Comparison of genotyping and weight of evidence results when applying different genotyping strategies on samples from a DNA transfer experiment

Author(s)
Duijs, Francisca E. ; Meijers, Erin; Kokshoorn, Bas; Sijen, Titia

DOI
10.1007/s00414-022-02918-7

Publication date
2023

Document Version
Author accepted manuscript (AAM)

Published in
International Journal of Legal Medicine

Comparison of genotyping and weight of evidence results when applying different genotyping strategies on samples from a DNA transfer experiment

Francisca E. Duijs[1], Erin Meijers[1], Bas Kokshoorn[1,2] and T. Sijen[1,3]*

[1]Division of Biological Traces, Netherlands Forensic Institute, The Hague, The Netherlands

[2]Faculty of Technology, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands

[3]University of Amsterdam, Swammerdam Institute for Life Sciences, Amsterdam, The Netherlands

*Corresponding author. E-mail address: t.sijen@nfi.nl

Abstract

In this study we assessed to what extent data on the subject of TPPR (transfer, persistence, prevalence, recovery) that are obtained through an older STR typing kit can be used in an activity-level evaluation for a case profiled with a more modern STR kit. Newer kits are generally more sensitive and consist of more loci which could increase the evidential value at the source-level. On the other hand, the increased genotyping information may invoke a higher number of contributors in weight of evidence calculations, which could affect the evidential values as well. An activity scenario well explored in earlier studies [1,2] was redone using volunteers with known DNA profiles. DNA extracts were analyzed with three different approaches, namely using the optimal DNA input for 1) an older and 2) a newer STR typing system, and 3) using a standard, volume-based input combined with replicate PCR analysis with only the newer STR kit. The genotyping results were analyzed for various aspects such as percentage detected alleles and relative peak height contribution for background and the contributors known to be involved in the activity. Next, source-level LRs were calculated and the same trends were observed with regards to inclusionary and exclusionary LRs for persons who had or had not been into direct contact with the sampled areas. We subsequently assessed the impact on the outcome of activity-level evaluation in an exemplary case by applying the assigned probabilities to a Bayesian network. We infer that data of different STR kits can be combined in activity-level evaluations.

Keywords NGM - PowerPlex Fusion 6C – DNA - Activity scenario – TPPR - LR calculations – Bayesian Network

Introduction

Since STR typing was introduced to forensic DNA analysis, several generations of STR systems have been developed and implemented, always aiming for an increased number of loci and higher sensitivity. The increased sensitivity can affect not only the amount of genotyping information that is revealed for a donor, but also increase detection of unspecified background DNA [3,4] and enlarge the number of contributors that is required to explain the data. This can affect source-level likelihood ratios (LRs) [5].

When activity-level questions are raised during the criminal investigation, evaluations are performed that use amongst others data from transfer, persistence, prevalence, recovery (TPPR) studies that relate to the evaluated scenario [3,4]. During the last decade, multiple TPPR studies have been published as described in [6-9] and the data were generated through various STR typing systems. It has been proposed that collecting these data in a searchable DNA transfer database [7] would increase the data available to reporting officers for activity-level case evaluations. It is unknown whether the STR typing system that was used has an effect on TPPR rates inferred from such data. Differences are seen when different laboratories examine similar samples [10-12] but it is not fully established what the specific effect of different STR systems is.

In this study we compare the results obtained by two different STR typing systems in one specific activity scenario, namely dragging a person by the trouser ankles [1,2]. The systems are the Next Generation Multiplex kit (NGM, Life Technologies) having 15 autosomal STRs and the PowerPlex Fusion 6C kit (PPF6C, Promega) carrying 23 autosomal STRs. For the PPF6C system, we used two approaches: an optimal PCR input of 500 to 1000 pg for which regularly the DNA extract needed to be concentrated due to low DNA concentrations and a standard PCR input guided by the maximum volume of DNA extract that can be accommodated in the PCR. Since most often the standard input represents a lower amount of DNA, three PCR replicates were performed. The dragging experiments were executed by volunteers whose DNA profiles were known. Therefore, various aspects could be determined, such as percentages of detected alleles and proportions of peak heights for the known contributors and information on unspecified background alleles. In addition, source-level LRs were calculated by DNAStatistX [13] and compared between the two STR typing systems and two PPF6C genotyping approaches. An exemplary Bayesian network was constructed and the impact of the analysis method on probability assignment was assessed.

Materials and Methods

Experimental design and sampling

For this study eight volunteers, four males and four females, signed informed consent and were grouped in couples of one male and one female with lowest similarity in their genotypes (a maximum of 11 shared alleles when considering the NGM loci and 16 when considering the PPF6C loci) to maximally discern both donors. All volunteers were asked to wear jeans during this experiment. Although information on for example whether jeans were previously worn or

freshly washed was collected in a questionnaire, this information was not considered due to the large variation. Couples dragged each other on a carpeted flooring while one volunteer (denoted 'victim', 'V') was sitting in a chair with metal sled base and the other volunteer (denoted 'grabber', 'G') pulled the victim with his/her bare unwashed hands by both ankles for one minute over a carpeted flooring without any further instructions alike in [1,2]. Directly after dragging, samples were obtained using two tape lifts [14] per sampling area: one for the front and another for the back side of the jeans of the victim at the ankle area [1]. Control samples were taken from the jeans at each knee area also using two tape lifts at each area [2]. It was confirmed by Go-Pro filming that the grabber did not directly contact these knee areas. Samples from the left and the right leg were kept separate for a direct comparison. The roles of victim and grabber were reversed and a second set of samples was generated. Two weeks later the full experiment was repeated, resulting in four dragging events per couple. This way, a total of 64 samples were collected for the four couples.

DNA extraction, quantification and STR profiling

The two collected tape lifts per sampled area were taken together and processed in one extraction using the QIAamp DNA Mini kit (QIAGEN) according to manufacturer's instructions with one adjustment in that a total elution volume of 140 µL was used. DNA quantification using 2 µL of the DNA extract was performed using the *Alu* assay [15,16] on an ABI 7900HT real-time PCR system (Applied Biosystems™ (AB)) according to the manufacturer's protocol [16] with minor adjustments as described in [17].

Prior to STR profiling, the DNA extract was split into two portions; the first portion of approximately 28 µL was used to prepare DNA profiles according to standard procedures; the second portion of 110 µL was concentrated to 15 µL using ethanol precipitation as described in [17] to facilitate an optimal DNA input in the PCR.

The ethanol-precipitated DNA extract was split and equally divided for one NGM PCR and one PPF6C PCR taking care that not more than 1000 pg DNA (calculated from the concentration determined prior to ethanol precipitation) was added to the PCR. The optimal PCR input had been determined to be 500-1000 pg during in house validations of the PCR systems. Inputs were between 461 and 994 pg. The not-precipitated part of the DNA extract was used to generate three PPF6C PCR replicates each having maximal 7.5 µL input while taking care that the DNA input did not exceed 1000 pg (inputs were between 60 and 924 pg).

STR profiling with the PPF6C kit was performed using 29 amplification cycles according to the manufacturer's protocol except that a 12.5 µL PCR volume was used which is validated as default within our laboratory. STR profiling with the AmpFℓSTR® NGM™ PCR Amplification Kit (NGM, Life Technologies, LT) was performed using 29 amplification cycles and a standard 25 µL PCR volume. PCR products were separated using a 3500xL Genetic Analyzer (LT) with POP-4 (LT) separation matrix while following the recommended injection settings (1.2kV 24sec) and using 1 µL PCR product, 9.6 µL HiDi formamide (AB) and 0.4 µL internal size standard (for PPF6C WEN Internal Lane Standard 500 (Promega) and for NGM LIZ500 (LT)). For each STR system 1 µL of appropriate ladder was used: PowerPlex® Fusion 6C Allelic Ladder Mix (Promega) or the NGM Allelic ladder mix (LT).

For profile analysis GeneMarker HiD version 2.9.4 (SoftGenetics) was used with analysis settings described in [18]. Profile analysis parameters were established by an in-house validation. The dye-specific detection thresholds for PPF6C profiles were: FL=95, JOE=140, TMR=85, CXR=135 and TOM=95 relative fluorescent units (RFU) with a minimal heterozygote imbalance percentage of 3%. Note that the minimal heterozygote imbalance percentage is an extra threshold that uses a percentage of the highest peak in a marker to define the minimum peak threshold. NGM dye-specific detection thresholds were: 6-FAM=110, VIC=65, NED=120 and TAZ=120 RFU and a minimal heterozygote imbalance percentage of 1% was applied. For both PCR kits in-house determined locus-specific stutter ratios were applied.

Data analysis and LR calculations

For all donors PPF6C reference profiles were generated that were used to mark alleles as grabber's, victim's, shared or background. During data analysis, homozygote alleles were counted as two alleles. Percentage of detected alleles and percentage RFU were calculated for alleles not shared between victim and grabber. Y-chromosomal markers and AMEL (amelogenin) were excluded from comparison, resulting in 23 autosomal loci that were considered for PPF6C and 15 loci for NGM. For the PPF6C standard input the average of the three replicates was determined.

DNAStatistX v1.2.0 [13], based on EuroForMix [19] was used for LR calculations as described in [20] with a theta correction (F_{st}) of 0.01. For each analysis the number of contributors used under the propositions was determined using the maximum allele count (MAC). Three types of LR calculations were performed for each sample: 1) the victim is taken as person of interest (POI); 2) the grabber is taken as POI and 3) the grabber is regarded POI while conditioned on the victim as known contributor under both propositions. Under Hd, the POI is replaced by an unknown unrelated individual; all other parameters were kept as under Hp.

Results

Genotyping results

Our volunteers produced 16 dragging events, from which 64 samples were taken: 32 dragging and 32 control samples (ankle and knee area, respectively). As described in section 'DNA extraction, quantification and STR profiling', all 64 samples were genotyped with three approaches: 1) NGM optimal input; 2) PPF6C optimal input and 3) PPF6C standard input with three PCR replicates. Various features regarding the representation of victim, grabber and unspecified background in the profiles were analyzed per sample for each genotyping approach. The detailed results are presented in Suppl. Table 1; the average and range of results for the samples are given in Table 1.

All samples were mixtures (at least two contributors (Table 1 and Suppl. Table 1)) and background was detected in all samples (Table 1 and Suppl. Table 1). In the analysis we

assumed that the alleles that match grabber or victim derived from these contributors except for shared alleles that were considered as a separate category to prevent over-calculating values for the contributor (grabber or victim) that has a much lower contribution. The grabber was less present in the control samples when regarding the percentage of detected grabber alleles, the percentage of complete grabber profiles and the peak height ratios of grabber to background and grabber to victim. This was expected and supports earlier expectations [1,2] as the grabber did not have direct contact with this area of the trousers. The PPF6C data showed the trend that more alleles were detected both for grabber and victim when an optimal instead of a standard input was used. The detailed analysis in Suppl. Table 1 and Suppl. Fig. 1 indicated that the left and right leg samplings were very similar in the percentage of detected alleles for victim and grabber. Because of the limited number of dragging events per volunteer in the same role (namely two), we did not assess the effect of different shedder statuses for our volunteers.

Table 1 The average and range of results for the 32 dragging and 32 control samples for the three approaches: 1) NGM optimal input; 2) PPF6C optimal input and 3) PPF6C standard input with three PCR replicates. PH = peak height

	Dragging samples (*n*=32)			Control samples (*n*=32)		
	NGM Opt. input	PPF6C Opt. input	PPF6C Std. input	NGM Opt. input	PPF6C Opt. input	PPF6C Std. input
Avg. # of contributors	4 ±0.9	4 ±1.1	3 ±0.7	3 ±1.0	4 ±1.2	3 ±0.9
Range # of contributors	2-5	2-6	2-5	2-5	2-7	2-6
Avg. % detected alleles victim	97% ±10%	98% ±7%	86% ±20%	100% ±0%	100% ±0%	95% ±12%
Range % detected alleles victim	53% - 100%	68% - 100%	21% - 100%	100%	100%	56% - 100%
Complete victim profiles	88%	84%	44%	100%	100%	76%
Avg. % detected alleles grabber	95% ±17%	95% ±16%	80% ±29%	50% ±29%	47% ±27%	37% ±23%
Range % detected alleles grabber	12% - 100%	11% - 100%	3% - 100%	0% - 100%	7% - 97%	3% - 97%
Complete grabber profiles	78%	81%	52%	3%	0%	0%
Avg. PH ratio grabber/victim	2.3 ±0.0	2.3 ±2.8	2.2 ±2.4	0.2 ±0.1	0.2 ±0.1	0.4 ±0.4
Range PH ratio grabber/victim	0.0 - 9.2	0.1 - 8.8	0.1 - 9.1	0.0 - 0.6	0.0 - 0.6	0.0 - 1.6
Avg. # background alleles/locus	1.6 ±0.7	1.3 ±0.7	1.0 ±0.5	1.6 ±1.0	1.4 ±0.9	1.0 ±0.7
Range # background alleles/locus	0-3	0-2	0-3	0-4	0-2	0-3
Avg. % background RFU	8% ±6%	7% ±6%	8% ±6%	12% ±10%	12% ±9%	13% ±14%
Range % background RFU	1% - 23%	0% - 25%	1% - 28%	0% - 34%	0% - 32%	0% - 48%
Avg. % victim RFU	34% ±21%	35% ±21%	37% ±25%	58% ±12%	60% ±12%	61% ±17%
Range. % victim RFU	4% - 65%	4% - 64%	2% - 78%	32% - 80%	34% - 80%	19% - 80%
Avg. % grabber RFU	33% ±21%	33% ±21%	36% ±24%	6% ±6%	6% ±5%	9% ±8%
Range. % grabber RFU	0% - 72%	0% - 70%	0% - 76%	0% - 20%	0% - 19%	0% - 29%
Avg. PH ratio victim/background	7 ±5.4	6 ±5.4	4 ±4.0	11 ±7.8	8 ±5.0	6 ±5.9
Range PH ratio victim/background	1 - 20	1 - 21	1 - 21	2 - 27	1 - 18	1 - 29
Avg. PH ratio grabber/background	7 ±6.1	5 ±4.1	4 ±3.3	1 ±0.6	1 ±0.4	1 ±0.4
Range PH ratio grabber/background	1 - 26	1 - 15	1 - 13	1 - 4	1 - 2	0 - 2

Comparing NGM and PPF6C data with exactly the same DNA input

For each sample, an NGM and PPF6C amplification was performed with an optimal PCR input which was between 461 and 994 pg with only one sample that did not reach the minimum of 500 pg input (Suppl. Table 1). The portion in each profile representing victim or grabber alleles, their shared alleles and alleles of which the source could not be attributed (these are denoted

background alleles), were determined as their percentage of the total RFUs. In Fig. 1 we plotted the PPF6C data against the NGM data; Fig. 1A shows the dragging samples (ankle area) and Fig. 1B the control samples (knee area). Since we regarded the percentages of the total RFUs, the difference in the number of autosomal loci in the two types of STR systems is accounted for.

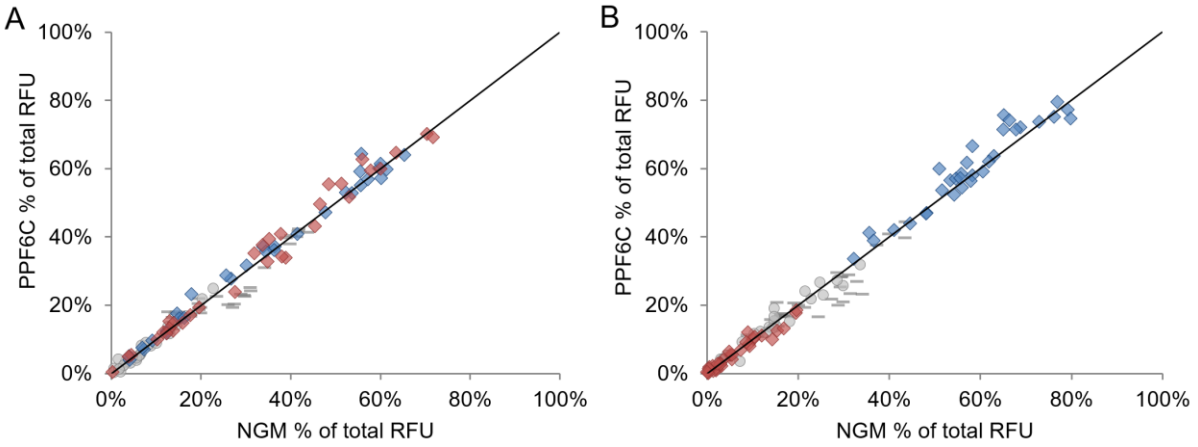

Fig. 1 STR kit comparison of the percentage of total RFU contributed by the grabber (red diamond), victim (blue diamond), background (gray bullet) and shared alleles (gray hyphen) in dragging samples (**A**) (ankle area, *n*=32) and control samples (**B**) (knee area, *n*=32)

All data points in both plots (Fig. 1) centered around the x=y line, indicating that similar results were obtained regardless of the STR typing system that was used. Also when the percentage of detected alleles was regarded, similar results were obtained as the data points for the two STR typing systems centered around the x=y line (results not shown). Although the newer generation STR systems are suggested to have increased sensitivity [21-23], we did not see an increased detection of low-level background in these dragging experiments: for the percentage background allele RFUs, the ratio PPF6C to NGM was on average 1.02 considering all samples. Apparently, PPF6C and NGM are of similar sensitivity with these samples. However, 17 of the 64 samples had a higher number of contributors with PPF6C than with NGM, while only nine samples showed a lower number for PPF6C. Per locus analysis showed that this is probably due to the presence of the highly discriminatory marker SE33 in the PPF6C system as the average MAC calculated over all samples was highest for this locus (Suppl. Table 2).

In the third genotyping approach, the input in the PCR was driven by the maximum volume of extract that could be given to the reaction (7.5 µL) while preventing overamplification by considering a maximum input of 1000 pg (Suppl. Table 1). Consequently, the standard DNA input was not a fixed ratio of the optimal DNA input: actually, of the 64 samples, 52 had a lower DNA input down to 13%, four samples had the same input and eight samples a higher input (up to 167%) compared to the optimal input. A direct comparison alike Fig. 1 has limited use therefore (but all data are presented in Suppl. Table 1) and we proceeded to source-level LR calculations.

Source-level LRs

In activity-level evaluations, the probability of recovering DNA of a POI is often based on the presence/absence of DNA. This categorization can for instance be based on above/below a certain amount of DNA [24] or inclusionary/exclusionary source-level LRs. The latter has the advantage that more information such as allele frequencies and peak heights are considered. Therefore, we calculated LRs for the three genotyping approaches applied to the 64 samples. We used either grabber or victim as POI; when in a case the trousers were found near or worn by the victim a suspect grabber is POI; when the trousers were found at the house of the grabber, an LR with the victim as POI may be informative.

LRs were calculated with DNAStatistX [13]. Three types of LR calculations were performed for each sample: 1) the victim is taken as POI; 2) the grabber is taken as POI and 3) the grabber is regarded POI while conditioned on the victim as known contributor under both propositions. In casework, the third proposition is preferred when conditioning on the victim is possible from the context of the case. Unlike for the analyses presented in Fig. 1, having 15 or 23 autosomal STR loci can have an effect on the LR, just as the triplicate PCR analysis with the standard input.

LRs were sorted in six categories, namely <1, 1-10; 10-10^4; 10^4-10^6; 10^6-10^9 and >10^9, in which LRs ≥10^4 are regarded inclusionary (10^6 and 10^9 are thresholds used in our casework [25]). Results are presented in Fig. 2 and Suppl. Table 3. In line with expectations [1,2] we observed the following:

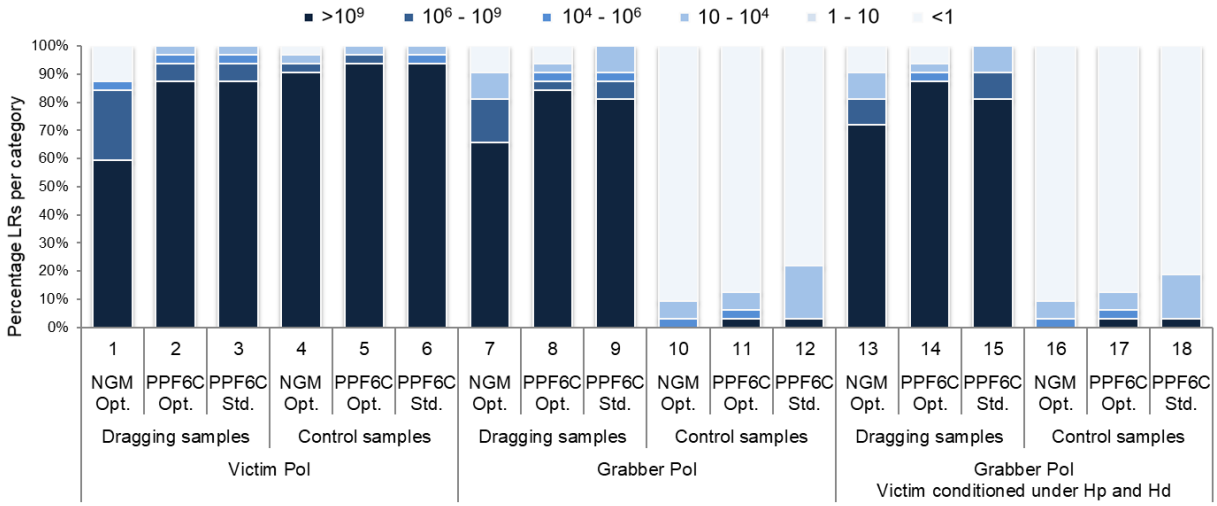

Fig. 2 LR distribution for all dragging and control samples. Opt. for the optimal DNA input, Std. for the standard DNA input. Per bar *n*=32 calculations

1) The control samples showed high LRs when the victim is used as POI (bars 4-6) and low LRs when the grabber is taken as POI (bars 10-12 and 16-18). This is understandable as the trousers were owned by the victim and the control areas have not been in direct contact with the grabber;

2) There is a great difference in the most frequent LR category for the dragging and control samples (category >10^9 and <1 respectively) when considering the grabber as POI (bars 7-9, 13-15 compared to bars 10-12, 16-18);

3) There is a higher frequency for LRs >10^9 when profiling using an optimal DNA input with the 23 loci PPF6C system than with the 15 loci NGM system (bars 1, 7, 13 compared to bars 2, 8, 14);

4) Conditioning on a donor, has a limited effect on the LRs (bars 7-9 compared to bars 13-15).

Next, we regarded the amount of POI DNA that was amplified. For this, we used the mixture contribution for a POI that is inferred in DNAStatistX [13] and multiplied it by the amount of DNA used in the PCR. Note that the mixture contribution estimates are not exactly the same for the NGM and PPF6C analyses which can be explained by stochastic effects during PCR amplification. Besides, the number of autosomal mismatches in the evidentiary trace profile was determined, being the number of reference alleles not present in the profiles from the experiments (note that volunteers with known profiles were used). The results are presented in Fig. 3. Details are given in Suppl. Table 4.

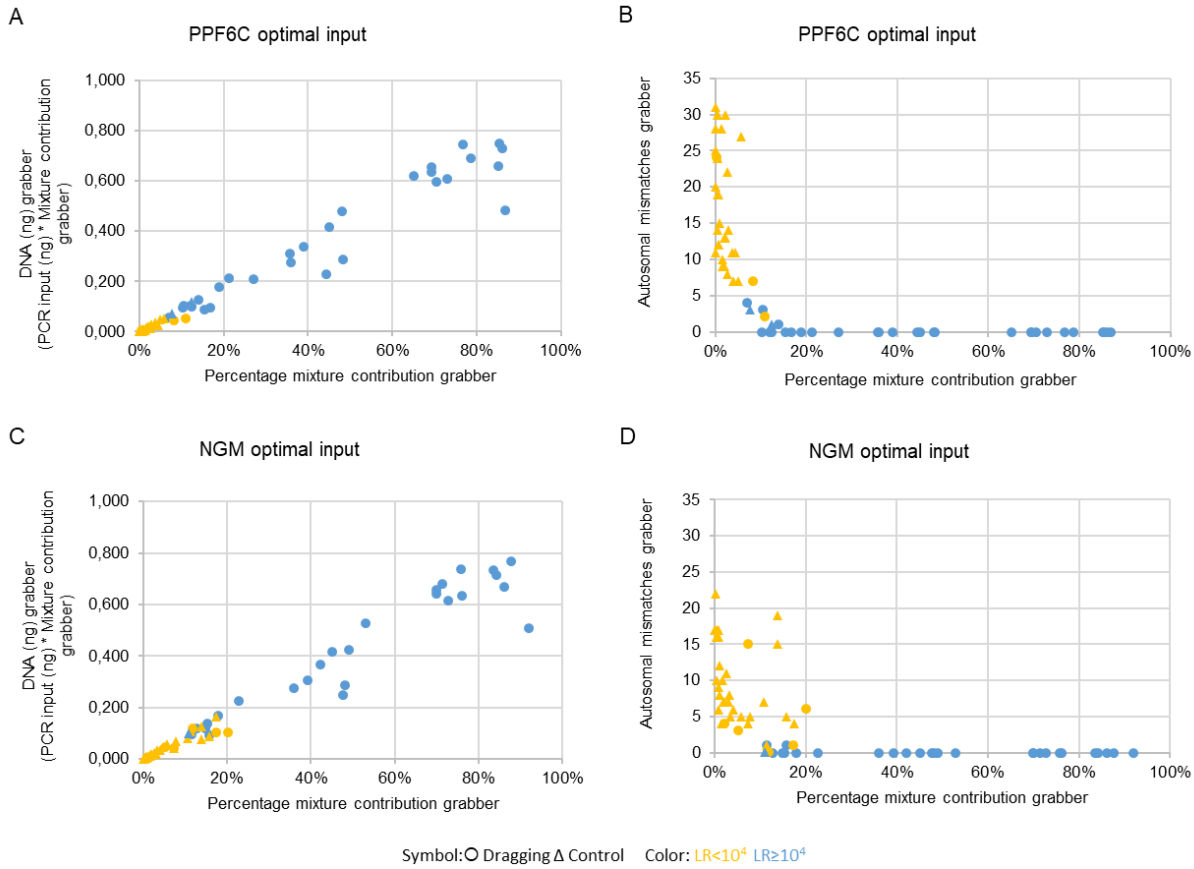

Fig. 3 Percentage mixture contribution of the grabber (POI in all calculations) plotted against the amount of DNA of the grabber in the PCR (**A, C**; PPF6C and NGM respectively) or the number of mismatches (considering only autosomal alleles) (**B, D**; PPF6C and NGM respectively). Circles indicate dragging samples; triangles control samples. LR results ≥10^4 are in blue; LR results <10^4 in yellow

Table 2 summarizes the information of Fig. 3. Although there is a clear trend that an LR ≥10^4 is more frequently achieved with less mismatches, more DNA for the POI or a higher mixture proportion, there are no clear-cut thresholds as there is overlap in the ranges observed for these features between the LR categories ≥10^4 and <10^4. The amount of DNA for the grabber estimated to be present in the amplification (based on mixture proportion and DNA input) appears more informative than the number of mismatches or the mixture proportion,

with a difference for PPF6C and NGM, which is explained from the loci in PPF6C. PPF6C carries 23 instead of 15 autosomal STRs and more high discriminatory loci (9 compared to 6 of the autosomal STRs have a polymorphism information content (PIC) value >0.800 in the European population [26]). Accordingly, slightly more samples had an LR $<10^4$ with NGM (for dragging samples 19% for NGM and 9% for PPF6C; with control samples 97% for NGM and 94% for PPF6C) (Table 2).

Table 2 Details for samples resulting in LRs $<10^4$ or $\geq 10^4$ for PPF6C and NGM using optimal input

Sample type	Profiling approach	LR	No. (%) samples	No. mismatches	No. with 0 mismatches	Pg grabber	Mix % grabber
Dragging	PPF6C Opt.	$<10^4$	3 (9%)	2-24	0	3-51	0.5-11.1%
		$\geq 10^4$	29 (91%)	0-4	26	58-749	7.0-86.9%
	NGM Opt.	$<10^4$	6 (19%)	0-15	1	10-117	2.2-20.2%
		$\geq 10^4$	26 (81%)	0-1	24	90-769	11.5-91.9%
Control	PPF6C Opt.	$<10^4$	30 (94%)	7-31	0	0-54	0.0-5.8%
		$\geq 10^4$	2 (6%)	1-3	0	70-117	7.8-12.5%
	NGM Opt.	$<10^4$	31 (97%)	1-22	0	1-163	0.2-17.5%
		$\geq 10^4$	1 (3%)	0	1	99	11%

Data used in an exemplary Bayesian Network

A case scenario related to the experiments in this study is the finding of the body of a victim (V) after a small social gathering. The state of the victim's clothing indicates dragging of their body at the ankles. After investigations, a defendant (D) is charged but this person claims to have had only social interaction with the victim during the gathering and never having touched the victim. The trouser ankles are sampled and DNA analyzed and this is the only result considered in the case. The source of the DNA is not disputed; the issue is how/when the DNA was transferred.

Fig. 4 shows an exemplary Bayesian network that is constructed using HUGIN Expert (v9.1) and the template in [24]. The propositions node [node 1] contains the propositions:
H1. The defendant dragged the victim by the ankles
H2. Somebody other than the defendant dragged the victim by the ankles
This node is the parent to the three activities of interest: social interaction between defendant and victim [node 2], dragging of the victim by the defendant [node 3], and the dragging of the victim by an alternative offender (AO) [node 4]. Each of these three activities informs a node that holds conditional probabilities on DNA TPR (transfer, persistence, recovery) [nodes 5 to 7]. These three nodes each contain two states; yes and no. This means that we consider the probability of recovering DNA or not, without specifying the amounts of DNA recovered. The probability of recovering background DNA of one or more unknown individuals from the clothing of the victim is given by [node 8], again with two states: yes and no. [Node 9] and [node 10] are summary nodes, stating whether or not DNA of the defendant or one or more unknowns (Us) is present in the sample. These two nodes subsequently inform the results node [node 11], which holds four states: 1) DNA of D only; 2) DNA of D and one or more unknowns; 3) DNA of one or more unknowns only or 4) Other result (no profile, or not interpretable).

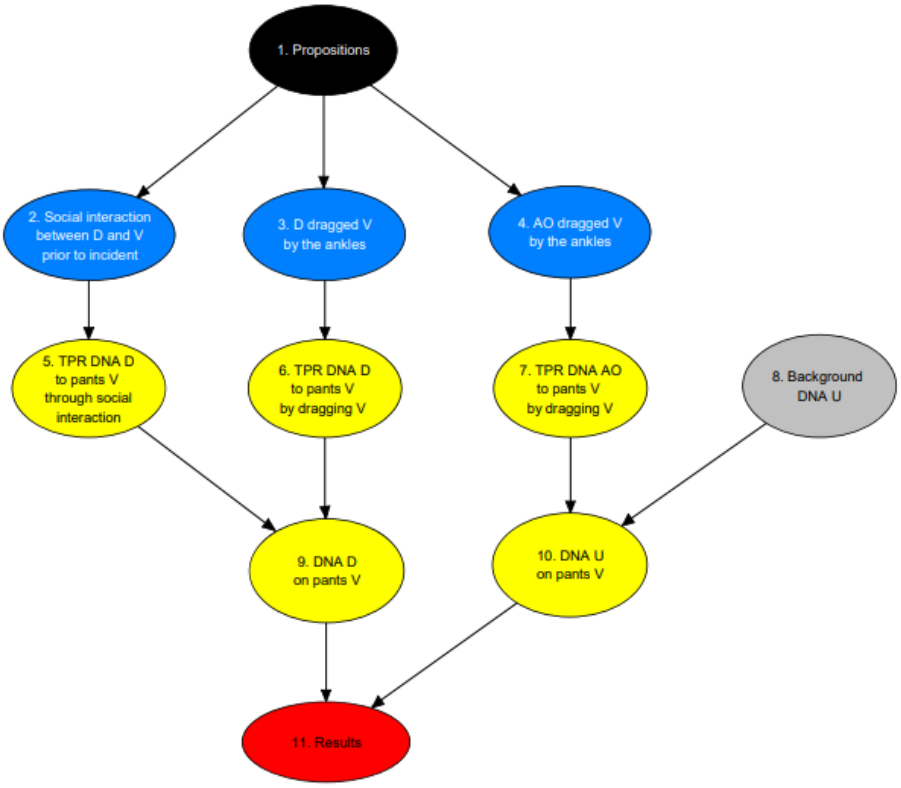

Fig. 4 Bayesian network for a case involving a dragged victim. Nodes are numbered and colored: black for propositions, blue for activities of interest, yellow for TPR probability (nodes 5-7) and summary (nodes 9 and 10), grey for probability of background DNA and red for results

Our experiments can be used to inform the probabilities assigned to nodes 5 to 7 relating to the presence or absence of DNA. For this exemplary Bayesian network, we assumed the experimental design aligns perfectly with the case circumstances and we translated the observed counts directly to probabilities with the assumption of a uniform prior count distribution as in [27,28]. We selected an LR of $\geq 10^4$ as threshold to categorize 'DNA regarded present' (LRs $<10^4$ are categorized as 'DNA regarded absent'). This threshold, while arbitrary, follows recent studies that were based on casework routines [29-31]. Node 5 (DNA prevalent after social interaction) is informed by the samples from the knees; nodes 6 and 7 (TPR through dragging) are based on the samples from the ankles. Node 8 contains (background DNA of one or more unknowns) also requires an arbitrary threshold to extract from the experimental data when background DNA is present. We define background DNA as the presence of more than 6 alleles that are not attributed to either victim or grabber (6 or fewer alleles are thus considered drop-ins such as elevated stutters). When the LRs for the grabber were below one, the alleles matching the grabber's reference profile were no longer considered grabbers' alleles and added to potential background. Using these thresholds we obtained counts and probabilities for the categories as shown in Table 3.

Table 4 shows the resulting LRs for the four result categories. Minor variations between the observed counts (Table 3) for the various approaches result in slightly different LRs at the activity-level (Table 4) that are in the same order of magnitude.

Table 3 Numbers of observations and between brackets the probabilities (using uniform prior count distribution) for the defined categories of Bayesian network (Fig. 4) nodes 5 to 8

Threshold	Node	Profiling approach	Yes	No
DNA regarded present (LR $\geq 10^4$)	Node 5	NGM Opt.	1 (0.06)	31 (0.94)
		PPF6C Opt.	2 (0.09)	30 (0.91)
		PPF6C Std.	1 (0.06)	32 (0.94)
	Nodes 6 and 7	NGM Opt.	26 (0.79)	6 (0.21)
		PPF6C Opt.	29 (0.88)	3 (0.12)
		PPF6C Std.	29 (0.88)	3 (0.12)
Background regarded present (>6 alleles)	Node 8	NGM Opt.	28 (0.85)	4 (0.15)
		PPF6C Opt.	31 (0.94)	1 (0.06)
		PPF6C Std.	32 (0.97)	0 (0.03)

Table 4 Activity-level LRs calculated for the four potential results in the exemplary Bayesian network (Fig. 4) for the three genotyping approaches

Activity-level LR	Profiling approach	DNA of D only	DNA of D and one or more unknowns	DNA of one or more unknowns only	Other result
	NGM Opt.	126	23	0.2 (5)	1
	PPF6C Opt.	164	19	0.11 (9)	1
	PPF6C Std.	245	29	0.11 (9)	1

When one has data for a node based on an older STR typing system while analyzing the case sample with a newer STR typing system, several approaches are possible:

Approach 1: One could only consider the loci present in both kits, which is clearly a loss of profiling information (eight PPF6C loci would not be regarded when data based on NGM would be used, including highly discriminatory loci such as SE33). **Approach 2**: One could make a rough estimate what the LRs for the samples typed by the older system would be when they would be analyzed with the newer system by multiplying the log10LRs for the older system by the proportion of loci in the newer system (PPF6C has 23/15≈1.5 times more loci). This estimation ignores the specifics of loci. Note that when we calculated the PPF6C/NGM ratio for the log10LRs, the average over all samples in our study was 1.48 (results not shown). **Approach 3**: One could use a higher LR threshold to consider an LR inclusionary when using more loci (for instance use 10^4 as inclusionary for NGM and 10^6 for PPF6C), but then one loses the advantage that with newer STR typing systems profiles may be less complete (for instance from DNA degradation) and still have high source-level LRs. **Approach 4**: One could use the counts observed for the older STR system even though the sample is typed with a newer STR system (as in Table 3).

In Table 4 we show that approach 4 has a limited effect on the activity-level LRs, but we tested whether approach 2 would perform better. Since PPF6C has ~1.5 times more loci, we adjusted the LRs for NGM by multiplying the log10LRs by 1.5 and counted the number of samples with an LR $\geq 10^4$. We observed no change for the counts for nodes 5, 6 and 7. This is of course dependent on the data set and the arbitrary threshold that was chosen. Also for deciding on the presence of background DNA, we adjusted the threshold to the lesser number of loci in the NGM system by regarding background present when more than four alleles not explained by victim and grabber were detected (four because this is 1.5 times fewer alleles than six). Now the counts for NGM for node 8 change: 28 times ‘no background’ (Table 3) becomes

30 times which means that the probabilities change from 0.85 to 0.91. As a consequence, the activity-level LR for NGM changes for the result 'DNA of D and one or more unknowns' from 23 to 25, which is a marginal difference.

Discussion

In an activity-level evaluation, several aspects may be questioned e.g. whether the victim was dragged, on what bodily location the victim was dragged, whether the grabber's DNA was transferred from dragging or a social/inoffensive interaction, whether bare or gloved hands were used while dragging. The experiment performed in this study will not inform on these questions; but it could contribute when assessing dragging by bare hands as the contact area (ankle region) and a control area (knee region) were analyzed as demonstrated by the exemplary Bayesian network.

When calculating activity-level LRs using a Bayesian network, one relies strongly on decisions where to set thresholds to categorize data. Even though these decisions are fed by expertise from case work and experimental research, these thresholds are per definition arbitrary. We have not extensively tested the effects of different thresholds but when we made adjustments to account for an increased number of loci with a newer system, only very limited effects were found. This is mainly because the source-level LRs for the grabber POI showed a very different distribution for the ankle and knee area that translate to the counts in nodes 6/7 and node 5 respectively. Furthermore, most of the samples carried background, which is expected on worn clothing, giving high counts for 'regarded present' in node 8.

Conclusion

We applied an older and a newer STR typing system on exactly the same DNA extract obtained from a dragging experiment and analyzed the profiling results, source-level LRs and activity-level LRs. For the newer STR system, two profiling approaches were performed: a single PCR with an optimal DNA input and three replicate PCRs with a standard DNA input. The newer STR system carries 23 instead of 15 autosomal loci. We find that with a profiling approach with more loci, more source-level LRs reside in the highest category ($>10^9$). Besides, the percentage samples with an inclusionary LR ($\geq 10^4$) was slightly higher and the activity-level LR for the result category 'DNA of defender only' was slightly higher. However, the results are in the same order of magnitude. This implies that data generated in TPPR studies with older STR systems or other approaches can be combined in activity-level evaluations.

Acknowledgements We would like to thank all volunteers for their participation. Margreet van den Berge, Corina Benschop and Klaas Slooten are thanked for helpful discussions and critically reading the manuscript. Jerry Hoogenboom is thanked for programming automized scripts to perform for LR calculations.

Author contribution Francisca Duijs: Investigation, Formal analysis, Visualization, Writing- Original draft preparation, Supervision. Erin Meijers: Investigation. Bas Kokshoorn: Writing- Reviewing and Editing. Titia Sijen: Conceptualization, Writing- Reviewing and Editing.

Declarations

Ethics approval Not applicable.

Informed consent Volunteers signed informed consent.

Data availability Test material is no longer available. Data beyond the details presented in the supplementary data and the Bayesian network in Hugin are sharable upon request and approval of release.

Competing Interests The authors declare no competing interests.

References

1. van den Berge M, Ozcanhan G, Zijlstra S, Lindenbergh A, Sijen T (2016) Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios. Forensic Sci Int: Genet 21:81-89
2. van den Berge M, van de Merwe L, Sijen T (2017) DNA transfer and cell type inference to assist activity level reporting: Post-activity background samples as a control in dragging scenario. Forensic Sci Int: Genet Suppl Ser 6:e591-e592
3. Gill P, Hicks T, Butler JM, Connolly E, Gusmão L, Kokshoorn B, Schneider PM (2018) DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence-Guidelines highlighting the importance of propositions: Part I: evaluation of DNA profiling comparisons given (sub-) source propositions. Forensic Sci Int: Genet 36:189-202
4. Gill P, Hicks T, Butler JM, Connolly E, Gusmão L, Kokshoorn B, Schneider PM (2020) DNA commission of the International society for forensic genetics: Assessing the value of forensic biological evidence-Guidelines highlighting the importance of propositions, Part II: Evaluation of biological traces considering activity level propositions. Forensic Sci Int: Genet 44:102186
5. Gill P, Benschop C, Buckleton J, Bleka Ø, Taylor D (2021) A Review of Probabilistic Genotyping Systems: EuroForMix DNAStatistX and STRmix™. Genes 12(10):1559
6. van Oorschot RA, Meakin GE, Kokshoorn B, Goray M, Szkuta B (2021) DNA Transfer in Forensic Science: Recent Progress towards Meeting Challenges. Genes 12(11):1766
7. Gosch A, Courts C (2019) On DNA transfer: the lack and difficulty of systematic research and how to do it better. Forensic Sci Int: Genet 40:24-36
8. Buckingham AK, Harvey ML, van Oorschot RA (2016) The origin of unknown source DNA from touched objects. Forensic Sci Int: Genet 25:26-33
9. Meakin GE, Kokshoorn B, van Oorschot RA, Szkuta B (2021) Evaluating forensic DNA evidence: Connecting the dots. Wiley Interdisciplinary Reviews: Forensic Sci 3(4):e1404
10. Steensma K, Ansell R, Clarisse L, Connolly E, Kloosterman AD, McKenna L, G, Kokshoorn B (2017) An inter-laboratory comparison study on transfer persistence and recovery of DNA from cable ties. Forensic Sci Int: Genet 31:95-104
11. Goray M, Kokshoorn B, Steensma K, Szkuta B, van Oorschot RA (2020) DNA detection of a temporary and original user of an office space. Forensic Sci Int: Genet 44:102203
12. Szkuta B, Ansell R, Boiso L, Connolly E, Kloosterman AD, Kokshoorn B, van Oorschot RA (2020) DNA transfer to worn upper garments during different activities and contacts: An inter-laboratory study. Forensic Sci Int: Genet 46:102268
13. Benschop CC, Hoogenboom J, Hovers P, Slagter M, Kruise D, Parag R, van Marion V (2019) DNAxs/DNAStatistX: Development and validation of a software suite for the data management and probabilistic interpretation of DNA profiles. Forensic Sci Int: Genet 42:81-89
14. Bhoelai B, Beemster F, Sijen T (2013) Revision of the tape used in a tape-lift protocol for DNA recovery. Forensic Sci Int: Genet Suppl Ser 4(1):e270-e271
15. Nicklas JA, Buel E (2003) Development of an Alu-based real-time PCR method for quantitation of human DNA in forensic samples. J For Sci 48(5):936-944
16. Nicklas JA, Buel E (2006) Simultaneous determination of total human and male DNA using a duplex real-time PCR assay. J For Sci 51(5):1005-1015

17. Lindenbergh A, de Pagter M, Ramdayal G, Visser M, Zubakov D, Kayser M, Sijen T (2012) A multiplex (m) RNA-profiling system for the forensic identification of body fluids and contact traces. Forensic Sci Int: Genet 6(5):565-577
18. Duijs F, van de Merwe L, Sijen T, Benschop CC (2018) Low-template methods yield limited extra information for PowerPlex® Fusion 6C profiling. Legal Med 33:62-65
19. Bleka Ø, Storvik G, Gill P (2016) 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 21:35-44
20. Bouwman I, Hoogenboom J, Sijen T, Benschop CC (2022) Performing LR calculations when loci are missing between reference and trace DNA profiles. Forensic Sci Int: Reports 5:100268
21. Tay JW, Murakami JA, Cooper PL, Rye MS (2019) Sensitivity and baseline noise of three new generation forensic autosomal STR kits: PowerPlex® Fusion VeriFilerTM Plus and Investigator® 24plex QS. Forensic Sci Int: Reports 1:100049
22. Parys-Proszek A, Wróbel M, Marcińska M, Kupiec T (2018) Dual amplification strategy for improved efficiency of forensic DNA analysis using NGM Detect™ NGM™ or Globalfiler™ kits. Forensic Sci Int: Genet 35:46-49
23. Oldoni F, Podini D (2019). Forensic molecular biomarkers for mixture analysis. Forensic Sci Int: Genet 41:107-119
24. Taylor D, Biedermann A, Hicks T, Champod C (2018) A template for constructing Bayesian networks in forensic biology cases when considering activity level propositions. Forensic Sci Int: Genet 33:136-146
25. Benschop CC, Hoogenboom J, Bargeman F, Hovers P, Slagter M, van der Linden J, Sijen T (2020) Multi-laboratory validation of DNAxs including the statistical library DNAStatistX. Forensic Sci Int: Genet 49:102390
26. Westen AA, Kraaijenbrink T, de Medina EAR, Harteveld J, Willemse P, Zuniga SB, de Knijff P (2014) Comparing six commercial autosomal STR kits in a large Dutch population sample. Forensic Sci Int: Genet 10:55-63
27. Samie L, Hicks T, Castella V, Taroni F (2016) Stabbing simulations and DNA transfer. Forensic Sci Int: Genet 22:73-80
28. Steensma K, Ansell R, Clarisse L, Connolly E, Kloosterman AD, McKenna LG, Kokshoorn B (2017) An inter-laboratory comparison study on transfer persistence and recovery of DNA from cable ties. Forensic Sci Int: Genet 31:95-104
29. Ramos P, Handt O, Taylor D (2020) Investigating the position and level of DNA transfer to undergarments during digital sexual assault. Forensic Sci Int: Genet 47:102316
30. De Wolff TR, Aarts LHJ, van den Berge M, Boyko T, van Oorschot RAH, Zuidberg M, Kokshoorn B (2021) Prevalence of DNA of regular occupants in vehicles. Forensic Sci Int 320:110713
31. Goray M, Kokshoorn B, Steensma K, Szkuta B, van Oorschot RA (2020) DNA detection of a temporary and original user of an office space. Forensic Sci Int: Genet 44:102203