MPS reveals isometric PCR artefacts in degraded samples

Paolo Fattorini[a,*], Barbara Gornjak Pogorelc[b], Solange Sorcaburu Cilieri[a], Tomaž Zupanc[b], Carlo Previderè[c], Irena Zupanič Pajnič[b]

[a] Department of Medicine, Surgery and Health, University of Trieste, Italy
[b] Institute of Forensic Medicine, Faculty of Medicine, University of Ljubljana, Slovenia
[c] Department of Public Health, Experimental and Forensic Medicine, Section of Legal Medicine and Forensic Sciences, University of Pavia, Italy

ARTICLE INFO

Keywords:
Massive parallel sequencing
Short tandem repeats
Globalfiler
PCR artefacts
DNA degradation

ABSTRACT

It is well known that DNA damage promotes PCR artefacts. In addition, as MPS provides the most accurate typing of PCR products, this technology should be able to highlight more PCR artefacts than the conventional CE approach. To test this hypothesis, a DNA sample was degraded by heating at 70 °C for 8, 16 and 24 h. The three resulting samples #1, #2 and #3 and the untreated control sample were then analysed by the Precision ID Globalfiler NGS STR Panel v2, in duplicate tests. The data analysis was performed by the Converge v2.1 software using the default setting parameters.

Artefacts were identified only in the degraded samples #2 and #3. Few well known artefacts (allelic imbalance, stutter products and allelic drop out) and several drop in phenomena were flagged by the software. The majority (18 out 22) of these drop ins were sequenced as "isometric artefacts", *e.g.* PCR artefacts which cannot be identified by the employment of CE. The molecular features of these PCR artefacts, as well as their risk to be confused with a DNA mixture, are described.

1. Introduction

PCR amplification of autosomal STR markers followed by CE separation of the amplicons is the current gold standard for human identification in forensics [1]. In the last ten years, however, massive parallel sequencing (MPS) technology has been under continuous development for forensic purposes and several commercial kits are nowadays available [2,3]. The PCR-MPS analysis provides the most accurate typing of the STR loci, as sequence variants are typed both within the tandem repeated motifs and in their flanking regions [2]. In addition, its application offers several benefits over the PCR-CE analysis, in particular when highly degraded and/or mixed samples need to be tested [2].

It is also true, however, that the extreme sensitivity of this new technology could be a disadvantage; for example, MPS could be able to highlight more PCR artefacts than CE, thus introducing further uncertainty to the interpretation of the results. To test this hypothesis, a DNA sample was degraded *in vitro* and analysed by the Precision ID Globalfiler NGS STR Panel v2. The results of this preliminary study are reported.

2. Materials and methods

DNA from a blood sample (sample "FM") was degraded by heating at 70 °C for 8, 16 and 24 h following the protocol described in [4]. The resulting samples #1, #2 and #3 showed a Degradation Index (as assessed by PowerQuant kit) of about 30, 180 and 500, respectively [5]. One nanogram of each of the three degraded samples and of the untreated control sample "FM" was analysed by the Precision ID Globalfiler NGS STR Panel v2, in duplicate tests, through 24 cycles of target enrichment PCR [6]. The libraries were prepared on the HID Ion Chef™ System, loaded into an Ion 520 Chip and run on an Ion S5 sequencer. The data analysis was performed by the Converge v2.1 software, using the default setting parameters [6]. A minimum coverage of 100x was adopted for locus call [7].

3. Results

The raw data of this run satisfied the recommended parameters [6] for the Ion 520 Chip on the Ion S5 platform (data not shown). Full profiles were obtained from each of the eight samples. No artefact was found in both the control and the less damaged sample #1. Well known artefacts, such as allelic imbalance, stutter products and allelic drop out

* Corresponding author at: Ospedale di Cattinara, Strada per Fiume 447, 34149, Trieste, Italy.
E-mail address: fattorin@units.it (P. Fattorini).

were found with low frequency in samples #2 and #3 (see Supplementary Table 1). Drop in phenomena, flagged by the software as additional alleles, were scored with remarkable frequency in samples #2 and 3# (see Supplementary Table 1). The majority (18 out 22) of these PCR products were sequenced as "isometric artefacts" while the remaining were either –1bp (two cases) or + n4 drop ins (two cases).

The molecular features of the "isometric artefacts" found in this study were the following: same length of the parental allele with at least one nucleotidic substitution within the STR motive; coverage from 29x to 363x (average value: 145x ± 104x; median value: 116x); coverage ratio with the parental allele of 0.1–0.9 (median value: 0.2). The results showed that loci D2S1338, D19S433 and TPOX were the most prone to this kind of artefacts.

4. Discussion

PCR-MPS analysis provides the most accurate definition of STR loci, both within the STR motifs and in their flanking regions [2]. In the present work, a well characterized DNA sample [4] underwent *in vitro* degradation by heating, a procedure which enriches the samples in Apurinic-Apyrimidinic (A-A) sites [8]. One nanongram of these trial samples were then analysed by the Precision ID Globalfiler NGS STR Panel v2, as recommended in its Application Guide, in duplicate tests, at the suggested number of PCR cycles [6].

This procedure allowed to obtain full profiles from each of the eight samples, but several artefacts occurred in the most damaged samples #2 and #3 (see Supplementary Table 1). Besides few well-known artefacts (such as "allelic imbalance" and "allele drop out"), most of the artefacts were represented by "drop in" phenomena. Out of these 22 "drop ins", 18 were sequenced as "isometric artefacts".

"Noise artefacts" (*e.g.* molecular products showing the same length of the parental alleles and at least one nucleotidic substitution within the STR motifs) have already been described by Young et al. as occurring at low coverage (usually less than 5-10x), even in the four control samples used there [9]. The present study shows that damaged samples promote the synthesis of "isometric artefacts" (*e.g.*, "noise artefacts" at high coverage) leading up to eight drop ins/test (see sample #3_test 1 of Supplementary Table 1). In addition, this study shows that the presence of these artefacts could interfere with the interpretation of the result in real-casework sample, as a DNA mixture could be suspected [2,10]. Therefore, replicate tests represent the recommended approach to tackle these ambiguous situations. Sequencing errors [11] are not expected to be the cause of these artefacts, while it is likely that they rose in the first cycles of the target enrichment procedure because of misinsertions promoted by the A–P lesions [8].

In conclusion, the present study shows that NSG-based technology is able to show "isometric artefacts", which cannot be identified by the employment of the conventional CE approach. Future studies are needed in order to evaluate if the molecular structure of the STR loci is involved in the origin of these artefacts.

Declaration of Competing Interest

The Authors declare that they have no conflict of interest.

Acknowledgements

The study was partially financially supported by the Slovenian Research Agency (project "Determination of the most appropriate skeletal elements for molecular genetic identification of aged human remains" (J3-8214)).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.fsigss.2019.10.015.

References

[1] J.M. Butler, Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers, 2nd ed., Elsevier Academic Press, New York, 2005.

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

[3] C. Hollard, L. Ausset, Y. Chantrel, S. Jullien, M. Clot, et al., Automation and developmental validation of the ForenSeqTM DNA Signature Preparation kit for high-throughput analysis in forensic laboratories, Forensic Sci. Int. Genet. 40 (2019) 37–45.

[4] P. Fattorini, G. Marrubini, S. Bonin, B. Bertoglio, P. Grignani, et al., Producing standard damaged DNA samples by heating: pitfalls and suggestions, Anal. Biochem. 549 (2018) 107–112.

[5] Promega Corporation, PowerQuant® System, Instructions for Use of Products PQ5002 and PQ5008, Technical manual Revised 3/18.

[6] ThermoFisher Scientific, Precision ID GlobalFiler™ NGS STR panel v2 with the HID ion S5™/HID ion GeneStudio™ S5 system, Appl. Guide 15 (November) (2018).

[7] A. Buchard, M.-L. Kampmann, L. Poulsen, C. Børsting, N. Morling, ISO 17025 validation of a next-generation sequencing assay for relationship testing, Electrophoresis 37 (2016) 2822–2831.

[8] K.A. Eckert, T.A. Kunkel, DNA polymerase fidelity and the polymerase chain reaction, Genome Res. PCR Methods Appl. 1 (1991) 17–24.

[9] B. Young, J.L. King, B. Budowle, L. Armogida, A technique for setting analytical thresholds in massively parallel sequencing-based forensic DNA analysis, PLoS One 12 (5) (2017) e0178005, https://doi.org/10.1371/journal.pone.0178005 Published online 2017 May 18.

[10] P. Müller, A. Alonso, P.A. Barrio, B. Berger, M. Bodner, et al., Systematic evaluation of the early access applied biosystems precision ID Globalfiler mixture ID and Globalfiler NGS STR panels for the ion S5 system, Forensic Sci. Int. Genet. 36 (2018) 95–103.

[11] E. Salata, A. Agostino, I. Ciuna, S. Wootton, L. Ripani, A. Berti, Revealing the challenges of low template DNA analysis with the prototype Ion AmpliSeq™ Identity panel v2.3 on the PGM™Sequencer, Forensic Sci. Int. Genet. 22 (2016) 25–36.