International Journal of Legal Medicine (2021) 135:727–738
https://doi.org/10.1007/s00414-021-02503-4

ORIGINAL ARTICLE

Towards developing forensically relevant single-cell pipelines by incorporating direct-to-PCR extraction: compatibility, signal quality, and allele detection

Nidhi Sheth[1] · Harish Swaminathan[2] · Amanda J. Gonzalez[3] · Ken R. Duffy[4] · Catherine M. Grgicak[1,2,3]

Received: 17 August 2020 / Accepted: 4 January 2021 / Published online: 23 January 2021

Abstract

Current analysis of forensic DNA stains relies on the probabilistic interpretation of bulk-processed samples that represent mixed profiles consisting of an unknown number of potentially partial representations of each contributor. Single-cell methods, in contrast, offer a solution to the forensic DNA mixture problem by incorporating a step that separates cells before extraction. A forensically relevant single-cell pipeline relies on efficient direct-to-PCR extractions that are compatible with standard downstream forensic reagents. Here we demonstrate the feasibility of implementing single-cell pipelines into the forensic process by exploring four metrics of electropherogram (EPG) signal quality—i.e., allele detection rates, peak heights, peak height ratios, and peak height balance across low- to high-molecular-weight short tandem repeat (STR) markers—obtained with four direct-to-PCR extraction treatments and a common post-PCR laboratory procedure. Each treatment was used to extract DNA from 102 single buccal cells, whereupon the amplification reagents were immediately added to the tube and the DNA was amplified/injected using post-PCR conditions known to elicit a limit of detection (LoD) of one DNA molecule. The results show that most cells, regardless of extraction treatment, rendered EPGs with at least a 50% true positive allele detection rate and that allele drop-out was not cell independent. Statistical tests demonstrated that extraction treatments significantly impacted all metrics of EPG quality, where the Arcturus® PicoPure™ extraction method resulted in the lowest median allele drop-out rate, highest median average peak height, highest median average peak height ratio, and least negative median values of EPG sloping for GlobalFiler™ STR loci amplified at half volume. We, therefore, conclude the feasibility of implementing single-cell pipelines for casework purposes and demonstrate that inferential systems assuming cell independence will not be appropriate in the probabilistic interpretation of a collection of single-cell EPGs.

Keywords Forensic DNA · Human identification · Single-cell forensic DNA · Direct-to-PCR · Forensic DNA Mixtures

Abbreviations

STR	Short tandem repeat
CE	Capillary electrophoresis
EPG	Electropherogram
PCR	Polymerase chain reaction
RFU	Relative fluorescence unit
b.p.	Base pairs
ILS	Internal lane standard
LoD	Limit of detection

✉ Catherine M. Grgicak
c.grgicak@rutgers.edu

[1] Center for Computational and Integrative Biology, Rutgers University, Camden, NJ 08102, USA

[2] Biomedical Forensic Sciences Program, Boston University School of Medicine, Boston, MA 02118, USA

[3] Department of Chemistry, Rutgers University, 315 Penn Street R306C, Camden, NJ 08102, USA

[4] Hamilton Institute, Maynooth University, Maynooth, Ireland

Introduction

Traditional forensic genotyping pipelines typically consist of a multi-step workflow that includes sample collection, DNA extraction, quantification, amplification of forensically relevant short tandem repeats (STRs), capillary electrophoresis (CE), peak detection, analysis, and interpretation. Forensic samples often contain material from an unknown number of unknown contributors, requiring sophisticated interpretation tools to evaluate the weight of the evidence. This is typically

reported as a log likelihood ratio (LR), which is the logarithm of the probability of observing the data given the suspect contributed divided by the probability of the evidence given they did not [1–7]. Although probabilistic evaluation of forensic evidence using increasingly sophisticated algorithms has gained traction in recent years, bulk processing samples that contain more than three contributors render electropherograms (EPGs) that are so complex [8] that significant computing power is often required to complete their statistical evaluation [5, 6]. Furthermore, even in the presence of nucleic acid material from the suspect, the likelihood ratio tends towards one as the number of contributors increases or as DNA template levels decrease [9, 10], resulting in reduced inferential power. In extreme circumstances, the number or height of the observed peaks may indicate that a sample contains more than four contributors, and in those cases it may not be feasible for the LR to be computed due to the computational complexity of inferential algorithms scaling exponentially as a function of the number of contributors. Data obtained from multiple cells that have undergone a bulk extraction/amplification process are, therefore, prone to inconsistent interpretation since each laboratory necessarily selects their own signal detection parameters, interpretation paradigm, and procedure [11–14].

An alternative to the bulk processing pipeline is one that relies on the extraction, amplification, and fragment analysis of each cell individually. In the single-cell process, a sample is collected and each cell is separated prior to extraction. As a result, the data consists of n EPGs, each from an individual cell, rather than a single EPG from n cells. Though single-cell pipelines have obvious advantages over bulk processing schemes [15–18], they necessitate the application and development of novel extraction and interpretation strategies that meet forensic requirements. While each step of the forensic pipeline contributes error to the final result [19–21], none has more of an effect on the signal quality than DNA extraction [22], as numerous studies have reported as much as a 96% loss in DNA yield [23–27] at this step. Not only are traditional extraction methods prone to low DNA yields; their relatively large end-volumes (i.e., 20–100 μL) [28, 29] coupled with the small volumes used to prepare the PCR reaction (i.e., 1–15 μL) [30, 31] cause fractionation of the extract leading to allele non-detection, also referred to as allelic drop-out, rates that are dictated by the volume of liquid pipetted to the PCR tube, V_{PCR}, divided by the extract's end-volume, V_{ext} [22]. In the case of single cells, any fractionation would necessarily be detrimental to obtaining a full profile. Thus, efficient direct-to-PCR extraction, where the PCR reagents are added to the extract vessel/tube containing all the DNA, is required before single-cell analysis can be introduced into forensic operations. This requires direct-to-PCR extraction reagents that are compatible with accepted PCR components to efficiently co-amplify tens of short tandem repeats (STRs) ranging from 80 to 400 b.p. in size.

Though DNA "fingerprints" from single cells were first described in 1997 by Findlay et al. [32], who reported 114 of 226 cells rendered a full profile, forensic signal from ones or tens of copies remains far from ideal. Recent studies continue to report high drop-out rates, e.g., 26.9% for 75 sperm cell using laser microdissection (LMD) [33] and 15.1% for extracts containing DNA from five cells [34]. Still others reported that the overall drop-out rate for single-cell sperm samples was 30% [35], while drop-out rates from 10 cells were approximately 13–19% in [36]. For white blood cells (WBC), drop-out rates of 18% per cell were reported by the authors of [37], while rates from 10 cells ranged from 0 to 3% in [36]. Studies focusing on epithelial cells corroborate previous findings, where the drop-out rate for 10 epithelial cells was reported to range from 2.4 to 5.2% across three donors [36].

Notably, the aforementioned studies significantly varied in their experimental protocols, possibly explaining the disparate reports. Allelic drop-out, generally, originates from two sources: (1) sampling effects, where no allele copies survive the pre-PCR process [20], and (2) detection effects, where copies that survive pre-PCR steps do not reach a sufficient number of amplified copies to elicit detection [38]. On the whole, these studies did not provide sufficient descriptions on post-PCR protocols to ascertain whether sub-optimal PCR and post-PCR effects confounded the reported drop-out rates, demonstrating that an end-to-end single-cell pipeline has yet to be well-articulated in the forensic domain. Despite this, simulations by the authors of [38, 39] show that by optimizing post-PCR protocols, the effects of (1) can be clearly described. This simulation work was recently experimentally corroborated in [20].

This study, therefore, presents a detailed comparison of four common direct-to-PCR extraction chemistries with the aim of producing a single-cell pipeline for forensic purposes. In this work, the post-PCR protocols have been chosen to ensure a post-PCR limit of detection (LoD) of one copy. That is, the post-PCR conditions were set so as to isolate the source of drop-out to the pre-PCR steps, thereby allowing a direct comparison between extraction chemistries. Specifically, 102 single-cell samples were extracted using four direct-to-PCR extraction treatments and amplified using accepted forensic PCR assays and common analysis platforms. To inform downstream single-cell interpretation strategies and our statistical testing, we evaluated whether allele drop-out rates are cell independent. Next, we explored differences between the four extraction methods by interrogating profile quality metrics such as allele drop-out, allele peak heights, peak height ratios, and peak heights as a function of molecular weight (i.e., EPG sloping). Extraction methods resulting in the lowest allele drop-out rates, highest peak heights and peak height ratios, and lowest EPG sloping are considered viable options for the single-cell forensic pipeline.

Materials and methods

Sample collection

Buccal swabs were obtained from three unknown individuals (samples A, B, and C) in accordance with the ethical standards of the Institutional Review Board, Protocol number Pro2018002536. Thirty-four cells per sample were tested across 4 extraction treatments, resulting in a total of 408 single cells being interrogated. When determining allele detection rates, peak height ratios, and other related statistics, only heterozygous loci with STRs at least two repeats apart were used. This represented 10, 13, and 15 loci for samples A, B, and C, respectively. In summary, for each extraction treatment, 10,336 alleles were evaluated.

Sample preparation and single cell collection

A total of 1 μL of whole saliva was washed three times by adding 999 μL TE buffer in a microcentrifuge tube and centrifuging at maximum angular velocity for 3 min. After the final wash was completed, 150 μL of TE buffer was added to the pellet and gently mixed to evenly distribute the cells in solution. Of the washed sample, 150 μL was then aliquoted onto a glass microscope slide, and the cells were visualized using a Nepa Gene 3 OmniClass Compass video microscope. Individual epithelial cells were pico-pipetted into a well containing 5 μL of extraction solution using the MPP-300 Micro Pick and Place System (Bulldog Bio, Portsmouth, NH). A total of 102 single cells (i.e., 34 cells per individual) were extracted using four different extraction methods, which are described in the following section.

DNA extraction protocols

In all cases we base the extraction protocol on the manufacturer's recommendations unless otherwise stated. When deviations from recommendation occur, we stipulate what they are and provide background information in Supplement 1 describing preliminary or optimization work describing the modification. In all cases, the final extraction volume was 5 μL.

Arcturus® PicoPure™ DNA extraction Each of the 102 cells was dispensed into one well of a 96-well plate containing 5 μL aliquots of PicoPure DNA extraction buffer (PicoPure DNA Extraction Kit, Arcturus, CA), where the extraction buffer was prepared by adding 155 μL of reconstitution buffer into one supplied vial of proteinase K according to recommendations [40]. Upon the addition of a cell to the extraction buffer, the sample was vortexed, centrifuged at 1000×*g* for 30 s, and incubated at 65 °C for 3 h. At the end of incubation, the sample was heated to 95 °C for 10 min to inactivate the proteinase K and immediately amplified using the amplification protocol outlined below.

DEPArray™ LysePrep DNA extraction Each of the 102 cells was dispensed into wells containing 5 μL of the LysePrep extraction mix. The LysePrep extraction mix (DEPArray™ LysePrep DNA Extraction Kit, Menarini Silicon Biosystems, Italy) was prepared based on the manufacturer's recommendations for FFPE (formalin-fixed paraffin-embedded) tissues [41] by adding 2 μL of TE buffer, 1 μL of PBS buffer, 1.28 μL of water, 0.2 μL of kit buffer, 0.13 μL each of reagent 1 and reagent 2, and 0.26 μL of proteinase K. Notably, the addition of TE buffer represents a modification to the component composition, though no significant impact on the final outcome was reported during preliminary testing (cf. Supplement 1). The samples were incubated at 42 °C for 15 h, briefly centrifuged, and heated at 80 °C for 10 min to inactivate the proteinase K. Incubation temperatures and times and the effects on downstream allele detection rate during preliminary/optimization studies are also presented in Supplement 1. Based on these results, the modified FFPE LysePrep procedure was chosen as the final candidate treatment. Amplification immediately followed extraction.

DirectPCR Lysis extraction Each cell was dispensed into wells containing 5 μL of DirectPCR Lysis extraction mix (DirectPCR Lysis Reagent (Cell), Viagen Biotech, CA), which was prepared by adding 4.28 μL of DI water, 0.48 μL of DirectPCR Lysis Reagent, and 0.25 μL of proteinase K. The samples were incubated at 55 °C for 6 h after briefly centrifuging. Proteinase K inactivation occurred at 85 °C for 45 min based on recommendations from the manufacturer [42], and amplification immediately followed.

ForensicGEM® Zygem extraction The cells were dispensed into wells containing 5 μL of the extraction mix. The forensicGEM extraction mix (forensic GEM® Zygem, MicroGEM International PLC, VA) was prepared according to the manufacturer's recommendations for saliva [43] by adding 4.45 μL of DI water, 0.5 μL of 10× Blue buffer (Tris buffer), and 0.05 μL of forensicGEM solution (EA1 enzyme containing proteinase K and glycerol). The samples were incubated at 75 °C for 15 min and heated at 95 °C for 5 min to inactivate proteinase K. As before, all amplifications occurred immediately following extraction.

Amplification and capillary electrophoresis

Direct single-cell STR amplification was performed with a total reaction volume of 12.5 μL using GlobalFiler™ PCR Amplification Kit (Life Technologies Corporation) where 7.5 μL of amplification reaction mix was directly added to each well. All thermal cycling temperatures, ramp speed,

and soaking times followed the manufacturer's recommendations [30] for the GeneAmp® 9700 PCR thermal cycler (Applied Biosystems™). A total of 30 PCR cycles were used. In preparation for fragment separation, 1 μL of PCR product was added to 9.7 μL of HiDi formamide and 0.3 μL of GeneScan™ 600 Liz® (60–460) Size Standard v2.0 (Applied Biosystems™). Products were separated and detected using an injection potential and time of 1200 V and 25 s, respectively, on an Applied Biosystems 3500 Genetic Analyzer. All profiles were analyzed using GeneMapper® ID-X v1.4 (Applied Biosystems™) with an analytical threshold of 30 relative fluorescence units (RFU). Notably, these post-PCR conditions have been shown to amplify and inject sufficient numbers of amplicons to exceed the detection threshold of 30 RFU at the single-copy regime in the absence of PCR inhibitors [20, 38, 44]; thus, any differences in signal quality are attributable to impacts imparted by extraction efficiencies or reagent incompatibility and are, therefore, not confounded by post-PCR non-detection effects.

Data preparation and data clean-up

In this study the genotype and source of the single-cell DNA are known. Known genotypes were determined by extracting a large volume of saliva with Qiagen Investigator DNA Extraction kit [28] and amplifying 0.3 ng of DNA using the manufacturer's recommended amplification and injection protocols for the GlobalFiler™ STR assay [30]. The single-cell data were authenticated by first confirming consistency between the observed alleles to the known genotype. Signal that could be classified as originating from uncommon artifact sources, such as pull-up from the ILS, dye blobs, and spikes, was manually removed during GeneMapper® ID-X v1.4 analysis, at which time the data was exported as a CSV file for further processing. The more common artifact peaks, such as pull-up (i.e., bleed-through from other fluorescent color channels) and incomplete adenylation (i.e., minus A), were automatically filtered using a script, named CleanIt, as per settings described in detail in [45]. Briefly, a peak was considered pull-up if it was the same size (± 0.3 b.p.) as a larger peak in another color and below 5% of the height of the larger peak. In the case of ILS pull-up, the range was increased to ± 1.6 b.p. Peaks were designated as "minus A" if they were one base pair smaller than an allele.

Profile analysis

EPG quality was ascertained by exploring allele detection rates, peak heights, peak height ratios, and the degree to which high-molecular-weight markers amplify in relation to low-molecular-weight markers. If the high-molecular-weight markers do not amplify as well, the peak heights are lower than those of their low-molecular-weight counterparts, and the EPG exhibits a "sloping" effect. Prior to statistically evaluating these quality metrics across extraction types, we test the hypothesis that allele drop-out is cell independent.

Allele drop-out cell independence

To ascertain if the number of detected alleles was cell independent, a permutation test, written in MATLAB R2018b, was conducted [46]. The statistic used was the variance in the measurement of the number of detected alleles across heterozygote loci for each cell. With a null hypothesis that each detected allele count was independent of the cell, but possibly dependent on the locus, the test was performed by creating 10^6 data sets by random permutations of the number of recovered alleles among cells, independently at each locus, and re-valuation of the variance in the number of per-cell detected alleles. A one-sided test, where the p value is the proportion of permutations that resulted in a variance that was higher than for the true statistic, was evaluated. A p value cutoff of 0.05 was used for rejecting the null hypothesis.

Allele drop-out between extraction treatments

We used the permutation test function in JMP® Pro 14 to test the null hypothesis that the number of alleles detected is interchangeable between extraction methods. To accomplish this, the number of detected alleles per cell was randomly shuffled among the four extraction treatments. Under the null hypothesis of no effect, any of these allocations are as likely as any other. The F-ratio obtained in this manner approximates the F-ratios under the null hypothesis such that the measured F-ratio can be compared to the null distribution to obtain the p value. A p value threshold of 0.05 was used to reject the null hypothesis, and 10^6 permutations were used to construct the null distribution.

Average peak height ratio and average peak height between extraction treatments

Permutation tests, using the F-ratio as the statistic, were again used to test if the average peak height ratio and average peak height within a cell were interchangeable between kits. Here the average peak height ratio and average peak height entry were randomly shuffled among extraction types. To compute the peak height ratio for the locus, the height of the less intense allele was divided by the height of the more intense allele. To compute the average peak height, the sum of the peak heights across the known alleles was divided by the number of known allele positions across the heterozygous loci. If the allele was not detected, a peak height of 1 RFU was used.

EPG contours between extraction treatments

The final quality metric examined was EPG sloping, which is a measure of DNA degradation or inhibition. Steep EPG contours are undesirable, and single-cell pipelines ought to implement extraction protocols that minimize this trait. Since all cells were prepared in the same manner immediately upon receipt, significant degrees of EPG sloping would indicate either an incompatibility between extraction and PCR reagents [47], or an excessively harsh extraction procedure causing DNA damage, or an inefficient extraction where the DNA is not adequately mined from the histones and other proteins that promote DNA packing. We therefore assess EPG sloping across extraction kits by first modeling the exponential decay in fluorescence as a function of molecular weight for each cell per EPG [48]:

$$H_l = \alpha e^{\beta \overline{s_l}} \quad (1)$$

where H_l is the sum of the peak heights associated with the known genotypes at locus l, $\overline{s}$ is the average base pair size of the STR alleles at locus l, and α and β are the exponential parameters obtained for each sample using least squares regression. In extreme cases of decay, the highest-molecular-weight peaks may not reach detectable levels. In these cases, a peak height of 1 RFU was assumed. If high-molecular-weight markers exhibited low peak heights, β will take a large negative value. In contrast, if there is good signal balance across all loci, indicating efficient PCR and high-quality template DNA, β will be near zero. Notably, unlike the other statistics, all loci are used to acquire the parameters of Eq. 1. Again, permutation tests, using the F-ratio statistic, were used to test if β values were exchangeable between extraction types.

Results and discussion

A total of 102 single cells, 34 from each of three individuals (labelled A, B, and C), were extracted using four direct-to-PCR methods. Thus, in total, 408 single-cell profiles were analyzed for this study. The extract end-volume was kept constant at 5 μL, and 7.5 μL of PCR master mix was directly added to the extract tube, thereby avoiding stochastic sampling effects that can drive allele drop-out [20].

Figure 1a depicts a representation of the forensic single-cell pipeline. Here, single nucleated epithelial cells are pipetted into a single tube or well, extracted, and amplified producing one electropherogram for each cell, where one cell may produce an EPG with high signal-to-noise ratio while another does not. In this process the DNA copies are not evenly distributed and the extract is not fractionated. In the analysis of traditional bulk cell mixtures, one only has a single EPG to assess, and so it is natural to assume that there is a single likelihood of allele drop-out. If single cell methods are to be employed for casework, however, it will be necessary to understand whether there is a statistical consistency in the quality of the signal of EPGs generated from distinct cells or whether some contain more information than others. Drop-out is one significant correlate to EPG quality as an EPG with fewer measured alleles holds less information on the source cell's genotype, so it is necessary to test if drop-out is consistent across single-cell EPGs, which will have implications for downstream interpretation of collections of single-cell EPGs.

Allele drop-out cell independence

Figure 1c is a histogram of the frequency of the proportion of heterozygous alleles detected per cell, across all 408 cells. The expected binomial distribution under the null hypothesis that allele dropout is cell and locus independent was calculated using the empirically determined overall drop-out rate of 0.33. Figure 1b is a pie chart that reports the percentage of samples rendering full, partial, little, and no allelic information. These data exhibit several interesting features. First, obtaining a complete profile is a rare event. Specifically, of the 408 single buccal cells, only 6% (Fig. 1b) resulted in full profile representations, suggesting allelic drop-out is common even when using post-PCR conditions with a detection limit of one copy. This is consistent with the findings of Williamson et al. [35], but inconsistent with the findings of Geng et al. who regularly produced full profiles [49]. Inconsistencies between the two studies may be explicated by an examination into the sample types: Geng et al. extracted and amplified the DNA of human lymphoid cells that were grown in a medium under controlled conditions, while Williamson et al. extracted and amplified the DNA of buccal, sperm, and blood cells contained on cotton-tipped applicators, which are more representative of sample qualities submitted to the forensic laboratory. Similar to the Williamson et al. study, Findlay et al. [32] analyzed the STR profiles of 226 single buccal cells from 4 different contributors amplified for 34 cycles and showed an overall drop-out rate of 39% where 114 of 226 (50%) cells produced full profiles. The same authors also reported complete drop-out in 20 of 226 (9%) single-cell profiles, which is consistent with the second feature depicted in Fig. 1b and c; that is, only 7% of single cells rendered full profile drop-out. The third notable feature seen in Fig. 1c is that the non-zero mode is (0.8–0.9], suggesting that the information content contained in most single-cell EPGs is high, and even the most temperamental of cell types (i.e., buccal cells) is likely to provide enough signal to adequately determine the weight of evidence against a person of interest. Notably, these detection rates were an improvement to the values reported by the authors of [32], who reported ≥ 4 alleles were detected in only 64% of the profiles showing significant progress towards single-cell analysis over the last two decades.

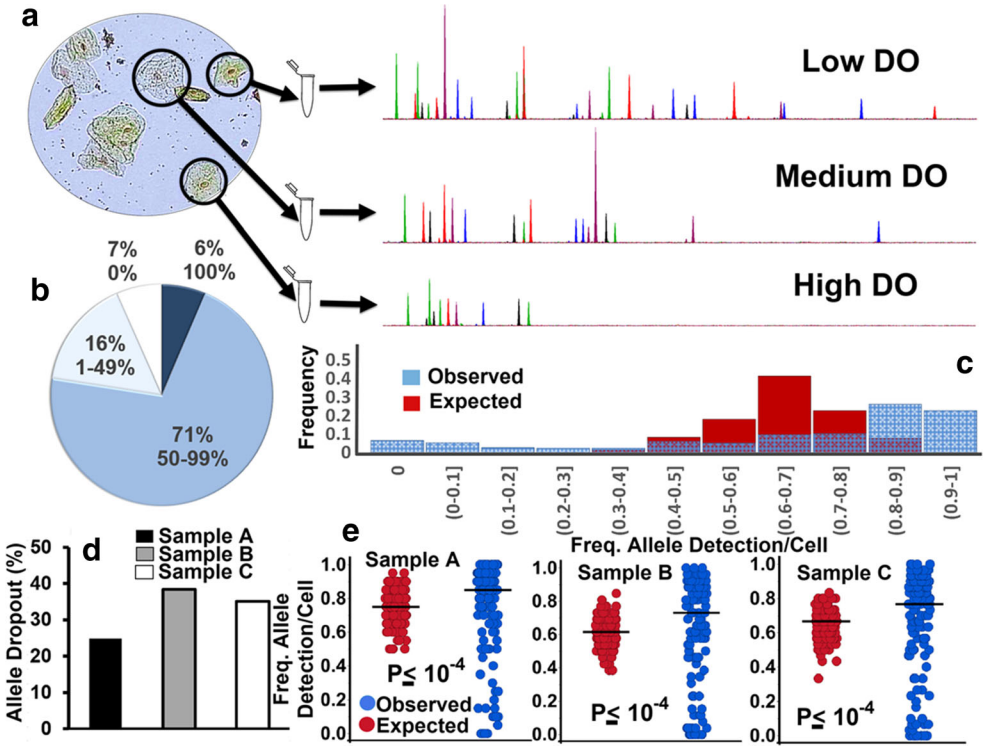

Fig. 1 **a** Representation of a sample, consisting of epithelial cells, undergoing forensically relevant direct-to-PCR single-cell processing. A slide containing epithelial cells and viewed at × 200 where single, nucleated cells are pico-pipetted into a tube/well and the DNA is extracted using direct-to-PCR treatments. Amplification and fragment analysis follow. The EPGs depict representative low, medium, and high allelic drop-out rates observed for three different cells collected from the sample. **b** Pie chart depicting the percentage of 408 single cells exhibiting (▪) all (100%), (▪) most (50–99%), (▪) some (1–49%), and (□) no signal from heterozygous loci, where the alleles are at least two STRs apart. **c** Histograms of the frequency of the proportion of alleles detected per cell for heterozygous alleles across all 408 single cells, where (▪) indicates the allele detection rates per cell and (▪) the expected detection distribution if drop-out was independent of the cell and locus. **d** A bar chart depicting the overall probability of drop-out per cell separated by sample. **e** A scatterplot expressing (•) the frequency of heterozygous alleles detected and (•) the frequency of expected heterozygous alleles for each sample had allele drop-out been cell and locus independent

Figure 1d is a bar chart of the probability of allele drop-out for each sample, calculated by dividing the total number of detected heterozygote alleles by the true number of heterozygous alleles. Overall, the drop-out rates across individuals were relatively consistent, ranging from 25 to 38%. Figure 1e presents scatterplots depicting the measured detection rates (blue) and the binomial distribution of the fraction of detections per cell that would be expected if allele drop-out was cell and locus independent (red), separated by sample. A clear mismatch between the null hypothesis and these data is evident. To precisely assess this observation, we performed the one-sided permutation test described in the "Allele drop-out cell independence" section. This gave a p value of $\leq 10^{-4}$, resulting in rejection of the null hypothesis that drop-out is cell independent, for every combination of kits and source genotypes.

Allele drop-out between extraction types

Figure 2 is a summary of the allele drop-out results obtained between direct-to-PCR extraction methods, where Fig. 2a depicts four representative EPGs from the same sample (i.e., sample B) with known allele positions marked in red. No unusual features such as peak broadening or an increase in the number or intensity of artifacts (e.g., EPG spikes or raised baselines) for any of the extraction treatments were observed for any extraction treatment. Next, we categorized each EPG as containing all (100%), most (50–99%), some (1–49%), or no (0%) heterozygous alleles per cell (Fig. 2b) for each extraction treatment and observed that of the four treatments, PicoPure™ (PP) and LysePrep (LP) resulted in the highest percentage of all, or most, heterozygous alleles detected at 87% and 80%, respectively. The scatterplot in Fig. 2c shows the median detection rates per cell across treatments, which were 0.7, 0.65, 0.85, and 0.90 for DirectPCR Lysis Reagent (DP), forensicGEM® Saliva (FG), DEPArray™ LysePrep (LP), and Arcturus® PicoPure™ DNA Extraction Kit (PP), respectively. A permutation test with the null hypothesis that the number of alleles detected is interchangeable across extraction methods resulted in a p value $\leq 10^{-4}$, suggesting systemic differences in kit performance and that single-cell allele drop-out can partially be mitigated by employing an extraction method that is well-designed for these strategies. Notably, the PicoPure™ treatment stands out as the treatment with the highest number of median alleles detected and rendered the highest percentage of full and near-full profiles.

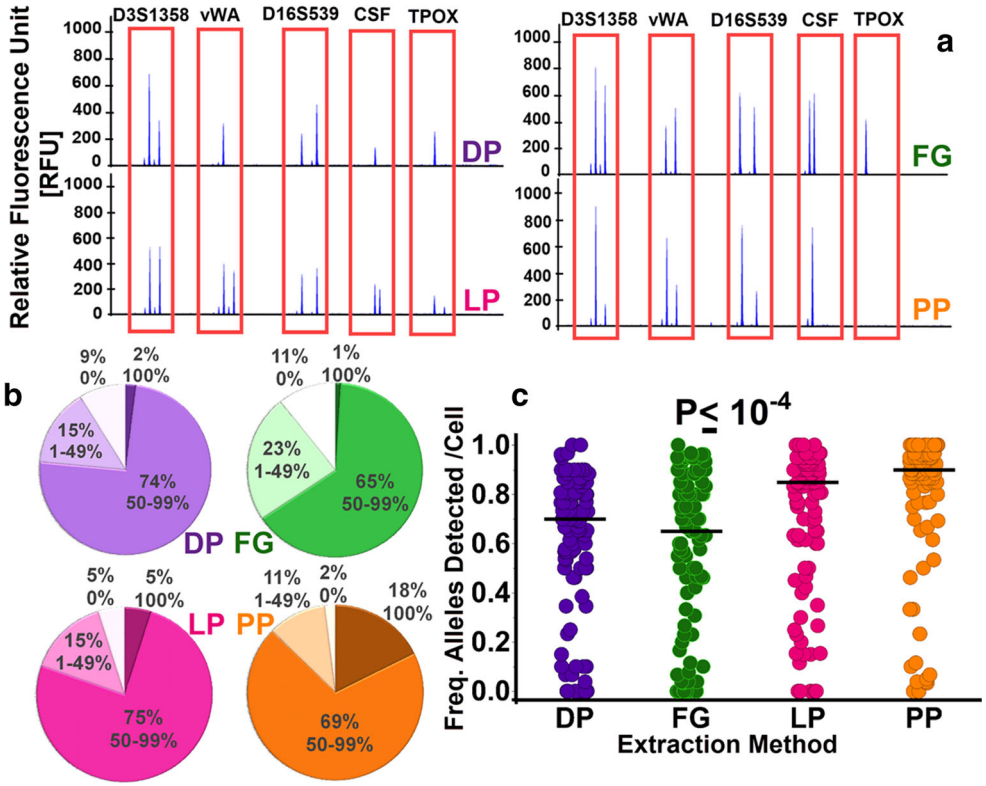

Fig. 2 **a** Four representative EPGs (blue channel) across extraction treatments for person B. The red boxes depict the locations of the known alleles. **b** Pie charts showing the number of profiles exhibiting full (100%), most (50–99%), some (1–49%), and no heterozygous alleles of 102 single-cell profiles for each extraction method. **c** A scatterplot depicting the frequency of detection for all heterozygous alleles separated by extraction method, with median values of 0.7, 0.65, 0.85, and 0.9 for DirectPCR Lysis Reagent (DP), forensicGEM® Saliva (FG), DEPArray™ LysePrep (LP), and PicoPure™ DNA Extraction Kit (PP), respectively

Average peak height ratio and average peak height between extraction types

Characterizing and developing models that reasonably predict peak height ratio or height distributions are an important component of many continuous probabilistic genotyping systems [7, 50–52]. Though applying reasonable models that describe peak height distributions is expected to take a smaller role in single-cell pipelines, good signal intensity and reasonable peak height ratios are still desirable, particularly when attempting to ascertain if a peak is likely stutter or allele or, perhaps, both. In addition, good signal is an indication of adequate amplification efficiencies and, therefore, reasonable extraction efficiencies and good compatibility between extraction and PCR reagents. Thus, to assess if the EPG signal is affected by the extraction methodology employed, the average peak height ratio and average peak height for each cell across treatments were evaluated (Fig. 3).

Figure 3a summarizes the peak height balance within a locus as a density plot of the height of the less intense allele plotted against the height of the more intense one. The x = y line is also shown. In cases where there is good heterozygous balance, the majority of the points would fall near or just below the x = y line, while extraction reagents that are incompatible with downstream PCR processes would result in a higher density of points approaching the origin. Across the four treatments, the peak heights ranged from 3117 (outlier not shown) to the analytical threshold, with the majority of peak heights falling below 1500 RFU (Fig. 3a). In particular, the PicoPure™ treatment resulted in plots with a greater density of points falling near the x = y line and further away from the origin, indicating good levels of peak balance and peak height. To evaluate this more closely, in Fig. 3b is a scatterplot that shows the distribution of the average peak height ratio within each cell for the four extraction treatments. The median average peak height ratio per cell was 0.51, 0.55, 0.59, and 0.61 for DirectPCR Lysis Reagent, forensicGEM® Saliva, DEPArray™ LysePrep, and Arcturus® PicoPure™ DNA Extraction Kit, respectively. A permutation test with the null hypothesis that the peak height ratio is exchangeable across extraction methods produces a p value $\leq 10^{-4}$, suggesting the extraction method does, indeed, impart statistically significant differences in heterozygous balance. Notably, however, at least 70% of the EPG/cells resulted in an average peak height ratio exceeding 0.7, regardless of treatment. Complementary to Fig. 3a and b is the scatterplot of Fig. 3c which depicts the average peak height per cell for all heterozygous alleles across treatments. As expected, most samples render an average per-cell peak height in the hundreds of RFU, with medians of 320, 306, 350, and 509 RFU for DirectPCR Lysis Reagent, forensicGEM® Saliva, DEPArray™ LysePrep, and

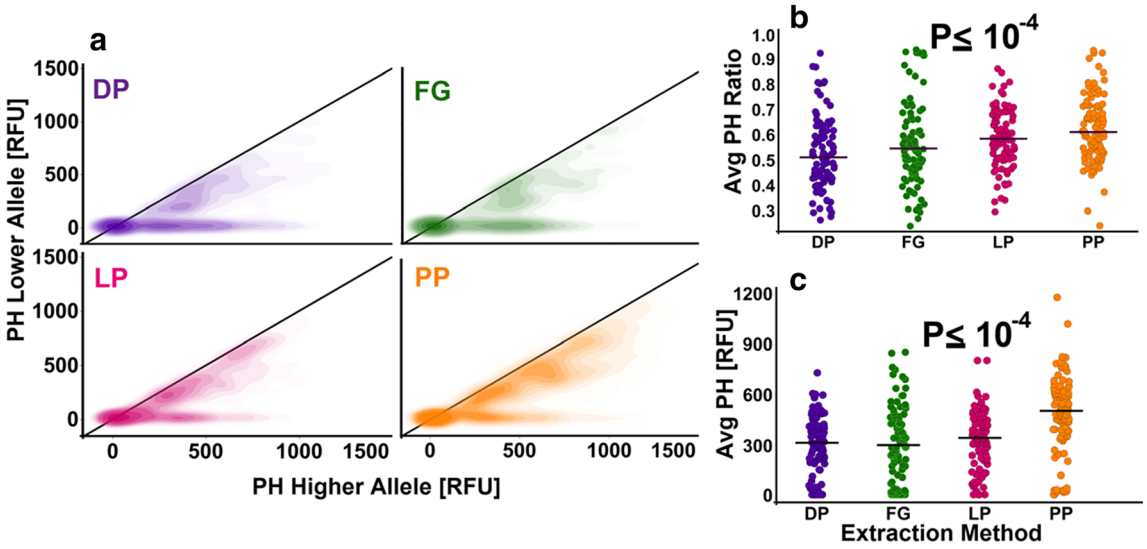

Fig. 3 **a** Density plots of the heights of the less intense peak plotted against heights of the more intense allele for all heterozygous loci where the known alleles were at least two STR units apart for DirectPCR Lysis Reagent (DP), forensicGEM® Saliva (FG), DEPArray™ LysePrep (LP), and PicoPure™ DNA Extraction (PP) treatments. Scatterplots of **b** the average peak height ratio per cell and **c** the average peak height for each cell across the four extraction treatments. Also shown are the associated medians

Arcturus® PicoPure™ DNA Extraction Kit, respectively. A permutation test with the null hypothesis that the average peak height is interchangeable across extraction methods rendered a p value of $\leq 10^{-4}$, suggesting that the extraction method also imparts differences in EPG signal intensity. Notably, the PicoPure™ treatment is the treatment associated with the highest median average peak height and peak height ratio per EPG.

EPG contours between extraction treatments

The EPGs of forensically relevant STRs typically show a slight downward trend with increasing molecular weight [48]. In this work, EPG sloping was determined by least squares regression to an exponential curve with average peak heights and average molecular size, in b.p., as the dependent and independent variables, respectively (Fig. 4a). The more negative the exponential parameter, β, the higher the degree of EPG downward sloping. Figure 4b summarizes the β values as a scatterplot across the four extraction methods. The median β values were -0.007, -0.004, -0.005, and -0.002 for DirectPCR Lysis Reagent, forensicGEM® Saliva, DEPArray™ LysePrep Kit, and Arcturus® PicoPure™, respectively, with the p value $\leq 10^{-4}$, suggesting rejection of the null hypothesis that EPG sloping is interchangeable between extraction types. Despite rare instances of severe sloping (i.e., $\beta < -0.02$), the majority of single-cell EPGs exhibited good peak heights across the length of the fragments, with the PicoPure™ exhibiting the least negative EPG sloping as measured by β. Though the origins of EPG sloping are unknown and can be due to a variety of factors including PCR inhibition due to co-elution of PCR inhibitors [53], inefficient extraction of the DNA from proteins that enhance DNA packing, or DNA damage that occur before extraction [54], we note that these samples were all processed upon arrival and in the same manner, thereby reducing the impact of confounding effects related to DNA differences induced by pre-extraction factors. Thus, significant differences in EPG sloping can be taken as evidence that the DNA extraction treatment had a significant impact on measures of EPG sloping.

Prior to extraction, each cell is pipetted into a well and this step is repeated until the required number of cells is reached. If each cell takes ca. 60–90 s to sequester, pipette, and dispense, then by the time the last cell is transferred, the first was in extraction buffer for some time. It is, therefore, of interest to assess if the length of time cells remains in extraction buffer impacts overall EPG quality. We measure this by interrogating the relationship between the exponential parameter, β, and its well number, as the latter increments in the order the cells were transferred. Figure 5 shows scatterplots of each cell's β value against well number and the corresponding linear regression parameters, as well as results from the F-test for linear regression, which expresses whether the independent variable explains some of the variation seen in the data. Interestingly, the LP treatment results marginally improve the longer the cell is in solution, while DirectPCR shows the reverse trend. In all cases, the highest R^2 value was 0.096, suggesting that the majority of the variation in β

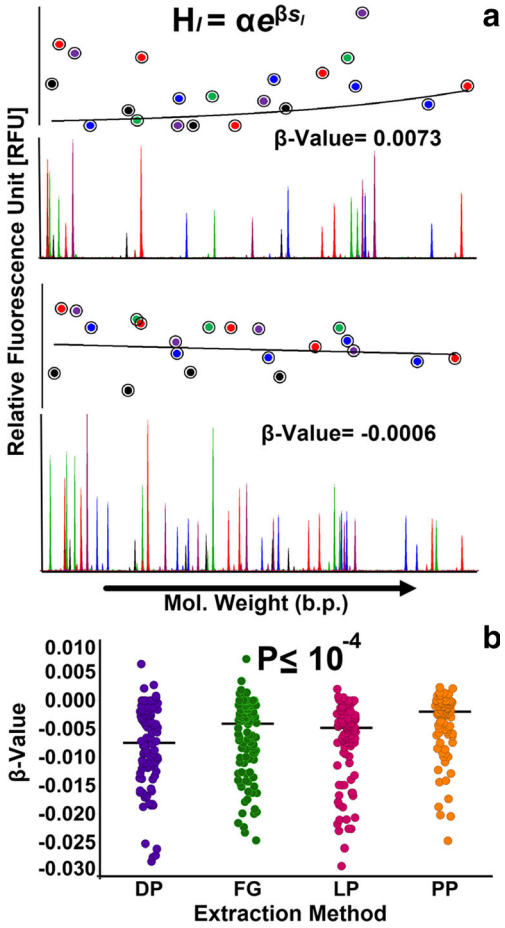

Fig. 4 **a** Representative EPGs obtained when amplifying DNA procured from a single epithelial cell for 30 cycles using the GlobalFiler™ assay, which contains 21 STR tri- or tetra-nucleotide repeats ranging from ca. 80 to 400 b.p. The exponential parameter, β, is a measure of EPG sloping and is obtained by least squares regression of the sum of the allele's heights versus the average molecular weight of the said alleles. Lower β values indicate a more dramatic decrease in the peak height as the fragment length increases that, based on this experimental design, would be indicative of incompatibilities between the extraction and PCR components or DNA damage induced during extraction. **b** The scatterplot indicates the calculated β value per cell across all four extraction methods, with the black bar representing the median values of − 0.007, − 0.004, − 0.005, and − 0.002 for DirectPCR Lysis Reagent (DP), forensicGEM® Saliva (FG), DEPArray™ LysePrep Kit (LP), and PicoPure™ (PP) extraction treatments, respectively

value originates from sources other than the time spent in extraction buffer.

Conclusion

The adoption of computational solutions for complex mixed data obliges the forensic DNA scientist to think deeply about the forensic laboratory process, which is a system of interleaving parts that interact with one another, impacting the final report. The data generated from the amplification of DNA from potentially partial genomes from any number of unknown contributors is complex and, therefore, compels the forensic domain to continuously implement technological advances related to data analysis and interpretation. Implementation of new technology, however, requires a system-based evaluation of the effect of each component to ensure compatibility and, therefore, integrity of the data and its interpretation.

To that end, we explore the feasibility of adopting a single-cell strategy for forensic purposes by assessing the compatibility of four extraction treatments within our current STR pipeline. Though the initial cost of implementing picopipetting and other micromanipulation techniques is relatively low, these techniques are not easily automated and, therefore, require much analyst engagement. Notably, epithelial cells are largely relative to other pertinent cell types, i.e., leukocytes and sperm, which makes this a reasonable method for larger cells, but requires more analyst precision and training as the cells become smaller. Despite these limitations, it is a viable option for operating laboratories engaged in single-cell work and can play an important role in early adoption of single-cell techniques by forensic practitioners.

Whatever single-cell sequestration technique is adopted into operations, it must be anchored by a cohesive analytical strategy and coupled with an interpretation procedure based on sound statistical principles. For these reasons we first explored and demonstrated that the allele drop-out rate is not cell independent, suggesting that statistical analysis should be performed on a per-cell basis. In addition, this observation informs us that any development of single-cell inference platforms ought not assume drop-out is cell independent, which is notably contrary to the commonly employed bulk-mixture platform assumptions [1, 4]. The reasons for this cell dependency are likely varied and vast but include cell apoptosis, the cell's life cycle, and cell rupture during processing. Though improved compatibility between pre- and post-PCR analytical procedures were observed with the PicoPure™ treatment, which are notably from the same vendor, all treatments were successfully implemented, and all showed high levels of allele detection rates and low EPG sloping. These results, therefore, suggest that pre- and post-PCR reagent compatibility cannot be assumed and will likely be driven by each laboratory's single-cell requirements. In conclusion, this work supports the position that forensic single-cell processing for resolving the complex DNA mixture problem is a viable alternative to bulk processing and is a valuable addition to the catalogue of bioanalytical techniques available to the forensic examiner. Though gaps associated with interpreting haploid sperm cells or Touch DNA originating from extra-cellular DNA will undoubtedly require additional research and development,

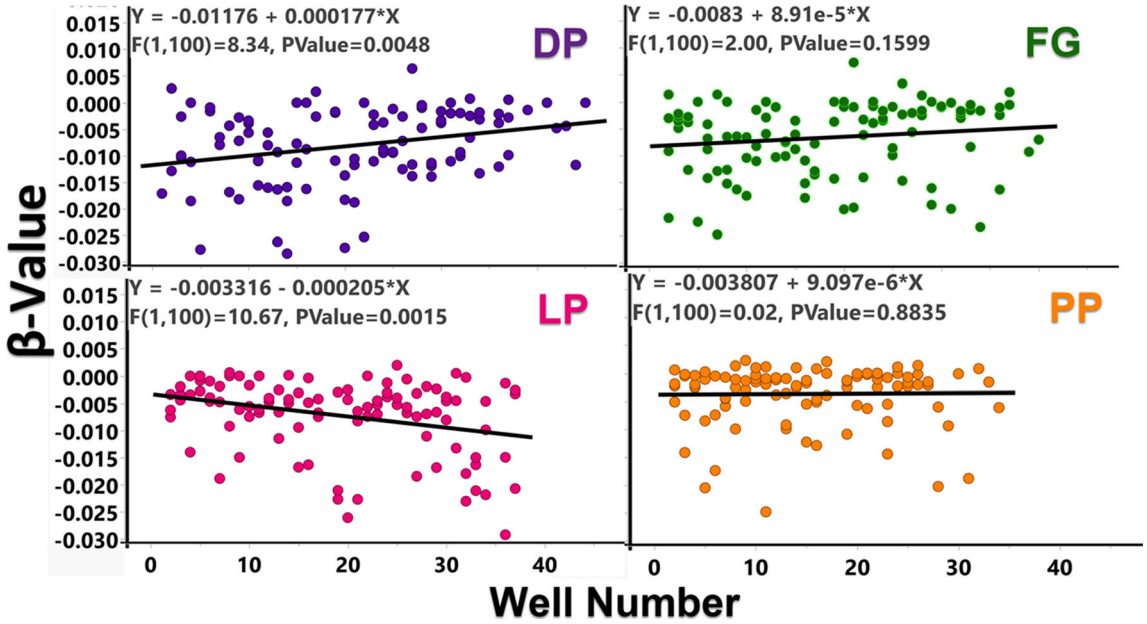

Fig. 5 Scatterplots of the β value versus well number, where the smallest well number represents the cell that remained in the extraction buffer the longest

cell sequestration at the front end continues to show great promise in solving the forensic DNA mixture conundrum.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00414-021-02503-4.

Funding This work was partially supported by NIJ2018-DU-BX-K0185 and NIJ2014-DN-BX-K026 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Disclaimer The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not reflect those of the Department of Justice.

References

1. Taylor D, Bright JA, Buckleton J (2013) The interpretation of single source and mixed DNA profiles. Forensic Sci Int Genet 7(5): 516–528. https://doi.org/10.1016/j.fsigen.2013.05.011
2. Swaminathan H, Garg A, Grgicak CM, Medard M, Lun DS (2016) CEESIt: a computational tool for the interpretation of STR mixtures. Forensic Sci Int Genet 22:149–160. https://doi.org/10.1016/j.fsigen.2016.02.005
3. Manabe S, Morimoto C, Hamano Y, Fujimoto S, Tamaki K (2017) Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model. PLoS One 12(11):e0188183. https://doi.org/10.1371/journal.pone.0188183
4. Bleka O, 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. https://doi.org/10.1016/j.fsigen.2015.11.008
5. Grgicak CM, Karkar S, Yearwood-Garcia X, Alfonse LE, Duffy KR, Lun DS (2020) A large-scale validation of NOCIt's a posteriori probability of the number of contributors and its integration into forensic interpretation pipelines. Forensic Sci Int Genet 47:14. https://doi.org/10.1016/j.fsigen.2020.102296
6. Slooten K, Caliebe A (2018) Contributors are a nuisance (parameter) for DNA mixture evidence evaluation. Forensic Sci Int Genet 37:116–125. https://doi.org/10.1016/j.fsigen.2018.05.004
7. Perlin MW, Legler MM, Spencer CE, Smith JL, Allan WP, Belrose JL, Duceman BW (2011) Validating TrueAllele® DNA mixture interpretation. J Forensic Sci 56(6):1430–1447. https://doi.org/10.1111/j.1556-4029.2011.01859.x
8. Lynch PC, Cotton RW (2018) Determination of the possible number of genotypes which can contribute to DNA mixtures: non-computer assisted deconvolution should not be attempted for greater than two person mixtures. Forensic Sci Int Genet 37:235–240. https://doi.org/10.1016/j.fsigen.2018.09.002
9. Peters KC, Swaminathan H, Sheehan J, Duffy KR, Lun DS, Grgicak CM (2017) Production of high-fidelity electropherograms results in improved and consistent DNA interpretation: standardizing the forensic validation process. Forensic Sci Int Genet 31:160–170. https://doi.org/10.1016/j.fsigen.2017.09.005
10. Taylor D, Buckleton J, Evett I (2015) Testing likelihood ratios produced from complex DNA profiles. Forensic Sci Int Genet 16: 165–171. https://doi.org/10.1016/j.fsigen.2015.01.008
11. Barrio PA, Crespillo M, Luque JA, Aler M, Baeza-Richer C, Baldassarri L, Carnevali E, Coufalova P, Flores I, Garcia O, Garcia MA, Gonzalez R, Hernandez A, Ingles V, Luque GM, Mosquera-Miguel A, Pedrosa S, Pontes ML, Porto MJ, Posada Y, Ramella MI, Ribeiro T, Riego E, Sala A, Saragoni VG, Serrano A, Vannelli S (2018) GHEP-ISFG collaborative exercise on mixture profiles (GHEP-MIX06). Reporting conclusions: results and evaluation. Forensic Sci Int Genet 35:156–163. https://doi.org/10.1016/j.fsigen.2018.05.005
12. Toscanini U, Gusmao L, Narvaez MCA, Alvarez C, Baldassarri L, Barbaro A, Berardi G, Hernandez EB, Camargo M, Carreras-Carbonell J, Castro J, Costa SC, Coufalova P, Dominguez V, de Carvalho EF, Ferreira STG, Furfuro S, Garcia O, Goios A,

Gonzalez R, de la Vega AG, Gorostiza A, Hernandez A, Moreno SJ, Lareu MV, Almagro AL, Marino M, Martinez G, Miozzo MC, Modesti NM, Onofri V, Pagano S, Arias BP, Pedrosa S, Penacino GA, Pontes ML, Porto MJ, Puente-Prieto J, Perez RR, Ribeiro T, Cardozo BR, Lesmes YMR, Satiage ASB, Saragoni VG, Serrano A, Streitenberger E, Morales MAT, Rey SAV, Miranda MV, Whittle MR, Fernandez K, Salas A (2016) Analysis of uni and bi-parental markers in mixture samples: lessons from the 22nd GHEP-ISFG Intercomparison Exercise. Forensic Sci Int Genet 25:63–72. https://doi.org/10.1016/j.fsigen.2016.07.010

13. Crespillo M, Barrio PA, Luque JA, Alves C, Aler M, Alessandrini F, Andrade L, Barretto RM, Bofarull A, Costa S, Garcia MA, Garcia O, Gaviria A, Gladys A, Gorostiza A, Hernandez A, Herrera M, Hombreiro L, Ibarra AA, Jimenez MJ, Luque GM, Madero P, Martinez-Jarreta B, Mascioveccchio MV, Modesti NM, Moreno F, Pagano S, Pedrosa S, Plaza G, Prat E, Puente J, Rendo F, Ribeiro T, Sala A, Santamaria E, Saragoni VG, Whittle MR (2014) EGHEP-ISFG collaborative exercise on mixture profiles of autosomal STRs (GHEP-MIX01, GHEP-MIX02 and GHEP-MIX03): results and evaluation. Forensic Sci Int Genet 10:64–72. https://doi.org/10.1016/j.fsigen.2014.01.009
14. Butler JM, Kline MC, Coble MD (2018) NIST interlaboratory studies involving DNA mixtures (MIX05 and MIX13): variation observed and lessons learned. Forensic Sci Int Genet 37:81–94. https://doi.org/10.1016/j.fsigen.2018.07.024
15. Binder V, Bartenhagen C, Okpanyi V, Gombert M, Moehlendick B, Behrens B, Klein HU, Rieder H, Krell PFI, Dugas M, Stoecklein NH, Borkhardt A (2014) A new workflow for whole-genome sequencing of single human cells. Hum Mutat 35(10):1260–1270. https://doi.org/10.1002/humu.22625
16. Lu X, Huang WH, Wang ZL, Cheng HK (2004) Recent developments in single-cell analysis. Anal Chim Acta 510(2):127–138. https://doi.org/10.1016/j.aca.2004.01.014
17. Peeters DJE, De Laere B, Van den Eynden GG, Van Laere SJ, Rothe F, Ignatiadis M, Sieuwerts AM, Lambrechts D, Rutten A, van Dam PA, Pauwels P, Peeters M, Vermeulen PB, Dirix LY (2013) Semiautomated isolation and molecular characterisation of single or highly purified tumour cells from CellSearch enriched blood samples using dielectrophoretic cell sorting. Br J Cancer 108(6):1358–1367. https://doi.org/10.1038/bjc.2013.92
18. Sinha N, Subedi N, Tel J (2018) Integrating immunology and microfluidics for single immune cell analysis. Front Immunol 9:16. https://doi.org/10.3389/fimmu.2018.02373
19. Rowan KE, Wellner GA, Grgicak CM (2016) Exploring the impacts of ordinary laboratory alterations during forensic DNA processing on peak height variation, thresholds, and probability of dropout. J Forensic Sci 61(1):177–185. https://doi.org/10.1111/1556-4029.12899
20. Duffy KR, Gurram N, Peters KC, Wellner G, Grgicak CM (2017) Exploring STR signal in the single- and multicopy number regimes: deductions from an in silico model of the entire DNA laboratory process. Electrophoresis 38(6):855–868. https://doi.org/10.1002/elps.201600385
21. Weusten J, Herbergs J (2012) A stochastic model of the processes in PCR based amplification of STR DNA in forensic applications. Forensic Sci Int Genet 6(1):17–25. https://doi.org/10.1016/j.fsigen.2011.01.003
22. Gill P, Curran J, Elliot K (2005) A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. Nucleic Acids Res 33(2):632–643. https://doi.org/10.1093/nar/gki205
23. Klein SB, Buoncristiani MR (2017) Evaluating the efficacy of DNA differential extraction methods for sexual assault evidence. Forensic Sci Int Genet 29:109–117. https://doi.org/10.1016/j.fsigen.2017.03.021
24. Ottens R, Templeton J, Paradiso V, Taylor D, Abarno D, Linacre A (2013) Application of direct PCR in forensic casework. Forensic Sci Int Genet Suppl Ser 4(1):e47–e48. https://doi.org/10.1016/j.fsigss.2013.10.024
25. van Oorschot RAH, Phelan DG, Furlong S, Scarfo GM, Holding NL, Cummins MJ (2003) Are you collecting all the available DNA from touched objects? Progr Forensic Genet 9(1239):803–807. https://doi.org/10.1016/s0531-5131(02)00498-3
26. Vandenberg N, vanOorschot RAH, Mitchell RJ (1997) An evaluation of selected DNA extraction strategies for short tandem repeat typing. Electrophoresis 18(9):1624–1626. https://doi.org/10.1002/elps.1150180924
27. Ip SC, Lin SW, Lai KM (2015) An evaluation of the performance of five extraction methods: Chelex® 100, QIAamp® DNA Blood Mini Kit, QIAamp® DNA Investigator Kit, QIAsymphony® DNA Investigator® Kit and DNA IQ™. Sci Justice 55(3):200–208. https://doi.org/10.1016/j.scijus.2015.01.005
28. QIAamp® DNA Investigator Handbook (2020) QIAGEN®
29. BioRad Chelex$^{(R)}$ 100 and Chelex 20 Chelating Ion Exchange Resin Instruction Manual
30. GlobalFiler™ PCR Amplification Kit (2015) Thermo Fisher Scientific Inc., version 4477604 Rev F
31. PowerPlex® Fusion System for Use on the Applied Biosystems® Genetic Analyzers (2020) Promega Corporation, version TMD045
32. Findlay I, Taylor A, Quirke P, Frazier R, Urquhart A (1997) DNA fingerprinting from single cells. Nature 389(6651):555–556. https://doi.org/10.1038/39225
33. Sanders CT, Sanchez N, Ballantyne J, Peterson DA (2006) Laser microdissection separation of pure spermatozoa from epithelial cells for short tandem repeat analysis. J Forensic Sci 51(4):748–757. https://doi.org/10.1111/j.1556-4029.2006.00180.x
34. Feng L, C-x L, Han J-p XC, Hu L (2015) Isolating cells from female/male blood mixtures using florescence in situ hybridization combined with low volume PCR and its application in forensic science. Int J Legal Med 129(6):1211–1215. https://doi.org/10.1007/s00414-014-1103-0
35. Williamson VR, Laris TM, Romano R, Marciano MA (2018) Enhanced DNA mixture deconvolution of sexual offense samples using the DEPArray (TM) system. Forensic Sci Int Genet 34:265–276. https://doi.org/10.1016/j.fsigen.2018.03.001
36. Fontana F, Rapone C, Bregola G, Aversa R, de Meo A, Signorini G, Sergio M, Ferrarini A, Lanzellotto R, Medoro G, Giorgini G, Manaresi N, Berti A (2017) Isolation and genetic analysis of pure cells from forensic biological mixtures: the precision of a digital approach. Forensic Sci Int Genet 29:225–241. https://doi.org/10.1016/j.fsigen.2017.04.023
37. Anslinger K, Graw M, Bayer B (2019) Deconvolution of blood-blood mixtures using DEPArray(TM) separated single cell STR profiling. Rechtsmedizin 29(1):30–40. https://doi.org/10.1007/s00194-018-0291-1
38. Hedell R, Hedman J, Mostad P (2018) Determining the optimal forensic DNA analysis procedure following investigation of sample quality. Int J Legal Med 132(4):955–966. https://doi.org/10.1007/s00414-017-1635-1
39. Hansson O, Egeland T, Gill P (2017) Characterization of degradation and heterozygote balance by simulation of the forensic DNA analysis process. Int J Legal Med 131(2):303–317. https://doi.org/10.1007/s00414-016-1453-x
40. ARCTURUS® PicoPure® DNA Extraction Kit (2010) Life Technologies Corporation, version 12637–00 Rev D
41. DEPArray™ Forensic Sample Prep Kit User Manual (2019) Menarini Silicon Biosystems S.p.A, version IFU_1005 Rev.4–2019 December
42. DirectPCR Lysis Reagent (Cell) Protocol (2007) Viagen Biotech, Inc., https://www.viagenbiotech.com/PDF/Protocol%2D%2D-Cell-2007_Feb.pdf. Accessed 12/07/2020

43. ZyGem Quick-Start Guide. DNA Extraction Using forensicGEM Saliva
44. Mönich UJ, Duffy K, Médard M, Cadambe V, Alfonse LE, Grgicak C (2015) Probabilistic characterisation of baseline noise in STR profiles. Forensic Sci Int Genet 19:107–122. https://doi.org/10.1016/j.fsigen.2015.07.001
45. Alfonse LE, Garrett AD, Lun DS, Duffy KR, Grgicak CM (2018) A large-scale dataset of single and mixed-source short tandem repeat profiles to inform human identification strategies: PROVEDIt. Forensic Sci Int Genet 32:62–70. https://doi.org/10.1016/j.fsigen.2017.10.006
46. Good P (2000) Permutation tests: a practical guide to resampling methods for testing hypotheses. https://doi.org/10.1007/978-1-4757-3235-1
47. Funes-Huacca ME, Opel K, Thompson R, McCord BR (2011) A comparison of the effects of PCR inhibition in quantitative PCR and forensic STR analysis. Electrophoresis 32(9):1084–1089. https://doi.org/10.1002/elps.201000584
48. Bright JA, Taylor D, Curran JM, Buckleton JS (2013) Degradation of forensic DNA profiles. Aust J Forensic Sci 45(4):445–449. https://doi.org/10.1080/00450618.2013.772235
49. Geng T, Novak R, Mathies RA (2014) Single-cell forensic short tandem repeat typing within microfluidic droplets. Anal Chem 86(1):703–712. https://doi.org/10.1021/ac403137h
50. Karkar S, Alfonse LE, Grgicak CM, Lun DS (2019) Statistical modeling of STR capillary electrophoresis signal. BMC Bioinformatics 20(16):584. https://doi.org/10.1186/s12859-019-3074-0
51. Goor RM, Forman Neall L, Hoffman D, Sherry ST (2011) A mathematical approach to the analysis of multiplex DNA profiles. Bull Math Biol 73(8):1909–1931. https://doi.org/10.1007/s11538-010-9598-0
52. Bright JA, Taylor D, Curran JM, Buckleton JS (2013) Developing allelic and stutter peak height models for a continuous method of DNA interpretation. Forensic Sci Int Genet 7(2):296–304. https://doi.org/10.1016/j.fsigen.2012.11.013
53. Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55(1):25–33. https://doi.org/10.1111/j.1556-4029.2009.01245.x
54. Hall A, Sims LM, Ballantyne J (2014) Assessment of DNA damage induced by terrestrial UV irradiation of dried bloodstains: forensic implications. Forensic Sci Int Genet 8(1):24–32. https://doi.org/10.1016/j.fsigen.2013.06.010