

Article

Application of Probabilistic Genotyping Software to Paternity Cases Involving Low-Template DNA [†]

Alessia Riem [‡] , Elena Chierito [‡] , Federica Bertolotto, Marco Parnigoni, Serena Aneli  and Carlo Robino ^{*}

Department of Public Health Sciences and Pediatrics, University of Turin, 10126 Turin, Italy; alessia.riem@unito.it (A.R.); elena.chierito@unito.it (E.C.); bertolottof@gmail.com (F.B.); marco.parnigoni@unito.it (M.P.); serena.aneli@unito.it (S.A.)

^{*} Correspondence: carlo.robino@unito.it; Tel.: +39-011-670-5917

[†] This article is a revised and expanded version of a paper entitled ‘Exploring different software options for likelihood ratio calculations in paternity cases with low template DNA’, which was presented at 30th Congress of the International Society for Forensic Genetics, Santiago de Compostela, Spain 9–13 September 2024.

[‡] These authors contributed equally to this work.

Abstract

Background: Interpreting short tandem repeat (STR) profiles from low-template DNA (LT-DNA) requires consideration of the stochastic phenomena that can affect the reliability of genotypes. Although several probabilistic genotyping tools have been developed to model such uncertainties, most have only been used for direct comparisons between persons of interest and crime scene samples. Their application to kinship testing involving LT-DNA has received comparatively little attention. **Methods:** We evaluated the performance of two PGS, EuroForMix (EFM) and EFMrep, which support alternative hypotheses with relatedness, by comparing them with a standard paternity testing software (Familias) in 33 paternity cases involving LT-DNA samples categorised as ‘mildly’ (MD) or ‘highly’ (HD) degraded based on the quality of the STR profiles. The samples included formalin-fixed paraffin-embedded tissues, bone specimens, and stains collected from personal items. Pedigrees with (‘trio’) and without (‘duo’) maternal information were considered. **Results:** In MD and HD duos, the likelihood ratios (LRs) obtained with EFMrep were significantly higher compared to other software. In trios, Familias produced significantly higher LRs than PGS for MD samples, whereas the three software performed comparably for HD samples. Notably, in HD trios, EFMrep was the software most likely to maximise LR values, which were above 10,000 in 60% of the cases, compared to 50% of EFM and 40% of Familias. **Conclusions:** These findings provide preliminary evidence of the potential and limitations of using PGS for kinship assessments involving LT-DNA specimens.

Keywords: kinship analysis; paternity testing; low-template DNA; probabilistic genotyping



Academic Editor: Fulvio Cruciani

Received: 15 December 2025

Revised: 30 December 2025

Accepted: 26 January 2026

Published: 1 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

In forensic genetics, the PCR amplification of short tandem repeat (STR) loci in low-template DNA (LT-DNA) samples is affected by stochastic variation [1]. This can lead to the unbalanced amplification of alleles, generating so-called drop-out and drop-in events. Drop-out events consist of the failure to amplify an allele that is actually present in the genotype of the source individual, while drop-in events consist of the observation of additional alleles that are not part of the source individual’s genotype [2]. The occurrence of stochastic events is favoured by suboptimal amounts of input DNA in the amplification reactions of STR loci and by the reduced number of DNA templates accessible to PCR primers in

cases of DNA degradation [3]. Another feature of LT-DNA samples that may complicate interpretation is the possible occurrence of large stutters [4]. These are artefactual peaks, typically one repeat unit shorter or longer than the actual allelic peak, that are the result of polymerase slippage during the process of amplifying repetitive STR sequences [5]. These peaks may be mistaken for real alleles in unbalanced PCR results.

Previous studies have shown that all these PCR artefacts can be interpreted within the framework of likelihood ratio (LR) calculations [6]. Probabilistic genotyping software (PGS) that can accommodate drop-out, drop-in and stutter events has been developed and validated accordingly [7,8]. This software can also handle the other complexities of forensic STR profiles, such as mixtures involving multiple contributors. In their most advanced form, known as ‘continuous’ methods, PGS can use quantitative information derived from peak heights in the DNA profile to calculate the probability of the observed peaks, given all possible genotype combinations for the individual contributors [9]. However, most of this software has been specifically designed to facilitate the interpretation of direct matches between a reference profile and a sample of interest in criminal investigations. Less attention has been given to probabilistic genotyping in the context of kinship and paternity testing [10], even though LT-DNA samples are not uncommon in such cases. Consider, for instance, cases of alleged paternity involving a deceased father, where the STR profile must be derived from archival material consisting of formalin-fixed, paraffin-embedded (FFPE) tissue, whose DNA is often fragmented due to the degradative effect of formalin [11]. Similarly, DNA degradation due to biochemical post-mortem modifications of the body is expected in exhumed human remains used for paternity testing [12]. Post-mortem and peri-mortem events, such as those involving fire [13], can also lead to DNA degradation in human remains. This has an impact on kinship calculations, which are carried out when a reference sample of the missing person is unavailable and DNA comparisons with relatives are needed for identification purposes [14].

Software capable of calculating LR in paternity and kinship testing, incorporating features such as the ability to consider maternal genotype information, mutation rates, and different mutation models for STRs, has existed for a considerable time [15]. More recently, these software programs have progressively been implemented to include the probability of drop-out events, which can result in apparent mismatches between genotypes [16,17]. However, software that can fully handle quantitative information to estimate genotype probabilities given all possible genotype sets and combine replicate analyses of STR loci into a single likelihood ratio (LR) value is still limited in availability and diffusion [18,19].

The current recommendations for the biostatistical evaluation of paternity, as developed by the International Society for Forensic Genetics (ISFG), do not address LT-DNA cases and the associated stochastic events [20]. In our laboratory, at least two STR amplification replicates are carried out for paternity/kinship testing of potentially LT-DNA samples, i.e., those derived from unconventional sources other than reference buccal swabs and fresh blood/tissue samples. A ‘consensus’ profile [1] is then obtained and analysed using the drop-out function available in the Familias software [17].

Some continuous PGS support the option of considering the hypothesis that a relative of the person of interest (POI) from whom the reference profile was obtained is the actual donor of the crime stain, rather than the POI him/herself [21]. It is therefore possible to calculate LR values given the proposition that the LT-DNA profile of interest comes from the father–child of a reference person, whose DNA profile is available and uncontested, rather than from a random individual [22]. With this in mind, we decided to investigate the possibility of applying such software to paternity cases involving LT-DNA samples. The software programs considered here were EuroForMix (EFM) [23] and EFMrep [24]. The latter is an extension of EuroForMix, which enables DNA profiling results obtained with

different STR kits to be combined. Like EFM, it includes the option of specifying related individuals in alternative propositions.

2. Materials and Methods

2.1. Samples

A total of 33 LT-DNA samples from 15 different cases submitted to the laboratory between 2019 and 2025 were analysed. Research was authorised by the Bioethics Committee of the University of Turin (Authorisation No. 0532888, 18 July 2025). The cases included: (i) paternity tests involving a deceased alleged father for whom ante-mortem samples were available and consisted of FFPE samples ($n = 12$ samples from 7 cases) or dry blood/saliva stains ($n = 6$ samples from 2 cases); (ii) identification of unknown skeletal remains through comparison with reference samples belonging to an alleged first-degree relative ($n = 15$ samples from 6 cases). Maternal information was available for 21 samples ('trio' cases) and missing in 12 samples ('duo' cases).

2.2. DNA Extraction

DNA extraction was performed using different methods depending on the starting material: (i) FFPE samples: Two to three sections of tissue, 5–10 μm thick, were cut via a microtome from each FFPE sample and deparaffinized using Deparaffinization Solution (Qiagen, Hilden, Germany) followed by DNA purification with a QIAamp DNA FFPE Tissue kit (Qiagen, Hilden, Germany) performed either manually (2019–2021) or automatically (2022–2025) with QIAcube Connect (Qiagen, Hilden, Germany). (ii) Bone: The PrepFiler BTA Forensic Kit (ThermoFisher Scientific, Waltham, MA, USA) was used, following the manufacturer's instructions. The procedure started from 50 mg of bone powder previously obtained through mechanical scraping followed by pulverisation in liquid nitrogen. (iii) Dried stains: The QIAamp DNA Investigator Kit (Qiagen, Hilden, Germany) was used, following the manufacturer's instructions, either manually (2019–2021) or automatically (2022–2025) with QIAcube Connect (Qiagen, Hilden, Germany). (iv) Reference samples: Buccal swabs were extracted using the Investigator STR GO! lysis buffer (Qiagen, Hilden, Germany) following the manufacturer's instructions.

2.3. DNA Quantification

Total human DNA and male-specific DNA isolated from the LT-DNA samples were quantified by qPCR using the Plexor[®] HY System Kit (Promega, Madison, WI, USA) and the CFX96 Touch[™] Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA, USA). Results were analysed with the Plexor Forensics software version 1.6.0 (Promega, Madison, WI, USA).

2.4. Genotyping

Samples were amplified with the PowerPlex ESI 17 Fast (Promega, Madison, WI, USA) ('ESI') and PowerPlex ESX 17 Fast (Promega, Madison, WI, USA) ('ESX') kits, targeting the same panel of STR loci with alternative PCR primers. Primers are designed so that STRs with shorter amplicon sizes in one kit have larger amplicon sizes in the other, and vice versa, so that the combination of the two kits maximises the number of genotyped STR loci [25]. For LT-DNA samples, duplicate amplifications were performed using either the ESI or ESX kit, in addition to a single amplification carried out with the kit not used for the duplicate reactions. The optimal amount of input DNA according to the manufacturer's instructions (0.5 ng) was employed. When this was not possible, due to reduced DNA concentration in LT-DNA samples, the maximum reaction volume available in PCR for input DNA was used. STR genotyping was performed by capillary electrophoresis using

the SeqStudio Genetic Analyzer for HID (ThermoFisher Scientific, Waltham, MA, USA). Data analysis was carried out with GeneMapper ID-X software version 1.7 (ThermoFisher Scientific, Waltham, MA, USA). The analytical threshold (AT), separating reported alleles from electrophoretic baseline noise was set at 50 relative fluorescence units (RFU) for all STR kits.

The degradation status of LT-DNA samples was assessed by considering, in each STR profile, the largest amplicon product with RFU peak height associated with an expected probability of drop-out <5% (5% stochastic threshold, 5ST) [26] according to logistic regression experiments [7]. Samples were classified as highly degraded (HD) when the 5ST largest amplicon length was <200 bp and mildly degraded (MD) when it was ≥ 200 bp.

2.5. Software Programs for LR Calculation

All LR calculations were performed using the allelic frequencies observed in the Italian population [27,28], except for locus SE33 [29], and by applying a 0.01 F_{ST} correction for population structure [30]. The interpretation of LR values was performed taking into account national guidelines for kinship testing, suggesting a threshold of 10,000 for DNA results to be acceptable in court as proof of paternity [31,32].

The specific parameters set according to the characteristics of the software Familias version 3.3.1, EuroForMix version 4.0.8, and EFMrep version 1.1.0 (Table 1) are summarised in the following sections.

Table 1. Characteristics of the software used in this study for the analysis of LT-DNA samples in paternity cases.

	Familias	EFM	EFMrep
Drop-out	Yes	Yes	Yes
Drop-in	No	Yes	Yes
Stutter	No	Yes	Yes
Mutation	Yes	No	No
Maternal information	Yes	No	No
Peak height	No	Yes	Yes
PCR replicas	No (consensus)	Yes (single STR kit)	Yes (multiple STR kits)

2.5.1. Familias

A ‘consensus’ profile was obtained using the ‘2×’ approach [33], which exclusively considers the alleles detected at least twice in the PCR experiments consisting, in this case, of two replicates performed with the same STR kit (ESI or ESX). The criteria for establishing the consensus profile and to assign the corresponding drop-out probabilities in Familias are outlined in Table 2. The probability distribution of drop-out was determined through internal validation and logistic regression (Supplementary Figure S1) [7].

Mean mutation rates for male and female individuals, as described in the 2003 report of the American Association of Blood Banks [34], were used. For loci not considered in that report, a standard mutation rate of 0.2% was adopted. The ‘equal probability’ mutation model was applied in all calculations.

In trio cases, calculations were performed both with and without considering maternal information.

Table 2. Method used to create consensus STR profiles and to assign dropout probabilities.

Peaks in Replicate 1	Peaks in Replicate 2	Consensus Genotype	Drop-Out Probability
A	None or B	None	Locus not considered in calculations
A	A	A–A	Calculated according to logistic regression, using the highest A allele peak among the two replicates
A–B	A	A–A	Calculated according to the logistic regression of A allele peak of replicate 2
A–B	A–C	A–A	Calculated according to logistic regression using the lowest A allele peak among the two replicates
A–B	A–B	A–B	No drop-out

2.5.2. EFM

LT-DNA samples underwent two PCR replicates with the same STR kit (ESI or ESX). The two alternative hypotheses used in the calculations were: (i) Hp: sample comes from an individual unrelated to the reference person; and (ii) Hd: sample comes from the parent/child of the reference person.

Lambda parameters to model drop-in peak height were calculated with the EFM ‘Fit-in drop-in data’ tool, based on internal validation data obtained for the ESI (0.038) and ESX (0.061) STR kits. The exponential drop-in functions for the two STR kits are displayed in Supplementary Figure S2.

The reciprocal of the maximum likelihood estimation (MLE) LR value obtained for the combination of parameters (degradation, forward and backward stutter) that maximised the adjusted loglikelihoods of the competing propositions was then used for comparisons between software.

2.5.3. EFMrep

Each LT-DNA sample was subjected to amplification using two PCR replicates with different STR kits (ESI and ESX). Lambda values were the same as those for EFM. The same combination of parameters maximising LR according to the results obtained with EFM using a single STR kit was selected, and it was assumed that they also applied to the STR kit, which had not been used for EFM calculations. As EFMrep, differently to EFM, supports relatedness as the numerator proposition, the two alternative hypotheses were: (i) Hp: sample comes from the parent/child of the reference person; and (ii) Hd: sample comes from an individual unrelated to the reference person.

The obtained MLE LR value was then used for comparisons between software.

2.6. Biostatistical Analysis

Statistical comparisons of data distributions were computed using the Wilcoxon signed-rank test. Distribution correlations were evaluated using the Pearson correlation. A statistical significance threshold of $p < 0.05$ was applied. All statistical analyses were conducted using R statistical software (version 4.4.1) [35].

3. Results

Of all 33 LT-DNA samples included in the study, 61% were classified as HD ($n = 20$) and 39% as MD ($n = 13$), according to the 5ST stochastic threshold derived from electropherograms. Separated analyses were conducted for duo and trio pedigrees. LT-DNA paternity cases in which maternal information was available ($n = 21$) were used for calculations in both the trio and the duo scenario, disregarding maternal data in the latter case.

The general characteristics of LT-DNA samples and the results of the software analysis are detailed in Supplementary Table S1. The number of loci that were informative of the genotype for LR calculations using the three different software is shown in Table 3. For Familias, only loci of the consensus profile with at least one detected allele were considered, whereas for EFM and EFMrep, loci with at least one allele above AT in one of the two PCR replicates were counted. It was seen that, for HD samples, the number of informative STR loci was significantly lower in Familias compared to EFM ($p = 0.0002$) and EFMrep ($p = 0.0002$). No differences between the three software were seen for MD samples. In HD samples, a significant positive correlation was observed between the number of informative STR loci and the obtained LR values in Familias ($r = 0.75$, $p = 0.0002$), EFM ($r = 0.76$, $p = 0.0001$), and EFMrep ($r = 0.48$, $p = 0.0323$).

Table 3. Mean number of informative STR loci ($\pm$ SD) for the three software in MD and HD LT-DNA cases.

	Familias	EFM	EFMrep
MD (n = 13)	15 $\pm$ 1.6	15.6 $\pm$ 1.1	15.9 $\pm$ 0.3
HD (n = 20)	10.4 $\pm$ 4.3	14 $\pm$ 2.2	15 $\pm$ 1.2

Comparisons of LR calculation results with standard kinship software and PGS are summarised in Figure 1.

It was found that, in duo cases, the LR values generated by EFMrep were significantly higher than those produced by Familias. This finding was consistent for both MD ($p = 0.0002$) and HD ($p = 8.2 \times 10^{-5}$) samples. Furthermore, for HD samples, the LR values generated by EFM were significantly higher than those of Familias ($p = 0.0027$), while no significant differences were detected between the two software programs for MD samples. In both cases, the LR values were significantly higher for EFMrep compared to EFM (MD $p = 0.0012$, HD $p = 0.0064$).

When maternal DNA information was available for MD samples (trios), the LR values produced by Familias were significantly higher than those obtained with EFM ($p = 0.0049$) and EFMrep ($p = 0.0140$). Within the MD samples, EFMrep also yielded significantly higher LR values than EFM ($p = 0.0049$). In contrast, no statistically significant differences among the three software were detected in the HD trio samples.

Figure 2 reports the distribution of LR values obtained with the three software in the different pedigree and degradation scenarios, highlighting the proportion of results below and above a set threshold of $LR \geq 10,000$.

Given the significant differences in the paired LR values between the three software in the MD-duo samples, EFMrep consistently yielded LR values exceeding 10,000 (100%), while EFM showed a lower percentage (84.6%), which was even lower for Familias (61%), as shown in Figure 2A. Similarly, for HD-duos, EFMrep showed a larger proportion of cases exceeding the $LR \geq 10,000$ threshold (65%) compared to EFM (55%) and Familias (30%). In MD-trio cases, although Familias produces on average higher LR values compared to PGS, it could be seen that, in 9% of the cases, Familias' LR values were below 10,000, whereas values obtained with EFMrep always passed the set threshold. Finally, for HD-trios, as shown in Figure 2B, where all three software did not show significant differences in terms of LR values, EFMrep showed a larger proportion of cases exceeding the $LR \geq 10,000$ threshold (60%) compared to EFM (50%) and Familias (40%).

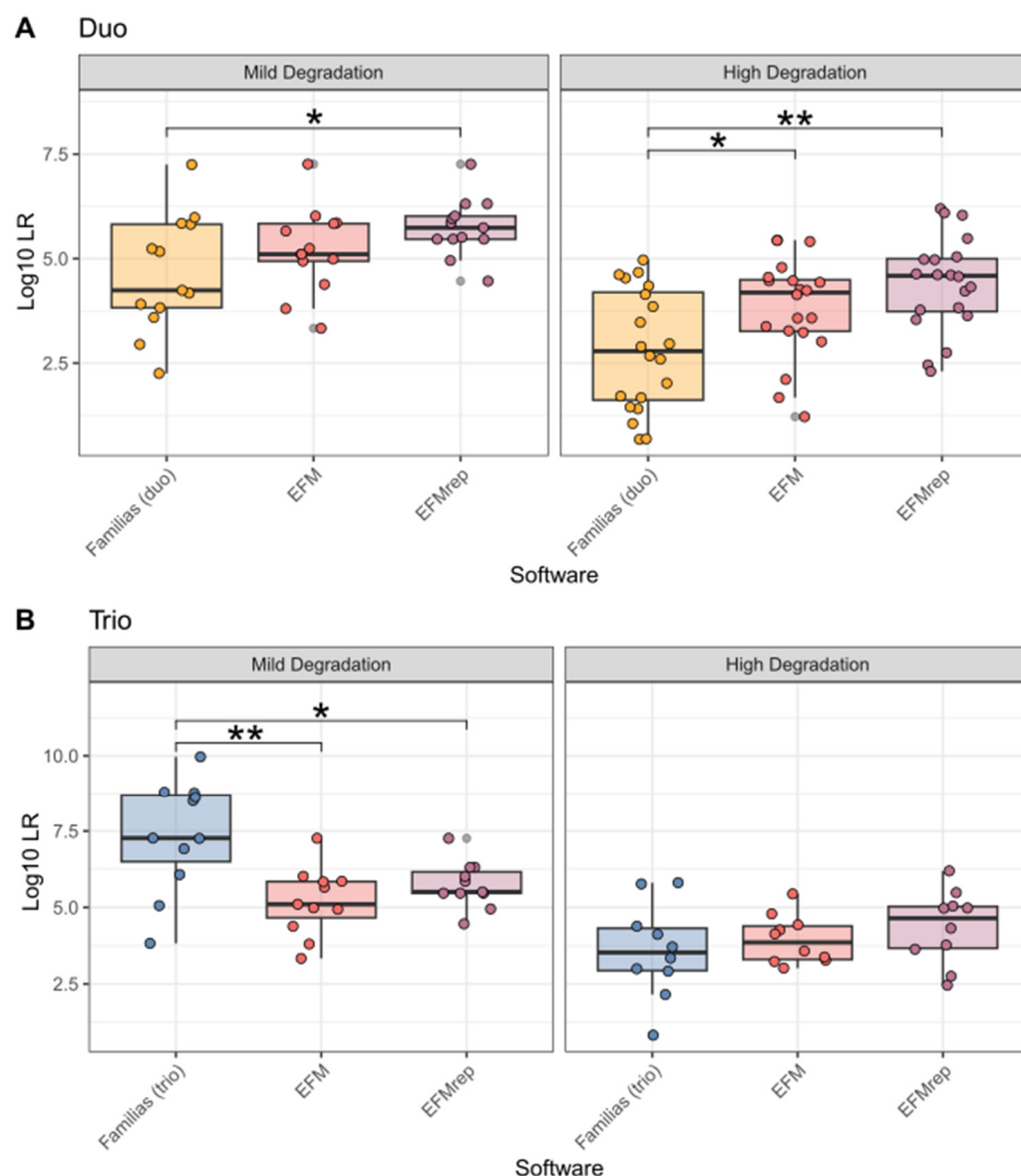


Figure 1. Boxplot showing LR values calculated by each software in duo (A) and trio (B) cases, divided by the degree of degradation of LT-DNA samples. Each dot represents a sample. “*”: $p < 0.05$, “**”: $p < 0.01$.

LR values obtained in each LT-DNA paternity case are depicted in Figure 3. As summarised in Table 4, the maximum difference between the LR values produced for the same LT-DNA samples by the three different software expressed in orders of magnitude (ΔLR_{max}) was 4. Table 4 also shows that, for all pedigree and degradation scenarios except for MD-trios, the software most likely to maximise LR values was EFMrep. On the other end, in the majority of MD-trios (82%), the software that generated the highest LR values was Familias. Overall, it could also be seen that in cases where the LR values obtained with Familias fell below the 10,000 threshold ($n = 10$, 30%), LR values above 10,000 could be obtained with PGS. In particular, LR values $\geq 10,000$ were observed with EFM in 20% of the cases ($n = 2$), with EFMrep in 20% of the cases ($n = 2$) and with both software in the remaining 60% of the cases ($n = 6$).

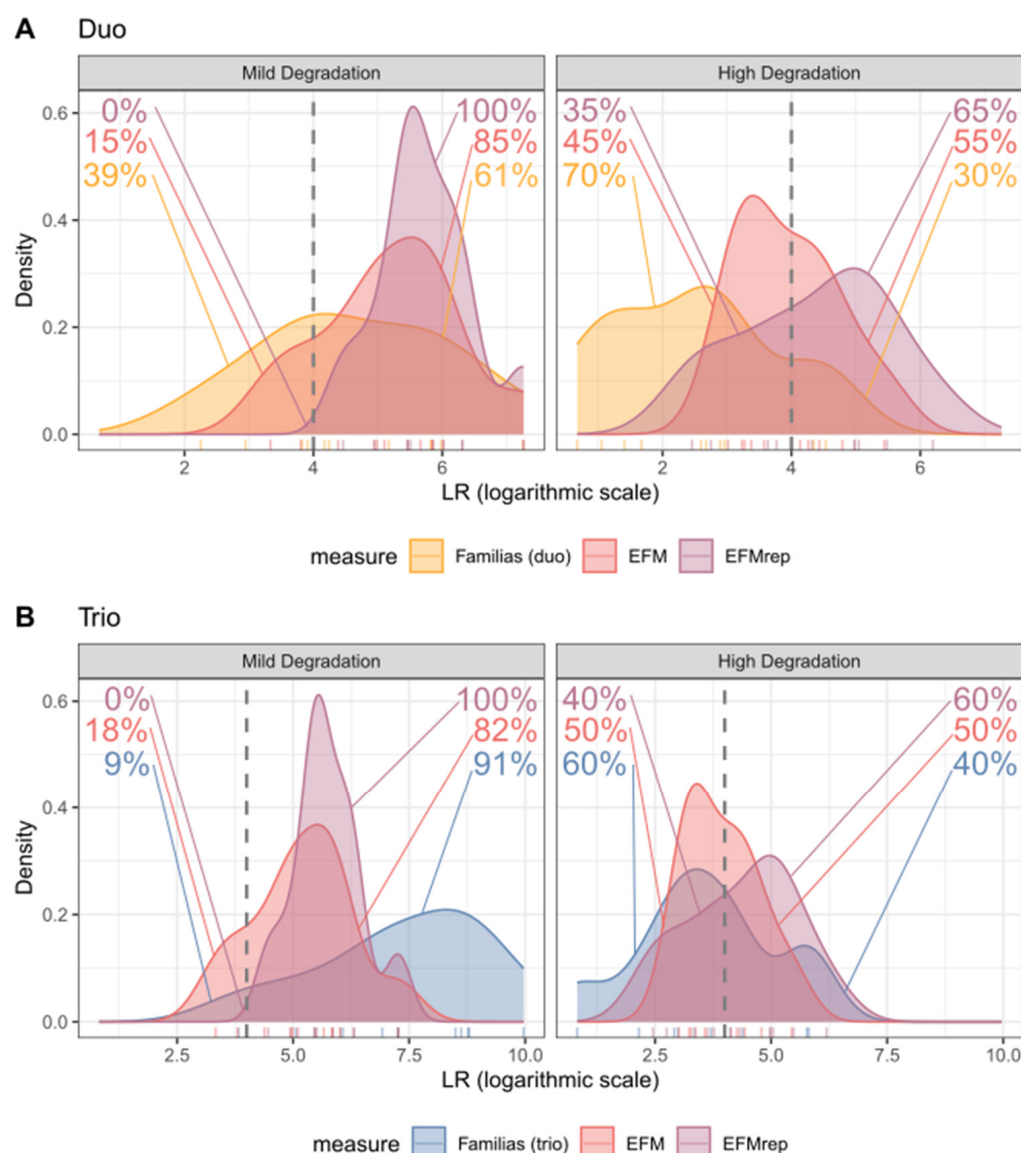


Figure 2. Density plot showing the LR values generated by each software in duo (A) and trio (B) cases, subdivided by the degree of degradation of the LT-DNA samples. The dashed line corresponds to LR = 10,000.

Table 4. Δ LR max values observed with the three software in sample categories subdivided by LT-DNA degradation level and pedigree type, shown on the left. Columns on the right report the percentage of cases in which, within a specific scenario, the indicated software produced the highest LR value.

	Δ LR Max	Familias	EFM	EFMrep
MD-duo (n = 13)	4	0%	15%	85%
HD-duo (n = 20)	4	10%	10%	80%
MD-trio (n = 11)	4	82%	0%	18%
HD-trio (n = 10)	3	20%	20%	60%

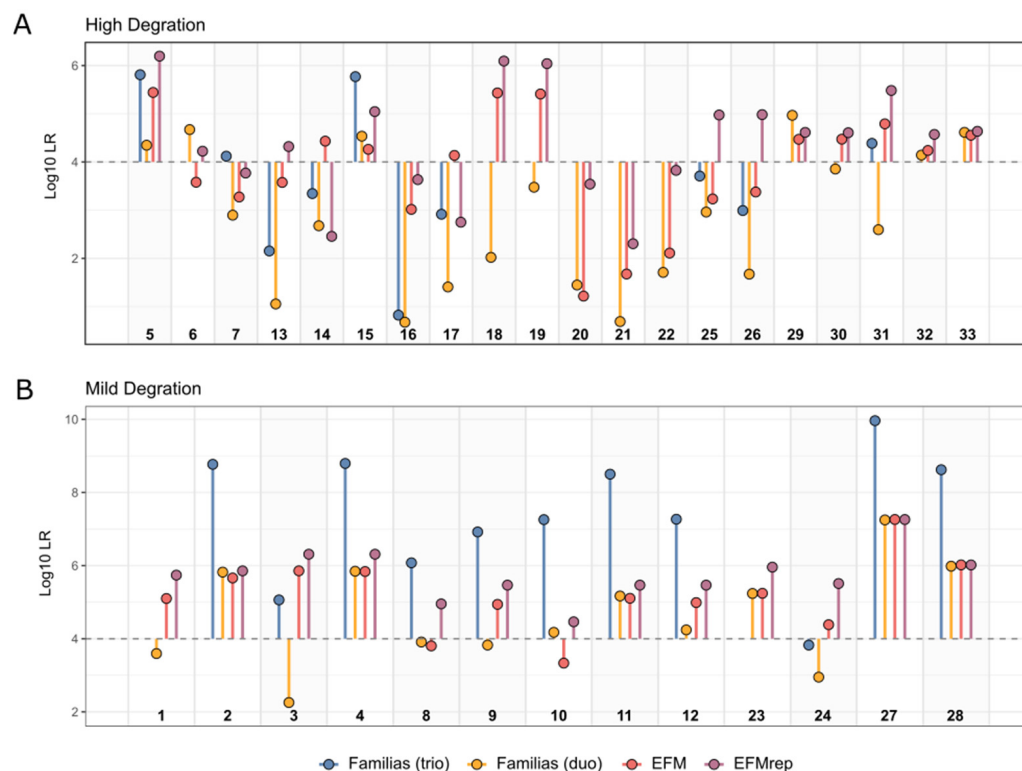


Figure 3. Lollipop plot of LR values computed with different software in HD (A) and MD samples (B). For the Familias software, LR values obtained taking into account maternal information when available (trio cases) are shown in blue. The dashed horizontal line indicates the threshold value of LR = 10,000. Sample ID numbers are shown below and correspond to those listed in Supplementary Table S1.

4. Discussion

The LT-DNA paternity cases selected for the study had all produced LR values > 1 , thus supporting the hypothesis of paternity rather than the alternative (unrelatedness) [36], when kinship software Familias and standard internal procedures, consisting of two replicate PCR experiments of the same STR kit, were applied. This approach relied on the combination of a well-established method to model for drop-out implemented in Familias [16,37] and the classic binary ‘biological’ solution strategy to LT-DNA samples, which deduces a consensus profile from replicate analyses based on all-or-none allele detection [1,33,38]. Reanalysis of STR data using PGS generated outputs that were always in agreement with Familias in producing LR values > 1 . Application of probabilistic genotyping to paternity analysis thus met the fundamental prerequisite for the internal validation of new software, that is, the consistency of results with those obtained on representative in-house generated data using standard laboratory-defined interpretative procedures [39].

Further steps of the software implementation process include the determination of robustness and limits in order to establish the types of profiles that are suitable for handling by the laboratory. The first factor to consider is whether the pedigrees being tested are standard trios or motherless cases (duos). In duo cases, uncertainty regarding the child’s maternal obligate alleles produces a loss of information with dual effect [40]. A general reduction in average LR values is observed, potentially leading to false-negative results, i.e., true fathers who do not reach the set LR threshold for inclusion [41]. Among the tested software programs, only Familias can take maternal information into account and, consequently, the higher LR values observed for Familias compared to EFM and EFMrep were expected, at least for cases in which degradation was mild and the effect of stochastic events was only marginally relevant.

The second effect of missing maternal information is the increased risk of false inclusions due to a reduction in the number of expected allelic inconsistencies between the child and an unrelated individual [42]. Thanks to the evolution of STR kits, which presently include more than 15 loci, it is now possible to overcome this pitfall of duo paternity testing, provided that the analysis is conducted on high-quality samples [43,44]. However, in LT-DNA samples, the number of STRs that can be genotyped effectively is often <15 due to fragmentation of template DNA. The problem is particularly evident when an extremely conservative approach ($2\times$) is adopted to establish consensus profiles in Familias, as in this study. Previously, it was shown that the $2\times$ method only detects an average of 50% of DNA profile alleles, which corresponds to the expected dropout of one-quarter of tested loci in LT-DNA conditions [33]. We have observed that, by applying probabilistic genotyping methods, the number of informative loci was significantly higher compared to Familias and was positively correlated with LR values. On average, the maximum number of informative STR loci and, consequently, significantly higher LRs were obtained with EFMrep, which leverages the combination of the results of PCR replicates obtained with different STR kits. The ESI and ESX STR kits employed in this research for EFMrep calculations, while targeting the same panel of loci, adopt distinct combinations of PCR primers [25]. Their combined use is particularly advantageous for the analysis of degraded or challenging samples, as it maximises the recovery of allelic information across loci and enables the confirmation of the results obtained with the alternate system [45].

In trio cases with highly degraded DNA, full quantitative analysis implemented in PGS compensated for the impossibility to take maternal information into account, producing LR values which were equivalent to those achieved with Familias, according to paired comparisons. However, in 80% of the cases, it was either EFM or EFMrep to maximise LR values.

Overall, irrespective of degradation status and pedigree type, in cases where Familias did not reach the nominal LR threshold of 10,000 for paternity inclusion according to national guidelines, this was possible by applying PGS. This threshold was set by policy based on high-quality DNA samples and standard LR calculation methods, and it represents a trade-off between the risks of false-positive and false-negative paternity test results [31,32]. It was observed that PGS could produce LR values up to four orders of magnitude larger than those calculated using Familias. It must be noted that the LR value produced by PGS and used for comparisons within the present study is of the MLE type. The determination of a more conservative LR through a sensitivity test, using the distribution of LRs based upon the errors of the parameters that were estimated to produce the MLE [23], is not feasible in EFM for calculations implying relatedness, and it is not implemented in the currently available version of EFMrep. This could potentially lead to an overestimation of evidence in favour of paternity. On the other hand, a risk of factitious underestimation of LR values when using PGS in paternity testing is linked to possible father–child mutations. The PGS investigated in the present study do not incorporate mutation rates, which range between 0.64% of STR SE33 and 0.01% of STR TH01 [34]. Consequently, any mismatch between the STR profiles of the LT-DNA sample and of the alleged father–child can only be interpreted as a PCR artefact. Consider, for example, a paternal LT-DNA sample displaying a heterozygous genotype that does not share any alleles with the alleged child. If paternity is hypothesised, with PGS used in the present study, the only possible explanation would be the drop-out of the shared allele together with the drop-in of a random allele. However, if the peak height of the allele assumed to have dropped in is significantly higher than the AT, the probability of this occurring, as calculated using the exponential drop-in function in EFM and EFMrep, is expected to be extremely low. This would result in an LR value that strongly rejects the proposition of paternity for the affected locus. If the paternal

LT-DNA sample is homozygous at a given STR locus with no shared allele with the child's genotype, one possible explanation under the proposition of paternity is that there was a drop-out of the locus accompanied by a drop-in of the observed allele. The effect on locus-specific LR results in this case will be the same as in the previous example. Another possible explanation would be drop-out in the paternal heterozygous genotype affecting the allele that is shared with the child. In this case, too, if the peak height of the paternal allele observed in the LT-DNA sample is well above the AT, the probability of drop-out of companion allele according to the continuous peak height model implemented in EFM and EFMrep is expected to be exceedingly low and the locus-specific LR will strongly support non-paternity. In any case, the model validation function implemented in EFM and EFMrep can be conveniently used to detect such events. It produces probability–probability plots that compare two empirical cumulative distributions of the observed peak heights with the theoretical underlying model. For reasonable peak height models, the probability values of STR loci will be arranged along a 45° straight line relationship, whereas single-STR loci affected by mutation will stand out as outliers.

We are aware that the PGS considered here (EFM and EFMrep) were not designed for the purpose of paternity testing. However, it has been suggested that statistical redundancy, i.e., applying a multi-software approach to challenging forensic cases, can be beneficial for the interpretation of the data and provide an additional diagnostic check [46,47]. The obtained results shed light on the potentialities and constraints of the statistical models underlying standard kinship software and PGS and contribute to the needed interpretational background required for the potential future use of PGS as a supplementary tool in paternity testing of LT-DNA samples [48].

Finally, it is important to note that the present study has limitations. Firstly, EFM and EFMrep are not the only PGS programs that can take the relatedness of DNA profile donors into account, so comparisons could be extended to other software with similar properties described in the literature [49,50]. Secondly, the analysis was conducted under simplified experimental conditions, i.e., only two PCR replicates were obtained using two STR kits with a limited number of loci (16). Given that a concrete balance must be reached in forensic investigations between test informativity and laboratory costs [51], it is known that by increasing the number of PCR replicates, it is possible to boost robustness and reliability [33,52]. While the use of multiplex PCR systems with additional STRs does not seem to provide a substantial gain in the number of detected alleles and informative loci in LT-DNA samples [53], the effect of combined results from different expanded STR kits has not been investigated so far. So, it could be interesting in the future to extend this study to include more PCR replicates and, for EFMrep, combinations of STR kits. The investigation of the performance of standard paternity software and PGS in more complex kinship scenarios, different from paternity, could be another intriguing development.

5. Conclusions

This preliminary study represents a first step for the internal validation of PGS in paternity investigations of LT-DNA samples and for the definition of an operational flow chart to identify the combination of calculation software most suitable for the starting analytical conditions in terms of the type of pedigree and the level of DNA degradation.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/genes17020187/s1>. Figure S1: Probability of drop-out as a function of peak height expressed in RFU for the PowerPlex ESI 17 Fast (A) and ESX 17 Fast (B) kit. Figure S2: Probability of drop-in as a function of peak height expressed in RFU for the PowerPlex ESI 17 Fast (A) and ESX 17 Fast (B) kit. Table S1: General characteristics of tested LT-samples: sample number, sample source, degradation status (MD: mildly degraded; HD: highly degraded), type of

pedigree (duo, trio), number of informative STRs and LR values obtained with the software Familias, EuroForMix, and EFMrep. For trio cases, Familias LRs calculated both considering and disregarding maternal information are given.

Author Contributions: Conceptualization, C.R.; formal analysis, S.A.; investigation, E.C. and A.R.; data curation, F.B. and M.P.; writing—original draft preparation, E.C., A.R., S.A. and C.R.; writing—review and editing, E.C., A.R. and C.R.; supervision, C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study design and research methods were approved by the University of Turin Research Ethics Review Committee (Authorization No. 0532888, 18 July 2025).

Informed Consent Statement: Informed consent for DNA testing and paternity calculations was obtained from living participants and authorised by the judicial authority for deceased individuals.

Data Availability Statement: General characteristics and LR values of anonymized LT-DNA cases used in this study can be found in the Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

5ST	5% Stochastic Threshold
AT	Analytical threshold
EFM	EuroForMix
EDNAP	European DNA Profiling Group
ENFSI	European Network of Forensic Science Institutes
ESI	PowerPlex® ESI 17 Fast kit
ESX	PowerPlex® ESX 17 Fast kit
ESS	European Standard Set
FST	Fixation Index (population genetics correction)
FFPE	Formalin-Fixed, Paraffin-Embedded
HD	Highly Degraded
Hd	Prosecution Hypothesis (paternity/relatedness)
Hp	Defence Hypothesis (alternative/unrelated)
ISFG	International Society for Forensic Genetics
LR	Likelihood Ratio
LT-DNA	Low-Template DNA
MD	Mildly Degraded
MLE	Maximum Likelihood Estimation
PGS	Probabilistic Genotyping Software
POI	Person of Interest
RFU	Relative Fluorescence Units
STR	Short Tandem Repeat

References

1. Gill, P.; Whitaker, J.; Flaxman, C.; Brown, N.; Buckleton, J. An Investigation of the Rigor of Interpretation Rules for STRs Derived from Less than 100 Pg of DNA. *Forensic Sci. Int.* **2000**, *112*, 17–40. [[CrossRef](#)]
2. Whitaker, J.P.; Cotton, E.A.; Gill, P. A Comparison of the Characteristics of Profiles Produced with the AMPFISTR® SGM Plus™ Multiplex System for Both Standard and Low Copy Number (LCN) STR DNA Analysis. *Forensic Sci. Int.* **2001**, *123*, 215–223. [[CrossRef](#)] [[PubMed](#)]
3. Gill, P.; Buckleton, J. A Universal Strategy to Interpret DNA Profiles That Does Not Require a Definition of *Low-Copy-Number*. *Forensic Sci. Int. Genet.* **2010**, *4*, 221–227. [[CrossRef](#)]

4. Balding, D.J.; Buckleton, J. Interpreting Low Template DNA Profiles. *Forensic Sci. Int. Genet.* **2009**, *4*, 1–10. [\[CrossRef\]](#)
5. Walsh, P.S.; Fildes, N.J.; Reynolds, R. Sequence Analysis and Characterization of Stutter Products at the Tetranucleotide Repeat Locus VWA. *Nucleic Acids Res.* **1996**, *24*, 2807–2812. [\[CrossRef\]](#)
6. Gill, P.; Brenner, C.H.; Buckleton, J.S.; Carracedo, A.; Krawczak, M.; Mayr, W.R.; Morling, N.; Prinz, M.; Schneider, P.M.; Weir, B.S.; et al. DNA Commission of the International Society of Forensic Genetics: Recommendations on the Interpretation of Mixtures. *Forensic Sci. Int.* **2006**, *160*, 90–101. [\[CrossRef\]](#)
7. Gill, P.; Gusmão, L.; Haned, H.; Mayr, W.R.; Morling, N.; Parson, W.; Prieto, L.; Prinz, M.; Schneider, H.; Schneider, P.M.; et al. 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.* **2012**, *6*, 679–688. [\[CrossRef\]](#)
8. Haned, H.; Gill, P.; Lohmueller, K.; Inman, K.; Rudin, N. Validation of Probabilistic Genotyping Software for Use in Forensic DNA Casework: Definitions and Illustrations. *Sci. Justice* **2016**, *56*, 104–108. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Coble, M.D.; Bright, J.-A. Probabilistic Genotyping Software: An Overview. *Forensic Sci. Int. Genet.* **2019**, *38*, 219–224. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Riem, A.; Chierito, E.; Aneli, S.; Bertolotto, F.; Robino, C. Exploring Different Software Options for Likelihood Ratio Calculations in Paternity Cases with Low Template DNA. *USC* **2025**, *1*, 483–489. [\[CrossRef\]](#)
11. Romero, R.L.; Juston, A.C.; Ballantyne, J.; Henry, B.E. The Applicability of Formalin-Fixed and Formalin Fixed Paraffin Embedded Tissues in Forensic DNA Analysis. *J. Forensic Sci.* **1997**, *42*, 708–714. [\[CrossRef\]](#)
12. Rucinski, C.; Malaver, A.L.; Yunis, E.J.; Yunis, J.J. Comparison of Two Methods for Isolating DNA from Human Skeletal Remains for STR Analysis. *J. Forensic Sci.* **2012**, *57*, 706–712. [\[CrossRef\]](#)
13. Emery, M.V.; Bolhofner, K.; Winingear, S.; Oldt, R.; Montes, M.; Kanthaswamy, S.; Buikstra, J.E.; Fulginiti, L.C.; Stone, A.C. Reconstructing Full and Partial STR Profiles from Severely Burned Human Remains Using Comparative Ancient and Forensic DNA Extraction Techniques. *Forensic Sci. Int. Genet.* **2020**, *46*, 102272. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Watherston, J.; McNevin, D.; Gahan, M.E.; Bruce, D.; Ward, J. Current and Emerging Tools for the Recovery of Genetic Information from Post Mortem Samples: New Directions for Disaster Victim Identification. *Forensic Sci. Int. Genet.* **2018**, *37*, 270–282. [\[CrossRef\]](#)
15. Egeland, T.; Mostad, P.F.; Mevåg, B.; Stenersen, M. Beyond Traditional Paternity and Identification Cases. Selecting the Most Probable Pedigree. *Forensic Sci. Int.* **2000**, *110*, 47–59. [\[CrossRef\]](#)
16. Dørum, G.; Kling, D.; Baeza-Richer, C.; García-Magariños, M.; Sæbø, S.; Desmyter, S.; Egeland, T. Models and Implementation for Relationship Problems with Dropout. *Int. J. Leg. Med.* **2015**, *129*, 411–423. [\[CrossRef\]](#)
17. Kling, D.; Tillmar, A.O.; Egeland, T. Familias 3-Extensions and New Functionality. *Forensic Sci. Int. Genet.* **2014**, *13*, 121–127. [\[CrossRef\]](#)
18. Green, P.J.; Mortera, J.; Prieto, L. Casework Applications of Probabilistic Genotyping Methods for DNA Mixtures That Allow Relationships between Contributors. *Forensic Sci. Int. Genet.* **2021**, *52*, 102482. [\[CrossRef\]](#)
19. Kruijver, M.; Taylor, D.; Bright, J.-A. Evaluating DNA Evidence Possibly Involving Multiple (Mixed) Samples, Common Donors and Related Contributors. *Forensic Sci. Int. Genet.* **2021**, *54*, 102532. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Gjertson, D.W.; Brenner, C.H.; Baur, M.P.; Carracedo, A.; Guidet, F.; Luque, J.A.; Lessig, R.; Mayr, W.R.; Pascali, V.L.; Prinz, M.; et al. ISFG: Recommendations on Biostatistics in Paternity Testing. *Forensic Sci. Int. Genet.* **2007**, *1*, 223–231. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Gill, P.; Benschop, C.; Buckleton, J.; Bleka, Ø.; Taylor, D. A Review of Probabilistic Genotyping Systems: EuroForMix, DNASTatistX and STRmixTM. *Genes* **2021**, *12*, 1559. [\[CrossRef\]](#)
22. Gómez, N.A.; Nieto, A.P.R.; Chacón, G.; Lizarazo, R.; Egeland, T. A Crime Case Involving a Mixture of Three Individuals Solved Using a Sample from the Daughter of the Person of Interest. *Int. J. Leg. Med.* **2025**, *139*, 1569–1574. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Bleka, Ø.; Storvik, G.; Gill, P. 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.* **2016**, *21*, 35–44. [\[CrossRef\]](#)
24. Bleka, Ø.; Prieto, L.; Gill, P. EFMrep: An Extension of EuroForMix for Improved Combination of STR DNA Mixture Profiles. *Forensic Sci. Int. Genet.* **2022**, *61*, 102771. [\[CrossRef\]](#)
25. Hill, C.R.; Duewer, D.L.; Kline, M.C.; Sprecher, C.J.; McLaren, R.S.; Rabbach, D.R.; Krenke, B.E.; Ensenberger, M.G.; Fulmer, P.M.; Storts, D.R.; et al. Concordance and Population Studies along with Stutter and Peak Height Ratio Analysis for the PowerPlex[®] ESX 17 and ESI 17 Systems. *Forensic Sci. Int. Genet.* **2011**, *5*, 269–275. [\[CrossRef\]](#)
26. Chierito, E.; Aneli, S.; Nocco, N.; Riem, A.; Onofri, M.; Carnevali, E.; Robino, C. Assessing DNA Degradation through Differential Amplification Efficiency of Total Human and Human Male DNA in a Forensic qPCR Assay. *Genes* **2024**, *15*, 622. [\[CrossRef\]](#)
27. Presciuttini, S.; Cerri, N.; Turrina, S.; Pennato, B.; Alù, M.; Asmundo, A.; Barbaro, A.; Boschi, I.; Buscemi, L.; Caenazzo, L.; et al. Validation of a Large Italian Database of 15 STR Loci. *Forensic Sci. Int.* **2006**, *156*, 266–268. [\[CrossRef\]](#) [\[PubMed\]](#)

28. Previderè, C.; Grignani, P.; Alessandrini, F.; Alù, M.; Biondo, R.; Boschi, I.; Caenazzo, L.; Carboni, I.; Carnevali, E.; De Stefano, F.; et al. The 2011 GeFI Collaborative Exercise. Concordance Study, Proficiency Testing and Italian Population Data on the New ENFSI/EDNAP Loci D1S1656, D2S441, D10S1248, D12S391, D22S1045. *Forensic Sci. Int. Genet.* **2013**, *7*, e15–e18. [\[CrossRef\]](#)
29. Welch, L.A.; Gill, P.; Phillips, C.; Ansell, R.; Morling, N.; Parson, W.; Palo, J.U.; Bastisch, I. European Network of Forensic Science Institutes (ENFSI): Evaluation of New Commercial STR Multiplexes That Include the European Standard Set (ESS) of Markers. *Forensic Sci. Int. Genet.* **2012**, *6*, 819–826. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Buckleton, J.; Curran, J.; Goudet, J.; Taylor, D.; Thiery, A.; Weir, B.S. Population-Specific FST Values for Forensic STR Markers: A Worldwide Survey. *Forensic Sci. Int. Genet.* **2016**, *23*, 91–100. [\[CrossRef\]](#)
31. Genetisti Forensi Italiani RACCOMANDAZIONI PER GLI ACCERTAMENTI PARENTALI. Available online: <https://www.gefi-isfg.org/temp/20112018100854.pdf> (accessed on 30 November 2025).
32. Società Italiana di Genetica Umana Analisi Genetiche Di Accertamento Parentale (Aggiornamento 2022). Available online: https://sigu.net/wp-content/uploads/2022/11/2023_05_02_AccertamentoParentale_2022_Rev03.pdf (accessed on 30 November 2025).
33. Benschop, C.C.G.; van der Beek, C.P.; Meiland, H.C.; van Gorp, A.G.M.; Westen, A.A.; Sijen, T. Low Template STR Typing: Effect of Replicate Number and Consensus Method on Genotyping Reliability and DNA Database Search Results. *Forensic Sci. Int. Genet.* **2011**, *5*, 316–328. [\[CrossRef\]](#)
34. Butler, J.M. *Advanced Topics in Forensic DNA Typing: Interpretation*; Academic Press: New York, NY, USA, 2015; ISBN 978-0-12-405213-0.
35. R: The R Project for Statistical Computing. Available online: <https://www.r-project.org/> (accessed on 26 November 2025).
36. European Network of Forensic Science Institutes ENFSI Guideline for Evaluative Reporting in Forensic Science. Available online: https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (accessed on 30 November 2025).
37. Buckleton, J.; Triggs, C. Dealing with Allelic Dropout When Reporting the Evidential Value in DNA Relatedness Analysis. *Forensic Sci. Int.* **2006**, *160*, 134–139. [\[CrossRef\]](#)
38. Kloosterman, A.D.; Kersbergen, P. Efficacy and Limits of Genotyping Low Copy Number (LCN) DNA Samples by Multiplex PCR of STR Loci. *J. Soc. Biol.* **2003**, *197*, 351–359. [\[CrossRef\]](#)
39. Coble, M.D.; Buckleton, J.; Butler, J.M.; Egeland, T.; Fimmers, R.; Gill, P.; Gusmão, L.; Guttman, B.; Krawczak, M.; Morling, N.; et al. DNA Commission of the International Society for Forensic Genetics: Recommendations on the Validation of Software Programs Performing Biostatistical Calculations for Forensic Genetics Applications. *Forensic Sci. Int. Genet.* **2016**, *25*, 191–197. [\[CrossRef\]](#)
40. Wenk, R.E.; Houtz, T.; Chifari, F.A. Maternal Typing and Test Sufficiency in Parentage Analyses. *Transfusion* **2006**, *46*, 199–203. [\[CrossRef\]](#)
41. Aguiar, V.R.C.; de Castro, A.M.; Pinto, L.M.; Ferreira, A.C.S.; Dos Santos, E.V.W.; Louro, I.D. Assessing False Paternity Risk in Simulated Motherless Cases from More than 20 000 Real Exclusion Trios. *Transfusion* **2021**, *61*, 678–681. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Lee, H.S.; Lee, J.W.; Han, G.R.; Hwang, J.J. Motherless Case in Paternity Testing. *Forensic Sci. Int.* **2000**, *114*, 57–65. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Elizabeth García-Aceves, M.; Gabriela Jasso-Razo, D.; Xitlally Díaz-Navarro, X.; Rangel-Villalobos, H. A Posteriori Parameters from Paternity Tests of a Mexican Laboratory with the Powerplex Fusion System. *Leg. Med.* **2023**, *64*, 102296. [\[CrossRef\]](#)
44. Poetsch, M.; Preusse-Prange, A.; Schwark, T.; von Wurmb-Schwark, N. The New Guidelines for Paternity Analysis in Germany—How Many STR Loci Are Necessary When Investigating Duo Cases? *Int. J. Leg. Med.* **2013**, *127*, 731–734. [\[CrossRef\]](#)
45. Pfeifer, C.M.; Klein-Unseld, R.; Klintschar, M.; Wiegand, P. Comparison of Different Interpretation Strategies for Low Template DNA Mixtures. *Forensic Sci. Int. Genet.* **2012**, *6*, 716–722. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Alladio, E.; Omedei, M.; Cisana, S.; D’Amico, G.; Caneparo, D.; Vincenti, M.; Garofano, P. DNA Mixtures Interpretation—A Proof-of-Concept Multi-Software Comparison Highlighting Different Probabilistic Methods’ Performances on Challenging Samples. *Forensic Sci. Int. Genet.* **2018**, *37*, 143–150. [\[CrossRef\]](#)
47. Riman, S.; Iyer, H.; Vallone, P.M. Examining Performance and Likelihood Ratios for Two Likelihood Ratio Systems Using the PROVEDIt Dataset. *PLoS ONE* **2021**, *16*, e0256714. [\[CrossRef\]](#)
48. Buckleton, J.; Susik, M.; Curran, J.M.; Cheng, K.; Taylor, D.; Bright, J.-A.; Kelly, H.; Wivell, R. A Diagnosis of the Primary Difference between EuroForMix and STRmix™. *J. Forensic Sci.* **2024**, *69*, 40–51. [\[CrossRef\]](#)
49. Buckleton, J.S.; Bright, J.-A.; Gittelson, S.; Moretti, T.R.; Onorato, A.J.; Bieber, F.R.; Budowle, B.; Taylor, D.A. The Probabilistic Genotyping Software STRmix: Utility and Evidence for Its Validity. *J. Forensic Sci.* **2019**, *64*, 393–405. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Adamowicz, M.S.; Rambo, T.N.; Clarke, J.L. Internal Validation of MaSTR™ Probabilistic Genotyping Software for the Interpretation of 2–5 Person Mixed DNA Profiles. *Genes* **2022**, *13*, 1429. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Hedell, R.; Hedman, J.; Mostad, P. Determining the Optimal Forensic DNA Analysis Procedure Following Investigation of Sample Quality. *Int. J. Leg. Med.* **2018**, *132*, 955–966. [\[CrossRef\]](#) [\[PubMed\]](#)

52. Steele, C.D.; Greenhalgh, M.; Balding, D.J. Verifying Likelihoods for Low Template DNA Profiles Using Multiple Replicates. *Forensic Sci. Int. Genet.* **2014**, *13*, 82–89. [[CrossRef](#)]
53. Duijs, F.; van de Merwe, L.; Sijen, T.; Benschop, C.C.G. Low-Template Methods Yield Limited Extra Information for PowerPlex® Fusion 6C Profiling. *Leg. Med.* **2018**, *33*, 62–65. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.