ion S5 MPS system.

the S5 and MiSeq platforms, probably due to changes in PCR conditions. Coverage distribution within each locus, i.e., at the SNP level, was also evaluated (Supplementary Fig. S3), indicating sequence reads uniformly covered the amplicon.

As represented in Fig. 2 and Supplementary Fig. S1, the non-template control showed low sequence coverage levels easily distinguishable from the reference samples with some allele calls representative of spurious, stochastic amplification of DNA. The mean total depth across markers reached values of 3.51 reads, with only nine loci showing values above 0 and ranging from 9 to 112 reads.

Second, allele balance was assessed in composite SNPs by calculating the major allele frequency, shown in Supplementary Fig. S4. SNPs consistently (in more than 66% of the cases, 10/15 reference samples) showing major allele frequencies in the range of 60–90%, not expected for single source samples, were identified for MHs 5qA (rs400452), 6qD (rs9365583), 1qD (rs12141154), 5qC (rs256257) and 6qB (rs2799638). Allele balance values show similar patterns to those previously reported when analyzing the same MH panel with Ion S5 and MiSeq MPS systems (see Supplementary Fig. S5 in [6]).

Third, strand bias was evaluated for composite SNPs (Supplementary Fig. S5); with SNPs showing uniform values across the same haplotype. The small intervals observed in the boxplots for most MHs indicates marker-based strand bias is consistent across the reference sample set.

Last, mean misincorporation rates per composite SNP (Supplementary Fig. S6) ranged from 0% to 10.90% and reached a mean value across SNPs of 2.69%, a higher value compared to those previously reported when analyzing the same MH panel using Ion S5 and MiSeq MPS (0.25% and 0.73%, respectively). Mean misincorporation values higher than 5% were observed in 35 of 325 SNPs.

Supplementary Fig. S7 shows the IGV visualization of SNP rs6702428 (MH 1pD), which showed the highest mean misincorporation rate for sample NA06994. In contrast to results from other MPS platforms, IGV visualization of the nanopore reads for this SNP revealed a higher frequency of misincorporations and Indels across the amplified sequence, that also influence the detection of the other composite SNP sites and might be related to the basecalling algorithm.

This higher rate of misincorporations and the comparatively low coverage values obtained for rs6702428 in the reference samples might hinder the analysis of non-reference samples, such as those with low-level DNA and DNA mixtures.

3.1.2. *Genotype concordance*

Genotype concordance for the reference samples was evaluated by comparison with 1000 Genomes Project Phase III genotypes and those obtained with Ion S5 MPS. Genotype comparisons are summarized in Fig. 3.

First, the four Coriell cell-line DNA controls reached an overall genotype call rate of 96.70% (352/364 genotypes) and a locus drop-out rate of 3.30% (12/364 genotypes), as shown in Fig. 3 and Supplementary Fig. S8. Amongst the called genotypes, 96.30% were concordant with the genotype reported by the 1000 Genomes Phase III data (339/352 genotypes) while 1.14% (4/352) showed an allele drop-in and 2.57% (9/352) an allelic drop-out.

Second, the six donor samples reached an overall genotype call rate of 93.59% (511/546 genotypes) and a locus drop-out rate of 6.41% (35/546 genotypes), as shown in Fig. 3 and Supplementary Fig. S8. Amongst the called genotypes, 96.67% were concordant with genotypes obtained with the Ion S5 system (494/511 genotypes) while 0.19% (1/511) showed an allele drop-in and 2.93% (15/511) showed an allele drop-out, respectively. A single genotype (1/511) could not be compared due to the lack of reference data.

These results show that the new MinION workflow cannot provide full profiles for the complete set of 91 autosomal MHs previously characterized with an MPS workflow for both the Ion S5 and MiSeq systems.

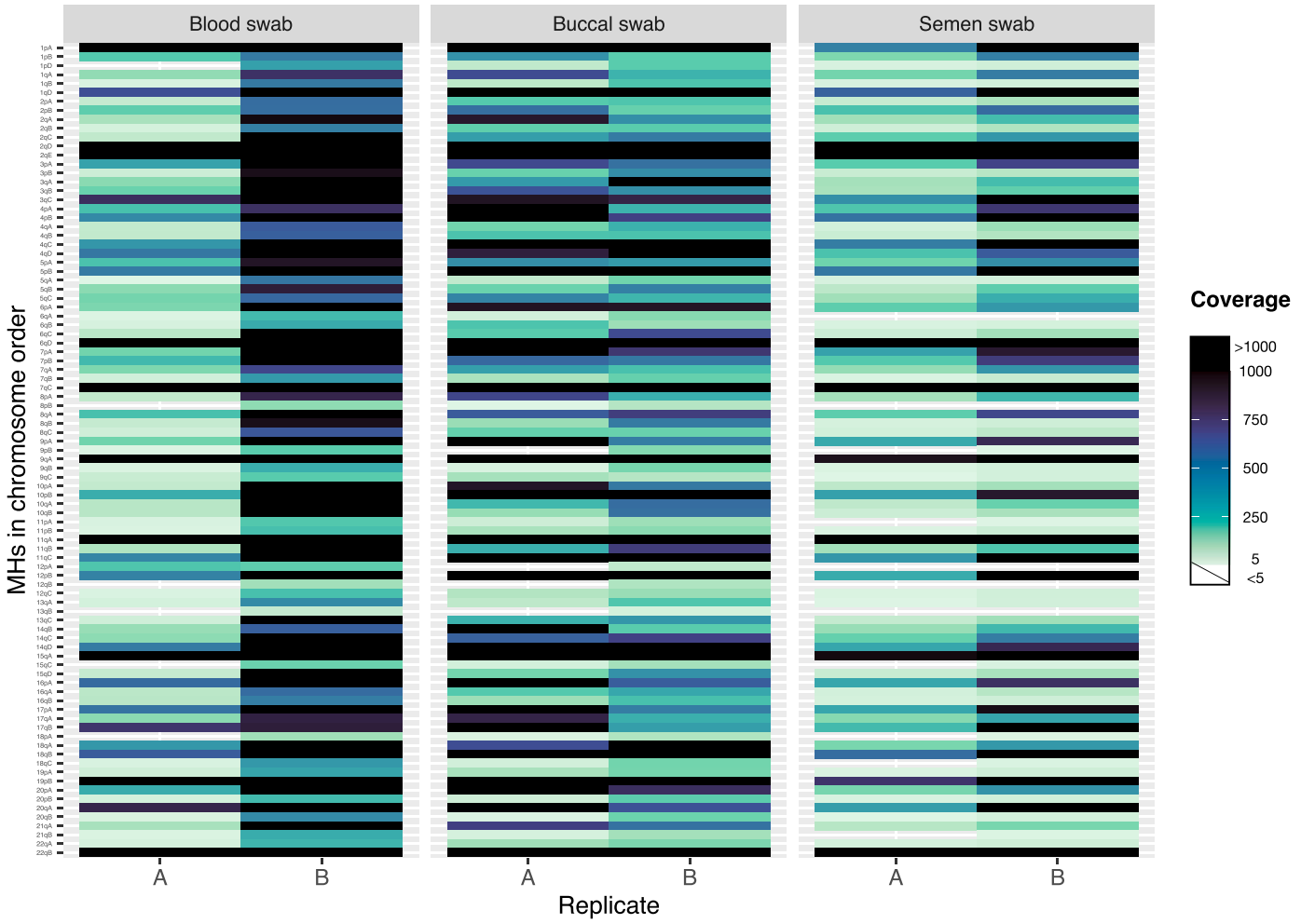


Fig. 6. Raster plot representations of sequence read depth (coverage) of the 91 autosomal MHs in chromosome order for replicates A and B of the direct PCR tests from blood, buccal and semen swabs.

However, a call rate of ~95% ensures high informativeness levels given the degree of polymorphism exhibited by the MHs, superior to that achieved using the widely used GlobalFiler CE-STR kit. The significant drop-in and allele drop-out rates found suggest that further refinement of the calling parameters applied to the MinION sequence output could be required.

3.2. Sensitivity to low level DNA

Fig. 4 displays the sequence read depth per locus recorded in each replicate of the 007 control DNA dilution series. Mean read depth per locus between replicates of the same input amount decreases from 1037.56 reads for 1 ng to 104.50 reads for 62.5 pg, while the 5 ng replicates show unexpected low values of 335.50 reads. These results, together with the lack of uniformity of the reference samples, suggests high inter-run variability for the MinION system that could be affected by the quality and age of the flow cells. The Flongle flow cells can be stored for up to four weeks according to the manufacturer's guidance. Furthermore, the experience of the analyst in loading the flow cell or preparing the libraries can affect nanopore sequencing performance. Nevertheless, as shown in Supplementary Fig. S9, most of the MHs were amplified successfully and were able to reach a sequence read threshold of five reads per allele.

The genotypes obtained from each replicate of the 007 dilution series were compared to those previously obtained with the Ion S5 system and results are shown in Fig. 5 and Supplementary Fig. S10. Mean genotype call rates maintained similar values over 92% for input amounts from 5 ng to 0.25 ng, but these reduced to 86.26% and 78.02% for 0.125 ng and 0.0625 ng of input DNA, respectively. However, mean allele concordant rates decrease to 91.72% (144/157 called genotypes considering the two replicates) and 83.80% (119/142 called genotypes) while mean drop-out rates increased to 8.28% (13/157 called genotypes considering the two replicates) and 16.20% (23/142 genotypes) for input amounts of 0.125 ng and 0.0625 ng, respectively. A single drop-in was observed in one 5 ng replicate.

These results support the idea that partial but informative MH profiles can be obtained from DNA quantities as low as 62.5 pg, reaching a sensitivity adequate for forensic genetic analyses. For the analysis of low-level DNA, a higher interpretation threshold or a drop-out modeling could be implemented to avoid the misinterpretation of single allele calls producing false homozygotes from a failed allele detection.

3.3. Direct PCR

One of the goals of adapting the USC-MHs panel to the portable sequencing solution represented by the MinION system was to allow for a direct PCR protocol from swabs, circumventing a DNA extraction step requiring both extra equipment and processing time. The results obtained in terms of sequence coverage from the two replicates of blood, buccal and semen swabs undergoing direct PCR, are shown in Fig. 6 indicating a mean locus sequence read depth between replicates of 943.16, 919.30 and 505.47 for blood, buccal and semen swabs,

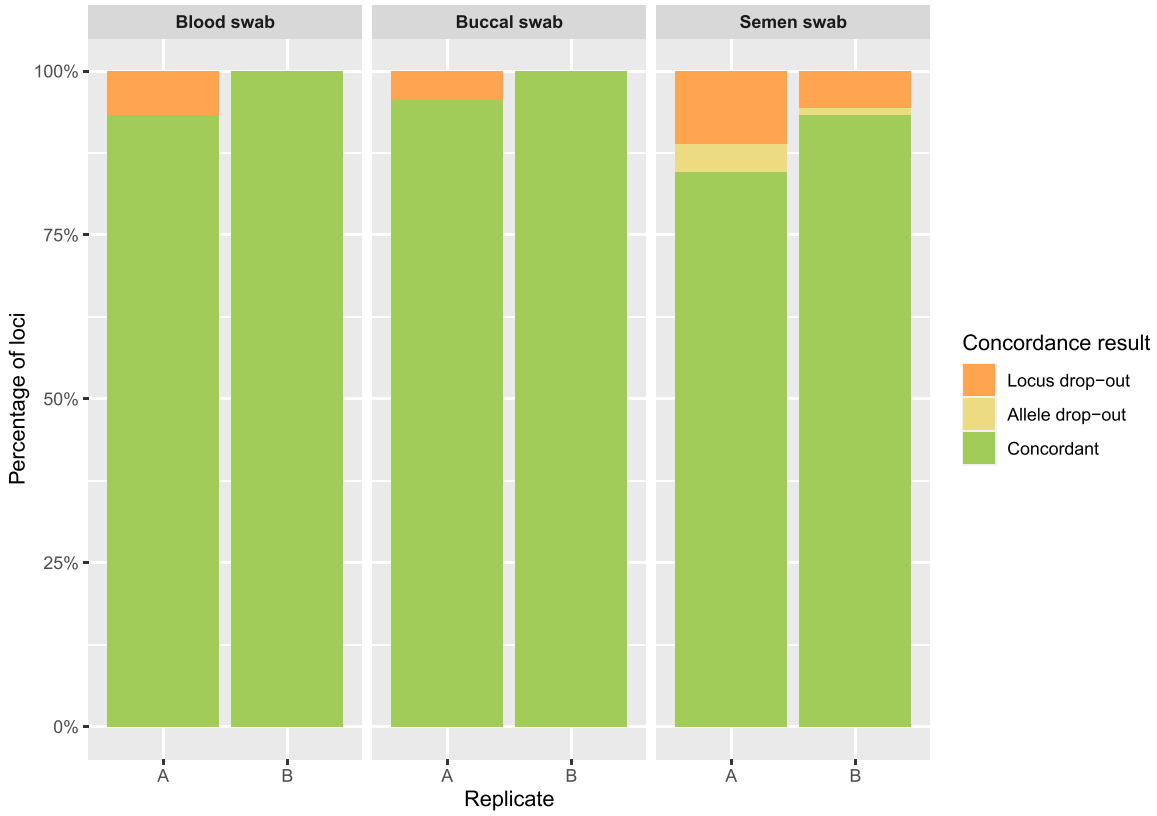


Fig. 7. Bar plot represents concordance results as concordant (green), allele drop-out (yellow), locus drop-out (orange) and drop-in (red) for replicates A and B of the direct PCR tests from blood, buccal and semen swabs, compared to genotypes obtained with the Ion S5 MPS system.

Table 1
Likelihood ratio values for two relationship tests detailed in Section 2.3.2 using the different combination of markers.

Marker set	Case 1	Case 2
GlobalFiler	3.90	6.14
GlobalFiler + HDplex	82.47	571.98
GlobalFiler + HDplex + pentameric STRs	125.30	28.44
91 MHs genotyped by MinION	27,402.60	82,454.31

respectively.

Concordance of the direct PCR replicates (Fig. 7 and Supplementary Fig. S11) was evaluated by comparison with MH genotypes obtained with the Ion S5 system. Mean genotype call rates reached values of 96.70%, 97.80% and 91.75% for blood, buccal and semen swabs, respectively. Mean allele concordant rates reached 100% for blood and buccal swabs (176/176 and 178/178 called genotypes) and 97% for semen swabs (162/167 called genotypes), the latter reduced rate due to three allele drop-outs (3% allele drop-out, 5/167 called genotypes).

These results are consistent with the findings for the reference DNAs, supporting the idea of a robust direct PCR protocol when working with buccal swabs from the donor or swabs sampling a surface with fresh blood or semen. However, a further validation is required to extend the assay to other forensically relevant DNA sources such as fabric, cigarette butts or FTA cards.

3.4. Kinship testing

Table 1 shows the likelihood ratio (LR) values obtained from two casework scenarios evaluated with the routinely applied panel of STRs from GloblaFiler, HDplex and the in-house pentameric STR set, compared to those obtained from the MH panel sequenced with the MinION. These show that the MinION provides LR values that are 2–3 orders of magnitude higher.

Fig. 8 plots the distribution of the expected LR values generated from 10,000 simulations of third-degree relatives and unrelated individuals, for STRs (left) and MHs (right). Mean LR values for related vs unrelated show a greater differentiation with the MH set. Moreover, the values obtained for cases 1 and 2 overlap with the distribution of the unrelated LRs when using STRs, but not when using MHs.

These results indicate that the high levels of informativeness expected to be provided by a large panel of MHs could be exploited to provide a greater degree of genetic information when single urgent distant kinship cases arrive at the laboratory and require detailed analyses in a cost-effective way and at a suitable scale of samples sequenced – making appropriate use of the sequencing capacity of the flexible MinION platform compared to MPS.

4. Concluding remarks

In this study, we adapted the USC-MHs panel to the MinION platform for nanopore sequencing. The results suggest good genotyping performance with reference DNA samples, with call rates of ~95% for an optimal subset of 91 autosomal MH loci. Low-level DNA sensitivity levels are well maintained down to 250 pg DNA input, while providing informative profiles down to input amounts of 62.5 pg. The novel protocol, we have developed for the MinION, together with previous development of extra functionality of our MH panel for biogeographical ancestry inference [13] (despite being initially conceived as a panel for individual identification of degraded DNA and kinship testing) broadens the applicability of microhaplotypes. This broad applicability, increased flexibility to analyze small sets of samples in short timeframes, and simple direct PCR protocols potentially useful in the field allows most forensic laboratories to benefit from the low initial implementation costs of the MinION platform compared to the Ion S5 and MiSeq MPS

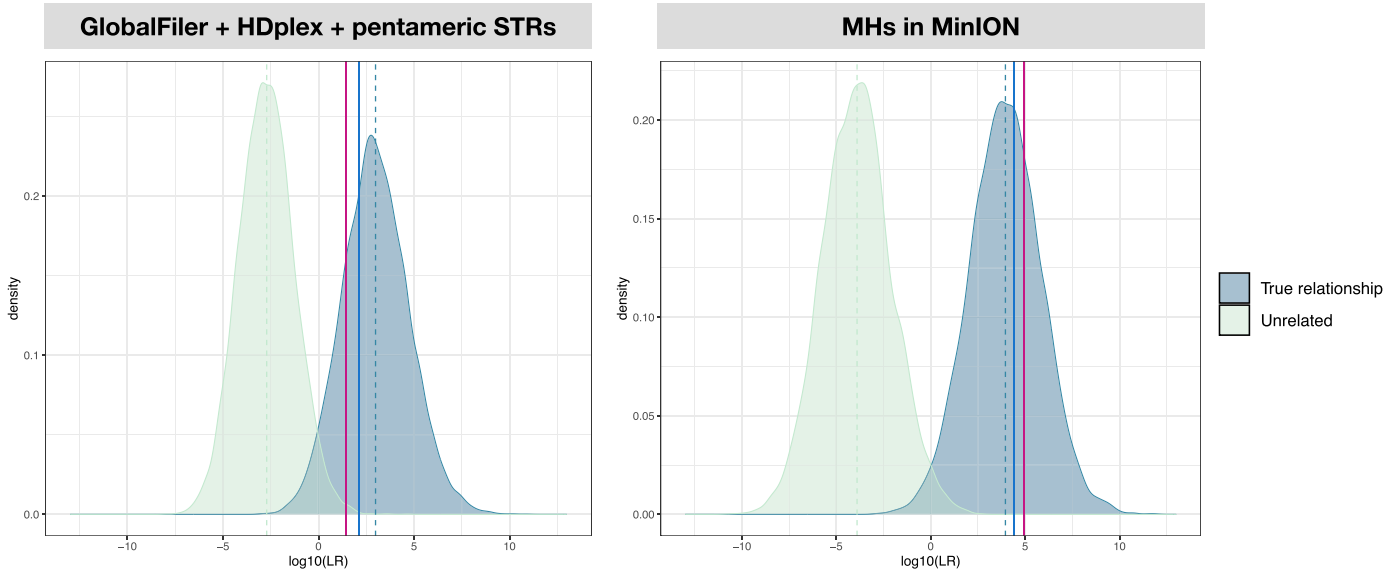


Fig. 8. Plots of $\log_{10}$LR obtained from 10,000 simulations of related-as-tested (third degree relative) in blue and unrelated individuals in green. Dashed lines represent mean values. Data obtained from the genotypes of a set of 38 STRs (left) and the 91 autosomal MHs analyzed in this study (right). LR results obtained from the Casework 1 (blue) and Casework 2 (red) are represented as vertical lines.

platforms.

We demonstrated the ability of the developed MinION workflow to resolve two pairwise third-degree relationship tests that had low LRs from genotyping an extended panel of 38 STRs, providing LR values 2–3 orders of magnitude higher after the analyisis of MHs. These results support the idea of a highly scalable platform that can be adapted to a single urgent casework with reduced costs, in contrast to the need to run partially filled flowcells for the MiSeq instrument and sequencing chips for the Ion S5 instrument. For example, for distant kinship casework such as the examples outlined in Section 3.4, the new workflow with the MinION represents approximately one quarter of the Ion S5 analysis costs.

The direct PCR system we optimized for the MinION combines well with the portability of the MinION, by reducing the required equipment for sequencing in the field to a pocket-sized sequencer, a portable thermocycler, and a laptop. Taking advantage of this portability, suggests the MH panel described, the direct PCR workflow and the instrumentation for nanopore sequencing could be realistically applied to aid field identifications in mass disaster programs when distant kinship comparisons are required, and forensic laboratory infrastructure is lacking close to the site. The described panel and nanopore sequencing system could also complement rapid DNA testing in criminalistic casework by providing additional means to differentiate close relatives and add biogeographical ancestry intelligence. In this context, the workflow can be translated to other custom or commercial AmpliSeq based panels, either for forensic or other purposes.

Further development is required to optimize the workflow for degraded or mixed DNA, as the sequencing accuracy of the MinION platform is lower than that achieved by the Ion S5 and MiSeq systems. Luckily, new advances from ONT in the architecture of the nanopore version R10, including a longer barrel and dual reader head have provided improved resolution of homopolymeric regions [40], and could help achieve levels of accuracy necessary for improved analysis of mixed and degraded DNA in the near future [26].

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.

Data availability statement

Fastq files from the Coriell cell-line and 007 forensic DNA control for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB63471 (https://www.ebi.ac.uk/ena/browser/view/PRJEB63471).

Acknowledgements

This work was partially supported by the Consellería de Cultura, Educación e Ordenación Universitaria e da Consellería de Economía, Emprego e Industria from Xunta de Galicia, Spain (ED481D-2021–008). JRR is supported by the “Programa de axudas á etapa predoutoral” funded by the Consellería de Cultura, Educación e Ordenación Universitaria e da Consellería de Economía, Emprego e Industria from Xunta de Galicia, Spain (ED481A-2020/039). MVL is supported by the Ministerio de Educación, Cultura y Ciencia, Spain (PID2019–107876RB-I00). MdlP is supported by a postdoctoral fellowship awarded by the Gobierno de España: IJC2020–042638-I, funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU/PRTR".

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2023.102937.

References

[1] K.K. Kidd, A.J. Pakstis, W.C. Speed, R. Lagaci, 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, https://doi.org/10.1016/j.fsigen.2014.06.014.

[2] K.K. Kidd, A.J. Pakstis, W.C. Speed, R. Lagace, S. Wootton, J. Chang, Selecting microhaplotypes optimized for different purposes, Electrophoresis 39 (2018) 2815–2823, https://doi.org/10.1002/elps.201800092.

[3] F. Oldoni, K.K. Kidd, D. Podini, Microhaplotypes in forensic genetics, Forensic Sci. Int. Genet. 38 (2019) 54–69, https://doi.org/10.1016/j.fsigen.2018.09.009.

[4] K.K. Kidd, A.J. Pakstis, State of the art for microhaplotypes, Genes 13 (2022), https://doi.org/10.3390/genes13081322.

[5] K.K. Kidd, W.C. Speed, A.J. Pakstis, D.S. Podini, R. Lagaci, 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, https://doi.org/10.1016/j.fsigen.2017.03.014.

[6] M. de la Puente, C. Phillips, C. Xavier, J. Amigo, A. Carracedo, W. Parson, M. V. Lareu, 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.
[7] 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), 983811, https://doi.org/10.3389/fgene.2022.983811.
[8] R. Zhang, J. Xue, M. Tan, D. Chen, Y. Xiao, G. Liu, Y. Zheng, Q. Wu, M. Liao, M. Lv, S. Qu, W. Liang, An MPS-based 50plex microhaplotype assay for forensic DNA analysis, Genes 14 (2023), https://doi.org/10.3390/genes14040865.
[9] R. Wu, H. Chen, R. Li, Y. Zang, X. Shen, B. Hao, Q. Wang, H. Sun, Pairwise kinship testing with microhaplotypes: can advancements be made in kinship inference with these markers? Forensic Sci. Int. 325 (2021), 110875 https://doi.org/10.1016/j.forsciint.2021.110875.
[10] J. Xue, M. Tan, R. Zhang, D. Chen, G. Liu, Y. Zheng, Q. Wu, Y. Xiao, M. Liao, S. Qu, W. Liang, Evaluation of microhaplotype panels for complex kinship analysis using massively parallel sequencing, Forensic Sci. Int. Genet. 65 (2023), 102887, https://doi.org/10.1016/j.fsigen.2023.102887.
[11] E.Y.Y. Cheung, C. Phillips, M. Eduardoff, M.V. Lareu, D. McNevin, Performance of ancestry-informative SNP and microhaplotype markers, Forensic Sci. Int. Genet. 43 (2019), 102141, https://doi.org/10.1016/j.fsigen.2019.102141.
[12] C. Phillips, D. McNevin, K.K. Kidd, R. Lagaci, S. Wootton, M. de la Puente, A. Freire-Aradas, A. Mosquera-Miguel, M. Eduardoff, T. Gross, L. Dagostino, D. Power, S. Olson, M. Hashiyada, C. Oz, W. Parson, P.M. Schneider, M.V. Lareu, R. Daniel, MAPlex - A massively parallel sequencing ancestry analysis multiplex for Asia-Pacific populations, Forensic Sci. Int. Genet. 42 (2019) 213–226, https://doi.org/10.1016/j.fsigen.2019.06.022.
[13] M. de la Puente, J. Ruiz-Ramente, A. Ambroa-Conde, C. Xavier, J. Amigo, Mig Casares de Cal, A. G.sares de, 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.
[14] A.J. Pakstis, N. Gandotra, W.C. Speed, M. Murtha, C. Scharfe, K.K. Kidd, The population genetics characteristics of a 90 locus panel of microhaplotypes, Hum. Genet. 140 (2021) 1753–1773, https://doi.org/10.1007/s00439-021-02382-0.
[15] S. Huang, M. Sheng, Z. Li, K. Li, J. Chen, J. Wu, K. Wang, C. Shi, H. Ding, H. Zhou, L. Ma, J. Yang, Y. Pu, Y. Yu, F. Chen, P. Chen, Inferring bio-geographical ancestry with 35 microhaplotypes, Forensic Sci. Int. 341 (2022), 111509, https://doi.org/10.1016/j.forsciint.2022.111509.
[16] J. Ruiz-Ramdoi, M. de la Puente, C. Xavier, A. Ambroa-Conde, J. Á.broa-Conde, A. Freire-Aradas, A. Mosquera-Miguel, A. Ralf, C. Amory, M.A. Katsara, T. Khellaf, M. Nothnagel, E.Y.Y. Cheung, T.E. Gross, P.M. Schneider, J. Uacyisrael, S. Oliveira, M.D.N. Klautau-Guimarl/1, C. Carvalho-Gontijo, E. Porvalho, W. Branicki, W. Parson, M. Kayser, A. Carracedo, M.V. Lareu, C. Phillips, Development and evaluations of the ancestry informative markers of the VISAGE enhanced tool for appearance and ancestry, Forensic Sci. Int. Genet. 64 (2023), 102853, https://doi.org/10.1016/j.fsigen.2023.102853.
[17] P. Chen, C. Yin, Z. Li, Y. Pu, Y. Yu, P. Zhao, D. Chen, W. Liang, L. Zhang, F. Chen, Evaluation of the microhaplotypes panel for DNA mixture analyses, Forensic Sci. Int. Genet. 35 (2018) 149–155, https://doi.org/10.1016/j.fsigen.2018.05.003.
[18] L. Bennett, F. Oldoni, K. Long, S. Cisana, K. Madella, S. Wootton, J. Chang, R. Hasegawa, R. Lagaca, K.K. Kidd, D. Podini, Mixture deconvolution by massively parallel sequencing of microhaplotypes, Int. J. Leg. Med. 133 (2019) 719–729, https://doi.org/10.1007/s00414-019-02010-7.
[19] 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.
[20] F. Oldoni, D. Bader, C. Fantinato, S.C. Wootton, R. Lagaco, 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.
[21] Q. Zhu, H. Wang, Y. Cao, Y. Huang, Y. Wei, Y. Hu, X. Dai, T. Shan, Y. Wang, J. Zhang, Evaluation of large-scale highly polymorphic microhaplotypes in complex DNA mixtures analysis using RMNE method, Forensic Sci. Int. Genet. 65 (2023), 102874, https://doi.org/10.1016/j.fsigen.2023.102874.
[22] D. Plesivkova, R. Richards, S. Harbison, A review of the potential of the MinION™ single-molecule sequencing system for forensic applications, WIREs Forensic Sci. 1 (2019), e1323, https://doi.org/10.1002/wfs2.1323.
[23] C.L. Hall, R.R. Zascavage, F.J. Sedlazeck, J.V. Planz, Potential applications of nanopore sequencing for forensic analysis, Forensic Sci. Rev. 32 (2020) 23–54.
[24] S. Cornelis, Y. Gansemans, L. Deleye, D. Deforce, F. Van Nieuwerburgh, Forensic SNP genotyping using nanopore MinION sequencing, Sci. Rep. 7 (2017), 41759, https://doi.org/10.1038/srep41759.
[25] S. Cornelis, Y. Gansemans, A.-S. Vander Plaetsen, J. Weymaere, S. Willems, D. Deforce, F. Van Nieuwerburgh, Forensic tri-allelic SNP genotyping using nanopore sequencing, Forensic Sci. Int. Genet. 38 (2019) 204–210, https://doi.org/10.1016/j.fsigen.2018.11.012.
[26] O. Tytgat, Y. Gansemans, J. Weymaere, K. Rubben, D. Deforce, F. Van Nieuwerburgh, Nanopore sequencing of a forensic STR multiplex reveals loci suitable for single-contributor STR profiling, Genes 11 (2020), https://doi.org/10.3390/genes11040381.
[27] Z.-L. Ren, J.-R. Zhang, X.-M. Zhang, X. Liu, Y.-F. Lin, H. Bai, M.-C. Wang, F. Cheng, J.-D. Liu, P. Li, L. Kong, X.-C. Bo, S.-Q. Wang, M. Ni, J.-W. Yan, Forensic nanopore sequencing of STRs and SNPs using verogeni suitable for single-contributor STR Profili, Int. J. Leg. Med. 135 (2021) 1685–1693, https://doi.org/10.1007/s00414-021-02604-0.
[28] Z. Wang, L. Qin, J. Liu, L. Jiang, X. Zou, X. Chen, F. Song, H. Dai, Y. Hou, Forensic nanopore sequencing of microhaplotype markers using QitanTechfor Sing, Forensic Sci. Int. Genet. 57 (2022), 102657, https://doi.org/10.1016/j.fsigen.2021.102657.
[29] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinforma. Oxf. Engl. 25 (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.
[30] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence Alignment/Map format and SAMtools, Bioinforma. Oxf. Engl. 25 (2009) 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.
[31] N. Thomas, Micro: Micro Constr. Vis. (2019).
[32] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2020. ⟨https://www.R-project.org/⟩.
[33] J.T. Robinson, H. Thorvaldsdw.R-p, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat. Biotechnol. 29 (2011) 24–26, https://doi.org/10.1038/nbt.1754.
[34] A. Khanna, D.E. Larson, S.N. Srivatsan, M. Mosior, T.E. Abbott, S. Kiwala, T.J. Ley, E.J. Duncavage, M.J. Walter, J.R. Walker, O.L. Griffith, M. Griffith, C.A. Miller, Bam-readcount – rapid Gener. basepair-Resolut. Seq. Metr. (2021).
[35] A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korbel, J. L. Marchini, S. McCarthy, G.A. McVean, G.R. Abecasis, A global reference for human genetic variation, Nature 526 (2015) 68–74, https://doi.org/10.1038/nature15393.
[36] M. de la Puente, C. Phillips, M. Fondevila, M. Gelabert-Besada, Á. Carracedo, M. V. Lareu, A forensic multiplex of nine novel pentameric-repeat STRs, Forensic Sci. Int. Genet. 29 (2017) 154–164, https://doi.org/10.1016/j.fsigen.2017.04.007.
[37] C. Phillips, L. Fernandez-Formoso, M. Garcia-Magarimos, L. Porras, T. Tvedebrink, J. Amigo, M. Fondevila, A. Gomez-Tato, J. Alvarez-Dios, A. Freire-Aradas, A. Gomez-Carballa, A. Mosquera-Miguel, A. Carracedo, M.V. Lareu, Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel, Forensic Sci. Int. Genet. 5 (2011) 155–169, https://doi.org/10.1016/j.fsigen.2010.02.003.
[38] C. Phillips, L. Fernandez-Formoso, M. Gelabert-Besada, M. Garcbert-Besadas, J. Amigo, A. Carracedo, M.V. Lareu, Global population variability in Qiagen Investigator HDplex STRs, Forensic Sci. Int. Genet. 8 (2014) 36–43, https://doi.org/10.1016/j.fsigen.2013.07.006.
[39] D. Kling, A.O. Tillmar, T. Egeland, Familias 3.org/10.1016/j.fsigen.2013.07.006s, Forensic Sci. Int. Genet. 13 (2014) 121–127, https://doi.org/10.1016/j.fsigen.2014.07.004.
[40] Y. Ni, X. Liu, Z.M. Simeneh, M. Yang, R. Li, Benchmarking of nanopore R10.4 and R9.4.1 flow cells in single-cell whole-genome amplification and whole-genome shotgun sequencing, Comput. Struct. Biotechnol. J. 21 (2023) 2352–2364, https://doi.org/10.1016/j.csbj.2023.03.038.