



Statistical analysis tools of mixture DNA samples: When the same software provides different results

Camila Costa^{a,b,*}, Carolina Figueiredo^{a,b}, António Amorim^{a,b,c}, Lourdes Prieto^{d,e}, Sandra Costa^f, Paulo Miguel Ferreira^f, Nádía Pinto^{b,c,g}

^a FCUP – Faculdade de Ciências da Universidade do Porto, Portugal

^b i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal

^c IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Portugal

^d Grupo de Medicina Xenómica, Instituto de Ciencias Forenses, Universidad de Santiago de Compostela, Santiago de Compostela, Spain

^e Comisaría General de Policía Científica, Laboratorio ADN, Madrid, Spain

^f LPC-PJ – Biología, Laboratório de Polícia Científica da Polícia Judiciária, Lisboa, Portugal

^g CMUP – Centro de Matemática da Universidade do Porto, Portugal

ARTICLE INFO

Keywords:

Weight-of-evidence
Genotyping software
LRmix Studio
EuroForMix
STRmix
Software parameters

ABSTRACT

The high complexity of the genetic analysis of crime scene samples is mainly related to the unknown number of contributors, low DNA quantity and quality, and associated stochastic effects. The difficulty and subjectivity of interpreting casework samples was the motto for the development of software to mitigate these conditions and allow the quantification of the genetic evidence. Currently, there are several tools for statistical analysis of mixture samples based on either qualitative or quantitative models. The first considers the electropherograms' qualitative information, while the latter also considers the associated quantitative information. This work's main goal was to evaluate the effect that parameters' settings variation may have on the LR computation, specifically the drop-in frequency parameter. For that, a qualitative – LRMix Studio – and two quantitative software – STRmix™ and EuroForMix – were considered and an intra-software analysis was performed, using as input real casework samples. The drop-in frequency variation showed an impact, leading to differences higher than four units (log₁₀ scale) for some pairs of samples. In addition, for some cases, no comparisons were performed either because the tool computed a null LR value or displayed an error message. Thus, this work reinforces the importance of proper parameters' modeling and estimation in forensic casework evaluation.

1. Introduction

DNA samples recovered from crime scenes may have associated several characteristics that make them difficult and complex for the expert to interpret and analyze. DNA mixtures, i.e., samples containing contributions from more than one donor (exact number unknown), are typically encountered in this context. Beyond the unknown number of contributors (NOC), the high complexity also results from allele sharing between contributors, contributions in different proportions, stutters which may be confounded with alleles of a minor contributor (or vice versa), amplification stochastic effects (heterozygotic imbalance, drop-in, and drop-out) due to low-template DNA, or degradation [1–12].

To deal with the abovementioned conditions and quantify the genetic evidence, several probabilistic genotyping software were

developed. These tools are based on either qualitative or quantitative models. The first only considers the qualitative information of the electropherogram (observed alleles), while the other also considers the quantitative information (height of the detected peaks) [13,14]. Both models quantify the genetic evidence through the computation of a Likelihood Ratio (LR) [4], comparing the probabilities of observing the genetic evidence assuming two alternative and mutually exclusive hypotheses. This calculation can account for several parameters, regarding population and analytical factors, on which the LR depends, and which are introduced by the user, such as NOC, drop-in, co-ancestry coefficient, and threshold detection.

The main goal of this work is to evaluate the impact that the estimation of the drop-in frequency parameter may have on the quantification of the evidence, using casework samples. Both qualitative –

* Correspondence to: i3S, Rua Alfredo Allen, 208, 4200-135 Porto, Portugal.
E-mail address: camilac@ipatimup.pt (C. Costa).

<https://doi.org/10.1016/j.fsigss.2022.09.014>

Received 19 September 2022; Accepted 23 September 2022

Available online 24 September 2022

1875-1768/© 2022 Elsevier B.V. All rights reserved.

LRmix Studio v.2.1.3 [15] – and quantitative tools – EuroForMix v. 3.4.0 [16] and STRmix™ v.2.7 [17] were considered. A drop-in allele corresponds to a spurious allele in the problem sample profile, i.e., an allele that shows up in the problem sample profile but cannot be explained by any of the contributors [6]. This stochastic effect generates discordances between the mixture and the contributors' profiles, which can result in misinterpretations and misleading conclusions. So, understanding its impact is extremely important [10,11].

2. Material and methods

Resorting to evidence from *Laboratório de Polícia Científica da Polícia Judiciária* former cases, a set of 156 irreversibly anonymized mixture/single contributor sample pairs were selected. For each case the single contributor sample was either from sampled individuals or single-source problem sample profiles obtained in the same case of the analyzed mixture. It was assumed as belonging to the person of interest (POI) but are not necessarily known. The mixture samples considered were estimated to have either two or three NOC. Casework samples were chosen due to their unknown composition, uniqueness, and inability to predict and replicate, carrying a much higher complexity than mock ones.

For each pair selected, a LR value was computed assuming the alternative hypotheses: “The POI is a contributor of the mixture” and “The POI is unrelated to any contributor of the mixture”. The LR values were computed using both qualitative (LRmix Studio v.2.1.3 [15]) and quantitative (EuroForMix v.3.4.0 [16] and STRmix™ v.2.7 [17]) tools. To understand the impact that drop-in frequency may have on the LR computation, the frequency was varied to lower (0.00) and higher values (0.10) relative to the established default (0.05) [18]. Afterward, an intra-software analysis was carried out through a comparison between the values computed under the default and varied conditions for each pair of samples.

The allelic frequencies of the National Institute of Standards and Technology (NIST) database concerning the Caucasian population were used [19].

3. Results and discussion

In general, the variation of drop-in frequency did not have a great impact on the LR values provided by the same software, as most of the computed differences were within one unit in a log₁₀ scale. For all the qualitative software comparisons (LRmix Studio), no calculated difference exceeded one log₁₀ unit, reflecting its lower sensitivity to this parameter variation compared to quantitative tools.

For both quantitative tools, differences higher than four log₁₀ units were observed - in EuroForMix when the drop-in frequency was considered null (one case with a mixture with two, and one case with a mixture with three estimated NOC), and in STRmix™ when the drop-in frequency was considered either as null or equal to 0.10 (two cases for each comparison; mixtures with three estimated NOC).

The impact of parameter's variation also translates into cases for which no comparison was performed, as the software either computed a null LR value – as STRmix™ – or does not compute a value at all – as both LRMix Studio and EuroForMix. This was observed when a null drop-in frequency was considered. For LRMix Studio and EuroForMix, this occurred in 25% and 27% (respectively) of the cases involving mixtures with two estimated NOC, and in 4% and 5% (respectively) of the cases with mixtures with three estimated. In STRmix™, this occurred in only 1% of the cases for both mixtures with two and three estimated contributors.

4. Conclusion

Before performing any computation, the expert needs to establish and introduce a value for some software parameters, such as NOC, F_{ST}, drop-in, or detection threshold. Thus, it is crucial to understand the

impact those parameters may have on the resulting LR. Focusing on the drop-in frequency parameter, LR values were computed under both the established default and varied values of the drop-in frequency, for each pair analyzed case. The results obtained by LRMix Studio, EuroForMix, and STRmix™ were compared and the impact within software was evaluated.

LRmix Studio showed to be less sensitive to this parameter variation than the quantitative tools. However, in both quantitative differences of more than four units (log₁₀ scale) were observed. Furthermore, for all software, there were cases for which no comparisons were performed due to the computation of a null LR value or an error message display.

This work emphasizes the importance of experts' understanding of the models incorporated in the existing tools and of its proper estimation, specifically for the case of drop-in.

Conflict of interest statement

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.

Acknowledgements

This project was supported by the Laboratório de Polícia Científica da Polícia Judiciária (LPC-PJ), Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP) and Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal (i3S). This work was partially financed by FEDER – Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 – Operational Program for Competitiveness and Internationalization (POCI), Portugal 2020, and by Portuguese funds through FCT – Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Inovação in the framework of the projects “Institute for Research and Innovation in Health Sciences” (POCI-01–0145-FEDER-007274). NP is supported by FCT, under the program contract provided in Decree-Law no.57/2016 of August 29. CC is funded by a FCT doctoral grant 2021.05655. BD.

References

- [1] P.S. Walsh, N.J. Fildes, R. Reynolds, Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA, *Nucleic Acids Res.* 24 (14) (1996) 2807–2812.
- [2] P. Gill, Application of low copy number DNA profiling, *Croat. Med. J.* 42 (3) (2001) 229–232.
- [3] P.M. Schneider, K. Bender, W.R. Mayr, W. Parson, B. Hoste, R. Decorte, J. Cordonnier, D. Vanek, N. Morling, M. Karjalainen, et al., STR analysis of artificially degraded DNA-results of a collaborative European exercise, *Forensic Sci. Int.* 139 (2–3) (2004) 123–134.
- [4] P. Gill, C.H. Brenner, J.S. Buckleton, A. Carracedo, M. Krawczak, W.R. Mayr, N. Morling, M. Prinz, P.M. Schneider, B.S. Weir, et al., DNA commission of the International Society of Forensic Genetics: recommendations on the interpretation of mixtures, *Forensic Sci. Int.* 160 (2–3) (2006) 90–101.
- [5] M. Fondevila, C. Phillips, N. Naverán, M. Cerezo, A. Rodríguez, R. Calvo, L. M. Fernández, A. Carracedo, M.V. Lareu, D.N.A. Challenging, Assessment of a range of genotyping approaches for highly degraded forensic samples, *Forensic Sci. Int. Genet. Suppl. Ser.* 1 (1) (2008) 26–28.
- [6] D.J. Balding, J. Buckleton, Interpreting low template DNA profiles, *Forensic Sci. Int. Genet.* 4 (1) (2009) 1–10.
- [7] A.J. Gibb, A.L. Huell, M.C. Simmons, R.M. Brown, Characterisation of forward stutter in the AmpFISTR SGM Plus PCR, *Sci. Justice* 49 (1) (2009) 24–31.
- [8] A.A. Westen, J.H. Nagel, C.C. Benschop, N.E. Weiler, B.J. de Jong, T. Sijen, Higher capillary electrophoresis injection settings as an efficient approach to increase the sensitivity of STR typing, *J. Forensic Sci.* 54 (3) (2009) 591–598.
- [9] A. Freire-Aradas, M. Fondevila, A.K. Kriegel, C. Phillips, P. Gill, L. Prieto, P. M. Schneider, A. Carracedo, M.V. Lareu, A new SNP assay for identification of highly degraded human DNA, *Forensic Sci. Int. Genet.* 6 (3) (2012) 341–349.
- [10] S. Gittelson, A. Biedermann, S. Bozza, F. Taroni, Decision analysis for the genotype designation in low-template-DNA profiles, *Forensic Sci. Int. Genet.* 9 (2014) 118–133.
- [11] C.D. Steele, M. Greenhalgh, D.J. Balding, Evaluation of low-template DNA profiles using peak heights, *Stat. Appl. Genet. Mol. Biol.* 15 (5) (2016) 431–445.
- [12] H.R. Dash, P. Shrivastava, S. Das, Analysis of capillary electrophoresis results by GeneMapper® ID-X v 1.5 Software. Principles and Practices of DNA Analysis: A Laboratory Manual for Forensic DNA Typing, Springer, 2020, pp. 213–237.

- [13] E. Alladio, M. Omedei, S. Cisana, G. D'Amico, D. Caneparo, M. Vincenti, P. Garofano, DNA mixtures interpretation - a proof-of-concept multi-software comparison highlighting different probabilistic methods' performances on challenging samples, *Forensic Sci. Int. Genet.* 37 (2018) 143–150.
- [14] M.D. Coble, J.A. Bright, Probabilistic genotyping software: an overview, *Forensic Sci. Int. Genet.* 38 (2019) 219–224.
- [15] H. Haned, K. Slooten, P. Gill, Exploratory data analysis for the interpretation of low template DNA mixtures, *Forensic Sci. Int. Genet.* 6 (6) (2012) 762–774.
- [16] O. Bleka, G. Storvik, P. Gill, EuroForMix: an open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts, *Forensic Sci. Int. Genet.* 21 (2016) 35–44.
- [17] D. Taylor, J.A. Bright, J. Buckleton, The interpretation of single source and mixed DNA profiles, *Forensic Sci. Int. Genet.* 7 (5) (2013) 516–528.
- [18] H. Haned, C.C. Benschop, P.D. Gill, T. Sijen, Complex DNA mixture analysis in a forensic context: evaluating the probative value using a likelihood ratio model, *Forensic Sci. Int. Genet.* 16 (2015) 17–25.
- [19] C.R. Hill, D.L. Duewer, M.C. Kline, M.D. Coble, J.M. Butler, US population data for 29 autosomal STR loci, *Forensic Sci. Int. Genet.* 7 (3) (2013) e82–e83.