Title	**Population Genetic Issues for Forensic DNA Profiles.**
Award Number	NIJ 2017-DN-BX-0136
Draft Final Report	January 1, 2017 - December 31, 2020
Submitting Official	Carol Rhodes, Director, Office of Sponsored Programs osp@uw.edu (206) 543-4043
Recipient Organization	Office of Sponsored Programs University of Washington 4333 Brooklyn Avenue NE Seattle, WA 09195-9472

Submitted by

Lester Villaflor, Grant Analyst
Office of Sponsored Programs
(University of Washington)

Accomplishments

Major Goals of the Project

This project was planned to conduct research on the population genetic issues affecting the interpretation of forensic DNA profiles. The particular topics addressed were:

- Interpretation of forensic evidence.
- Interpretation of DNA sequence data.
- Interpretation of lineage marker evidence.
- Interpretation of DNA mixtures.
- Probabilistic genotyping.
- Other topics.

During the award period we published 36 papers addressing these topics, and they are summarized in the next section.

Summaries of Publications

Interpretation of Forensic Evidence.

Response to Lund and Iyer Lund and Iyer (2017) raised an argument regarding the use of likelihood ratios in court. In our view, their argument is based on a lack of understanding of the paradigm. Lund and Iyer argue that the decision maker should not accept the expert's likelihood ratio without further consideration. This is agreed by all parties. In normal practice, there is often considerable and proper exploration in court of the basis for any probabilistic statement. We concluded that Lund and Iyer argue against a practice that does not exist and which no one advocates. Further we conclude that the most informative summary of evidential weight is the likelihood ratio. We stated that this is the summary that should be presented to a court in every scientific assessment of evidential weight with supporting information about how it was constructed and on what it was based. [22. Gittelson et al., 2018]

A history of DNA profiling. The advent of DNA profiling in the 1980s has revolutionized forensic science. Forensic DNA profiling is a powerful tool that is used to both exonerate and implicate persons of interest in criminal cases. The technologies used to recover and detect DNA from crime scene stains have evolved over time. Whereas 30 years ago most forensic profiles were generated from visible stains such as blood or semen, nowadays, DNA profiles are often generated from trace amounts of DNA including touched items. The DNA in these profiles is often compromised in quantity and quality, and frequently originates from more than one individual. This improved sensitivity has complicated the interpretation of forensic DNA profiles, and consequently improved methods of profile interpretation have had to be developed. The adoption of these methods has meant that forensic scientists can now interpret mixed DNA profiles previously thought to be too complicated. We reviewed the history of forensic DNA profiling globally, introduced the challenges

of DNA profile interpretation, discussed why the community is adopting improved methods, and reflected on what new technology is on the horizon and the new challenges that may present. [8. Bright, Kelly et al., 2019]

Review of Likelihood Ratios We addressed criticisms by Stiffelman (2019) of the application of likelihood ratios (LRs) in forensic science, in particular their use in probabilistic genotyping (PG) software:— LRs do not infringe on the ultimate issue. The Bayesian paradigm clearly separates the role of the scientist from that of the decision makers and distances the scientist from comment on the ultimate and subsidiary issues.
LRs do not affect the reasonable doubt standard. Fact finders must still make decisions based on all the evidence and they must do this considering all evidence, not just that given probabilistically.
LRs do not infringe on the presumption of innocence. The presumption of innocence does not equate with a prior probability of zero but simply that the person of interest (POI) is no more likely than anyone else to be the donor.
Propositions need to be exhaustive within the context of the case. That is, propositions deemed relevant by either defense or prosecution which are not fanciful must not be omitted from consideration. [17. Buckleton, Robertson et al., 2020]

Low Likelihood Ratios To answer the question "Are low likelihood ratios reliable?" requires both a definition of reliable and then a test of whether low likelihood ratios (LRs) meet that definition. We offered, from a purely statistical standpoint, that reliability can be determined by assessing whether the rate of inclusionary support for non-donors over many cases is not larger than expected from the LR value. Thus, it is not the magnitude of the LR alone that determines reliability. Turing's rule is used to inform the expected rate of non-donor inclusionary support, where the rate of non-donor inclusionary support is at most the reciprocal of the LRJ i.e. $\Pr(\mathrm{LR} > x|H_a) < l/x$. There are parallel concerns about whether the value of the evidence can be communicated. We used a mixture of real and simulated data to show that the rate of non-donor inclusionary support for these data is significantly lower than the upper bound given by Turing's rule. We take this as strong evidence that low LRs are reliable. [16. Buckleyon, Pugh et al., 2020]

Interpreting DNA Sequence Data.

NGS Population Structure Match probabilities calculated during the evaluation of DNA evidence profiles rely on appropriate values of the population structure quantity F_{ST}. NGS-based methods will enhance forensic identification and with the transformation to such methods comes the need to facilitate NGS-based population genetics analysis. If NGS data are to be used for match probabilities there needs to be a way to accommodate population structure, which requires values for F_{ST} for those data. Such estimates have not been available. We assessed population structure for sequence-based data using a relatively new approach applied to STR data over 27 loci in five different geographic groups. Matching proportions between individuals or groups were used to obtain locus-specific F_{ST} estimates as well as estimates per geographic group and a global measure. The results demonstrated similar effects of sequencing data on F_{ST} estimates compared to what has been seen for CE-based results. [1. Aalbers and Weir, 2020]

Additivity of Read Counts. There has been an increase in the number of laboratories and researchers adopting new sequencing technologies, Known as nest-generation sequencing (NGS). An understanding of the behavior of NGS DNA profiles is needed to enable the development of probabilistic genotyping for the interpretation of such profiles. We investigated NGS analyte signal variation, specifically heterozygous balance and stutter variability from profiles generated using the ForenSeq^{TM}DNA Signature Prep Kit, DNA Primr Mix B. We also investigated additivity of analyte signals in NGS profiles for overlapping allelic and stutter signals originating from the same or different contributors. We have developed models that can be use to inform a continuous method for the interpretation of DNA profiling data. [20. Chenk, Skillman et al., 2020]

Interpreting Lineage Marker Evidence.

Y-haplotype Frequencies Estimating Y haplotype population frequencies is a demanding task in forensic genetics. Despite the suggestion of various methods, none these have yet reached a level of accuracy and precision that is acceptable to the forensic genetics community. At the basis of this problem is the complex dependency structure between the involved STR loci. We approximated this structure by the use of specific graphical models, namely t-cherry junction trees. We applied trees of order three by which dependencies between three STR loci can be taken into account, thereby extending the Chow-Liu method which is restricted to pairwise dependencies. We showed that the t-cherry tree method outperforms the Chow-Liu method as well as the well-established discrete Laplace method in estimation accuracy. [2. Andersen et al., 2020]

Y-STR Bayesian Networks We reconsidered the problem of population frequency estimation for Y-STR profiles by application of Bayesian networks and the Chow-Liu algorithm to model dependencies between loci. We found that the method based on the Chow-Liu algorithm performs almost as well as the discrete Laplace method. We have also made comparisons to the independence model and we have demonstrated once again that assuming independence of Y-STR loci cannot be supported. [3. Andersen et al., 2018]

We corrected an algebraic expression in an earlier paper [3]. The error in the original paper did not affect the numerical results shown there. [4. Andersen et al., 2019]

Y-STR Likelihood Ratios We introduced a new likelihood ratio method for evaluating mixed Y-STR profiles that is based on the premise that, given a haplotype has been seen in the person of interest, the most likely source of a second haplotype, matching at all or most loci, is in an individual with a recent common ancestor. We have called the new method the "Haplotype centered" (HC) method for likelihood ratio derivation. For single source, unambiguous haplotypes the HC method performs identically with the Kappa method proposed by Brenner (2012).

When attention is turned to mixtures we are required to assign a probability to many haplotypes seen neither in the database nor any person typed in the case. We derived a likelihood ratio formula in a way that allows a locus by locus approach. We demonstrated the application of the HC method on a series of Y-STR mixtures, originating from two to four individuals, in a manner that is still calculated locus-by-locus in nature. [33. Taylor et al., 2018]

Interpretation of DNA Mixtures.

Mixtures with unknown number of minor contributors. Modern interpretation strategies typically require an assignment of the number of contributors (NOC) to a DNA profile. This can prove to be a difficult task, particularly when dealing with higher order mixtures or mixtures where one or more contributors have donated low amounts of DNA. Differences in the assigned NOC at interpretation can lead to differences in the likelihood ration (LR). If the number of contributors cannot reasonably be assigned, then an interpretation of the profile may not be able to be progressed.

We investigated mixed DNA profiles of varying complexity and interpreted them altering the assigned NOC. We assigned LRs for true- and non-contributors and compared the results given different assignments of NOC over a range of mixture proportions. When a component of a mixture had a proportion of at least 10%, a ratio of at least 1.5:1 to the next highest component, and a DNA amount (as determined by STRmix^{TM}) of at least 50 rfu, the LR of the component for a true contributor was not significantly affected by varying NOC and was therefore suitable for interpretation and the assignment of an LR. LRs produced for minor contributors were found to vary significantly as the assigned NOC was changed. These heuristics may be used to identify profiles suitable for interpretation. [5. Bille et al., 2019]

Mixture to mixture matching. Standard practice in forensic science is to compare the reference DNA profile for a person of interest (POI) with an evidence DNA profile and calculate a likelihood ratio that considers propositions including and excluding the POI as a DNA donor. A method has recently been published that provides the ability to compare two evidence profiles (of any number of contributors and of any level of resolution) comparing propositions that consider the profiles either have a common contributor, or do not have any common contributors. Using this method, forensic analysts can provide intelligence to law enforcement by linking crime scenes when no suspects may be available. The method could also be used as a quality assurance measure to identify potential sample to sample contamination. We analyzed a number of constructed mixtures, ranging from two to five contributors, and with known numbers of common contributors, in order to investigate the performance of using likelihood ratios for mixture to mixture comparisons. Our findings demonstrated the ability to identify common donors in DNA mixtures with the power of discrimination depending largely on the least informative mixture of the pair being considered. The ability to match mixtures to mixtures may provide intelligence information to investigators by identifying possible links between cases which otherwise may not have been considered connected. [10. Bright et al., 2019]

Mixtures with varying numbers of contributors. Using a simplified model, we examined the effect of varying the number of contributors in the prosecution and alternate propositions for a number of simulated examples.

We compared the Slooten and Caliebe (2018) solution, with several existing practices. Our own experience is that most laboratories, and ourselves, assign the number of contributors, NOC, by allele count and a manual examination of peak heights. The LR for one or a very few values is calculated and typically one of these is presented, usually the most conservative. This gives an acceptable approximation.

Reassessing the number of contributors if LR = 0 and adding one to N under both Hp and Ha to "fit" the POI may lead to a substantial overstatement of the LR. A more reasonable option is

to allow optimization of the assignment under Hp and Ha separately. We show that an additional contributor explained the single locus profile better when PHR ≥ 0.51. This is pleasingly in line with current interpretation approaches.

Collectively these trials, and the solid theoretical development, suggest that the Slooten and Caliebe approach performs well. [12. Buckleton, Bright, Cheng et al., 2019]

We also considered uncertainty in the assignment of the number of contributors in higher order mixtures, where alleles shared between contributors may have dropped out or may be masked by the stutter artifacts or allelic peaks of a more dominant contributor. Most probabilistic genotyping software requires the assignment of NOC prior to interpretation. We explored the performance of this variable number of contributors (varNOC) method programmed within the probabilistic genotyping software STRmix^{TM}. The desired combination of performance and runtime was obtained using the default STRmix^{TM}version 2.7 MCMC settings in conjunction with a 2.5 % hyper-rectangle range, at least 10,000 naive MC iterations and 8 MCMC chains. The varNOC LR demonstrated the typical sensitivity and specificity behavior seen in previous studies, with a high level of reproducibility given repeat analyses. Profiles previously demonstrating ambiguity in the NOC assigned using conventional estimation methods, were able to be reliably interpreted and a varNOC LR assigned. [29. McGovern et al., 2020]

Probative value of mixed DNA profiles. Probabilistic genotyping typically proceeds by first deconvoluting a mixture into separate components and then computing a likelihood ratio for a potential contributor. The typical range of likelihood ratios for contributors and unrelated profiles depends, to a large extent, on how well the mixture is resolved. This in turn depends on the quality and complexity of the sample. Some samples are highly discriminatory while others are likely to yield only inconclusive results. Knowing the power of discrimination is helpful when deciding on whether or not to proceed with investigating a case or to do potential rework and for deciding on database inclusion.

We presented a method for exploring the range of likelihood ratios expected under a proposition of choice. We described the results of several case scenarios exploring questions such as: "Is the profile's minor contributor likely to yield a large likelihood ratio?", "Is it feasible to find a donor among a large database of unrelated profiles?" or "What is the probability that a brother of a true donor yields a large likelihood ratio?" We demonstrated the efficiency and speed of the method for a number of examples. [25. Kruijver et al., 2019]

Calibrating likelihood ratios Ramos and Gonzalez-Rodriguez (2013) introduced the concept of calibration in order to determine whether a system of evidence presentation is a reliable assessor of evidential weight. We applied this calibration method to a dataset of mixed forensic DNA profiles generated using the QIAGEN Investigator (R) 24plex QS Kit and interpreted using the probabilistic genotyping software STRmix^{TM}.We described the methodology for applying calibration to sets of forensic DNA profiles. We observed an approximate correspondence of the posterior probability, as assigned from the LR and the prior odds, with the observed rate of true donors for this dataset. [7. Bright, Jones Dukes et al., 2019]

Simulation studies of mixture deconvolution. If an unambiguous single source DNA profile is obtained from a crime scene, then a potential person of interest can either match or not match

the crime scene profile and the likelihood ratio for the single matching genotype can be easily computed. Mixed DNA profiles on the other hand are typically ambiguous and a vast number of different likelihood ratios can be obtained depending on the genotype of a potential person of interest that is compared with the mixture later. In the absence of a person of interest it can be unclear how suitable the profile is for discriminating between donors and non-donors. We introduced a simulation method to explore the range of likelihood ratios that is expected to be obtained when a non-donor or a donor is compared with the mixed DNA profile. Sampling is conditional on the mixture deconvolution obtained using probabilistic genotyping. These simulations help to decide whether or not a (mixed) profile is suitable for comparison to a person of interest. Moreover, the methods can be used to determine whether a profile is suitable for upload to a database and whether or not potential rework could be advised. [26. Kruijver et al., 2019]

Two-person mixture likelihood ratios. We analyzed 102 two-person DNA samples from simulated mixtures and male-female and male-male post-coital specimens and reported data on profile characteristics of these samples and likelihood ratios (LRs) generated using semi- and fully continuous systems. Both log(10) LRs from true and non-contributor tests are presented. These data may supplement studies comparing performance of different probabilistic systems for DNA evidence interpretation. [30. Rodriguez et al., 2019]

Use of Non-donor Distributions The reporting of a likelihood ratio (LR) calculated from probabilistic genotyping software has become more popular since 2015 and has allowed for the use of more complex mixtures at court. The meaning of "inconclusive" LRs and how to communicate the significance of low LRs at court is now important. We have developed a method using the distribution of LRs obtained from non-donors and studied. The non-donor distribution is useful for examining calibration and discrimination for profiles that have produced LRs less than about 10^4 and we constructed a range of mixed DNA profiles of varying quantity compared the LR distribution considering the minor contributor for a number of non-donors was compared to the expectation given a calibrated system. We demonstrated that conditioning genotypes should be used where reasonable given the background information to decrease the rate of non-donor LRs above 1. In all 17 cases examined, the LR for the minor donor was higher than the non-donor LRs, and in 12 of the 17 cases, the 99.9 percentile of the non-donor distribution was lower when appropriate conditioning information was used. The output of the tool is a graph that can show the position of the LR for the person of interest set against the non-donor LR distribution. This may assist communication between scientists and the court. [7. Schuerman et al., 2020]

Additivity of Peak Heights It is routinely assumed when interpreting forensic DNA profiles that peaks of the same molecular size, whether allelic or stutter in origin, "stack". That is, the height of a composite peak is approximately equal to the sum of its parts. There is strong theoretical reason to believe that this assumption should hold across the range of peak heights where fluorescent response is linear with respect to template. However, recent publications have called for empirical proof of, or directly questioned, this assumption. We examined the heights of allelic, stutter, and composite peaks, and demonstrated that peak heights are reliably predicted as the sum of their individual components. This work supports the long-held belief that peak heights stack in an additive fashion. [19. Cheng et al., 2020]

Probabilistic Genotyping

Response to PCAST Report To address concerns raised in the PCAST report (PCAST, 2016) we reported a large compilation of the internal validations of the probabilistic genotyping software STRmix^{TM}. Thirty one laboratories contributed data resulting in 2825 mixtures comprising three to six donors and a wide range of multiplex, equipment, mixture proportions and templates. Previously reported trends in the LR were confirmed including less discriminatory LRs occurring both for donors and non-donors at low template (for the donor in question) and at high contributor number. We were unable to isolate an effect of allelic sharing. Any apparent effect appears to be largely confounded with increased contributor number. [9. Bright et al., 2018]

Allelic Stuttering Modern probabilistic genotyping (PG) software is capable of modeling stutter as part of the profile weighting statistic. This allows for peaks in stutter positions to be considered as allelic or stutter or both. However, prior to running any sample through a PG calculator, the examiner must first interpret the sample, considering such things as artifacts and number of contributors (NOC).

Stutter can play a major role both during the assignment of the number of contributors, and the assessment of inclusion and exclusion. If stutter peaks are not filtered when they should be, it can lead to the assignment of an additional contributor, causing NOC contributors to be assigned as NOC+1. If peaks in the stutter position of a major contributor are filtered using a threshold that is too high, true alleles of minor contributors can be lost.

Until now, the software used to view the electropherogram stutter filters are based on a locus specific model. Combined stutter peaks occur when a peak could be the result of both back stutter (stutter one repeat shorter than the allele) and forward stutter (stutter one repeat unit larger than the allele). This can challenge existing filters.

We presented a novel stutter filter model in the ArmedXpert^{TM}software package that uses a linear model based on allele for back stutter and applies an additive filter for combined stutter. We term this the allele specific stutter model (AM).

We compared AM with a traditional model based on locus specific stutter filters (termed LM). This improved stutter model has the benefit of:
1. Increased detection of minor contributor alleles that are in a stutter position of a major contributor (we term the event of missing such alleles "over filtering"), and
2. Fewer false assignments of alleles (we term this event "under filtering").

Instances of over filtering were reduced 78% from 101 for a traditional model (LM) to 22 for the allele specific model (AM) when scored against each other. Instances of under filtering were reduced 80% from 85 (LM) to 17 (AM) when scored against ground truth mixtures. [23. Kalafut et al., 2018]

Robustness of STRMix STRmix^{TM}uses several laboratory specific parameters to calibrate the stochastic model for peak heights. These are modeled on empirical observations specific to the instruments and protocol used in the analysis. The extent to which these parameters can be borrowed from laboratories with similar technology and protocols without affecting the accuracy of the system was investigated using a sensitivity analysis. Parameters are first calibrated to a publicly available dataset, after which a large number of likelihood ratios are computed for true contributors and non-contributors using both the calibrated parameters and several borrowed parameters.

Differences in the LR caused by using different sets of parameter values were found to be negligible. [24. Kelly et al., 2018]

Interlaboratory Study for Mixture Analysis We reported the results of MIX13: an interlaboratory exercise directed by NIST in 2013. The goal of the exercise was to evaluate the general state of interpretation methods in use at the time across the forensic community within the US and Canada and to measure the consistency in mixture interpretation. The findings were that there was a large variation in analysts' interpretations between and within laboratories.

Within this work, we sought to evaluate the same mock mixture cases analyzed in MIX13 but with a more current view of the state-of-the-science. Each of the five cases were analyzed using the Identifiler^{TM}multiplex and interpreted with the combined probability of inclusion, CPI, and four different modern probabilistic genotyping systems. Cases 1-4 can be interpreted without difficulty by any of the four PG systems examined. Cases 1 and 4 could also be interpreted successfully with the CPI by assuming two donors.

Cases 2 and 3 cannot be interpreted successfully with the CPI because of potential of allele dropout. Case 3 demonstrated the need to consider relevant background information before interpretation of the profile. This case does not show that there is some barrier to interpretation caused by relatedness beyond the increased allelic overlap that can occur. Had this profile been of better template it might have been interpreted using the CPI despite the (potential) relatedness of contributors.

Case 5 suffers from over-engineering. It is unclear whether reference 5C, a non-donor, can be excluded by manual methods. Inclusion of reference 5C should be termed an adventitious match not a false inclusion. Beyond this statement this case does not contribute to the interlaboratory study of analyst/laboratory interpretation method performance, instead, it explores the limits of DNA analysis.

Taken collectively the analysis of these five cases demonstrates the benefits of changing from CPI to a PG system. [11. Buckleton et al., 2018]

Collaborative exercise on DNA mixture interpretation. An intra and inter-laboratory study using the probabilistic genotyping (PG) software STRmix^{TM}was conducted. Two complex mixtures from the PROVEDIt set, analyzed on an Applied Biosystems 3500 Series Genetic Analyzer^{TM}, were selected and 174 participants responded.

For Sample 1 (low template, in the order of 200 rfu for major contributors) five participants described the comparison as inconclusive with respect to the POI or excluded him. Where LRs were assigned, the point estimates ranging from 2×10^4 to 8×10^6. For Sample 2 (in the order of 2000 rfu for major contributors), LRs ranged from 2×10^{28} to 2×10^{29}. Where LRs were calculated, the differences between participants can be attributed to (from largest to smallest impact): varying number of contributors (NOC), the exclusion of some loci within the interpretation, differences in local CE data analysis methods leading to variation in the peaks present and their heights in the input files used, and run-to-run variation due to the random sampling inherent to all MCMC-based methods.

The study demonstrated a high level of repeatability and reproducibility among the participants. For those results that differed from the mode, the differences in LR were almost always minor or conservative. [6. Bright, Cheng et al., 2019]

We responded to a commentary on [6] by McNevin et al. (2019). those authors acknowledged that our earlier work was a "valuable first step in establishing a foundational validity for mixture interpretation using probabilistic software, as recommended by the PCAST report" and suggested further experiments to "better address this issue". Courts have a justifiable interest in whether the LR in any particular case may be trusted. In many US states the Judge's gatekeeper function is enshrined in either the Frye or Daubert standard. In our response we addressed the concerns of McNevin et al. [13. Buckleton, Bright et al., 2020]

Utility and validity of STRmix™ Forensic DNA interpretation is transitioning from manual interpretation based usually on binary decision-making toward computer-based systems that model the probability of the profile given different explanations for it, termed probabilistic genotyping (PG). Decision-making by laboratories to implement probability-based interpretation should be based on scientific principles for validity and information that supports its utility, such as criteria to support admissibility. The principles behind STRmix™include standard mathematics and modeling of peak heights and variability in those heights. All PG methods generate a likelihood ratio (LR) and require the formulation of propositions. Principles underpinning formulations of propositions include the identification of reasonably assumed contributors. Substantial data have been produced that support precision, error rate, and reliability of PG, and in particular, STRmix™. A current issue is access to the code and quality processes used while coding. There are substantial data that describe the performance, strengths, and limitations of STRmix™, one of the available PG software. [14. Buckleton, Bright, Gittelson et al., 2019]

Probabilistic genotyping software. The interpretation of mixed profiles from DNA evidentiary material is one of the more challenging duties of the forensic scientist. Traditionally, analysts have used a "binary" approach to interpretation where inferred genotypes are either included or excluded from the mixture using a stochastic threshold and other biological parameters such as heterozygote balance, mixture ratio, and stutter ratios. As the sensitivity of STR multiplexes and capillary electrophoresis instrumentation improved over the past 25 years, coupled with the change in the type of evidence being submitted for analysis (from high quality and quantity (often single-source) stains to low quality and quantity (often mixed) "touch" samples), the complexity of DNA profile interpretation has equally increased. We gave a historical perspective on the movement from binary methods of interpretation to probabilistic methods of interpretation. We described the two approaches to probabilistic genotyping (semi-continuous and fully continuous) and addressed issues such as validation and court acceptance. Areas of future needs for probabilistic software were discussed. [21. Coble and Bright, 2019]

Comparison of probabilistic genotyping methods. Alladio et al. (2018) have provided an important contribution to our ongoing understanding of PG interpretation landscape in forensic biology. They demonstrated the concordance between three continuous and two semi-continuous systems despite differences in the underlying models. The comparison of results from PG systems, however, must be based on systems that utilize the same data and function in a similar manner. To do otherwise will simply penalize the most discriminating system, by the performance of the least discriminating system. [32. Taylor et al., 2019]

Contamination detection. A recent publication has provided the ability to compare two mixed DNA profiles and consider their probability of occurrence if they do, compared to if they do not, have a common contributor. This ability has applications to both quality assurance (to test for sample to sample contamination) and for intelligence gathering purposes (did the same unknown offender donate DNA to multiple samples). We used a mixture-to-mixture comparison tool to investigate the prevalence of sample-to-sample contamination that could occur from two laboratory mechanisms, one during DNA extraction and one during electrophoresis. By carrying out pairwise comparisons of all samples (deconvoluted using probabilistic genotyping software STRmix^{TM}) within extraction or run batches we identify any potential common DNA donors and investigate these with respect to their risk of contamination from the two proposed mechanisms. While not identifying any contamination, we inadvertently found a potential intelligence link between samples, showing the use of a mixture to mixture comparison tool for investigative purposes. [34. Taylor, Rowe et al., 2019]

Additivity of stutter peaks. Peaks in an electropherogram could represent alleles, stutter product, or a combination of allele and stutter. Continuous probabilistic genotyping (PG) systems model the heights of peaks in an additive manner: for a shared or composite peak, PG models assume that the peak height is the sum of the allelic component and the stutter component. We examined the assumption that the heights of overlapping alleles from a minor contributor and stutter peaks from a major contributor are additive.

Any peak below the analytical threshold is considered unobserved; hence, in any dataset and particularly in low-template DNA profiles, some or many peaks may be unobserved or missing. Using simulation and empirical data, we show that an additive model can explain the heights of overlapping alleles from a minor contributor and stutter peaks from a major contributor as long as missing data are carefully considered. We use a naive method of imputation for the missing data which appears to perform adequately in this case. If missing data are ignored then the sum of stutter and allelic peaks is expected to be an overestimate of the average height of the composite peaks, as was observed in this study. [15. Buckleton, Lohmueller et al., 2019]

Mixed Profiles in Parentage Cases We reported the interpretation of three-person mixed DNA profiles constructed from DNA from one mother, father, and child trio using the probabilistic genotyping software STRmix^{TM}. A total of 40 mixtures were examined, with varying total template and mixture proportions of the three contributors. In addition, mixtures were artificially degraded at four different rates to test the effects of degradation on the interpretation of mother, father and child trios. A total of 560 STRmix^{TM}analyses were undertaken, examining four different interpretation strategies. Reasonable results were only achieved by conditioning on one parent as an assumed donor and applying a user-informed prior to the mixture proportion of both parents. [27,28. Lin et al., 2020a,b]

Other Topics

Transnational Crime Rapid growth in world trade has enabled transnational criminal networks to conceal their contraband among the one billion containers shipped worldwide annually. Forensic methods are needed to identify the major cartels moving the contraband into transit. In Wasser et al. (2018) we combined DNA-based sample matching and geographic assignment of tusks to show

that the two tusks from the same elephant are often shipped by the same trafficker in separate large consignments of ivory. The paired shipments occur close in time from the same initial place of export and have high overlap in the geographic origins of their tusks. Collectively, these paired shipments form a linked chain that reflects the sizes, interconnectedness, and places of operation of Africa's largest ivory smuggling cartels. [36. Wasser et al., 2018]

Laser Microdissection Laser microdissection (LMD) is a tool that is used in forensic laboratories for the analysis of DNA from specifically targeted cells. Since 2010, the Institute of Environmental Science and Research Limited's (ESR) Forensic Biology laboratory has applied LMD DNA testing to a variety of sexual assault casework samples where small numbers of sperm are present in cell mixtures. We reviewed the DNA profiling results obtained from semen-stained casework samples that have been analyzed using the LMD DNA methodology developed in our laboratory. Dissected sperm have been analyzed using the AmpFISTR Identifiler amplification kit at 28 cycle PCR, the AmpFISTR MiniFiler amplification kit at 30 cycle PCR, or the AmpFISTR SGM Plus amplification kit at 34 cycle PCR, depending on the number of sperm recovered and on consideration of other circumstantial case information, such as the time since intercourse (TSI). From a review of these data, success rates for different sample numbers of sperm recovered from semen-stained samples were determined. The DNA profiling results obtained from three cases where laser microdissection has been used were also presented to demonstrate the success of the LMD testing strategy in a forensic laboratory. [35. Vintiner et al., 2020]

References

- Alladio E, Omedei M, Cisana S, D'Amico G, Caneparo D, Vincenti M, Garofano P. 2918. DNA mixtures interpretation - A proof-of-concept multi-software comparison highlighting different probabilistic methods' performances on challenging samples. Forensic Science International: Genetics 37:143-150. DOI: 10.1016/j.fsigen.2018.08.002

- Brenner CH. 2010. Fundamental problem of forensic mathematics - the evidential value of a rare haplotype. Forensic Science International: Genetics 4: 281-191.

- Lund SP, Iyer H. 2017. Likelihood ratio as weight of forensic evidence: A closer look. J Res Natl Inst Stan 122:27. https://doi.org/10.6028/jres.122.027

- McNevin D, Wright K, Chaseling J, Barash M. 2019. Commentary on: Bright et al. (2018) Internal validation of STRmix^{TM}- a multi laboratory response to PCAST, Forensic Science International: Genetics 34: 1124. Forensic Science International: Genetics 41:E14-E17. DOI:https://doi.org/10.1016/j.fsigen.2019.03.01

- PCAST (President's Council of Advisors on Science and Techology). 2016. PCAST report on Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_

- Ramos D, Gonzalez-Rodriguez J. 2013. Reliable support: Measuring calibration of likelihood ratios. Forensic Science International 230:156-169. DOI: 10.1016/j.forsciint.2013.04.014

- Slooten K, Caliebe A. 2018. Contributors are a nuisance (parameter) for DNA mixture evidence evaluation. Forensic Science International: Genetics 37:116-125.

- Stiffelman B. 2019. No longer the gold standard: probabilistic genotyping is changing the nature of DNA evidence in criminal trials. Berkeley J. Crim. Law 24:110.

Dissemination of Results

The primary means of dissemination of results has been the publication of peer-reviewed papers, as listed below. In addition we made presentations at scientific conferences and taught courses, as now listed:

Scientific Presentations / Short Courses

- Aalbers S, Weir BS. 2019. Analyzing population structure for forensic STR markers in next generation sequencing data. International Symposium on Human Identification, Palm Springs, CA.
- Bright JA. 2019. Low-level Likelihood Ratios. In the Forensic Science Center of Excellence Webinars on Probabilistic Genotyping of Evidentiary DNA Typing.
- Buckleton JS. 2019. STRmix admissibility. In the Forensic Science Center of Excellence Webinars on Probabilistic Genotyping Evidentiary DNA Typing.
- Cheng K. 2020. Developing biological models for the probabilistic genotyping of next generation sequencing (NGS) data. American Academy of Forensic Sciences, Anaheim CA.
- Cheng K. 2019. Assessing Heterozygous Balance in NGS DNA Profiles. NZ Statistical Association Conference.
- Cheng K. 2020. Developing biological models for the probabilistic genotyping of NGS data. Department of Statistics, Auckland University, New Zealand.
- Weir BS. 2018, 2019, 2020. Forensic Genetics. University of Washington.
- Weir BS, Aalbers S. 2018, 2019, 2020. Forensic Genetics. Summer Institute in Statistical Genetics, University of Washington.
- Weir BS. 2018. Multi-locus match probabilities. NIJ Forensic Research and Development Symposium. Seattle, WA.
- Weir BS. 2019. Interpreting Y-STR forensic evidence. University of Washington.

Publications Acknowledging NIJ award 2017-DN-BX-0136

1. Aalbers S, Hipp MJ, Kennedy SR, Weir BS. 2020. Analyzing population structure for forensic STR markers in next generation sequencing data. Forensic Science International: Genetics 49:Article Number 102364. DOI: 10.1016/j.fsigen.2020.102364.
2. Andersen MM, Caliebe A, Kirkeby K, Knudsen M, Vihrs N, Curran JM. 2020. Estimation of Y haplotype frequencies with lower order dependencies. Forensic Science International: Genetics 46: Article Number: UNSP 102214. DOI: 10.1016/j.fsigen.2019.102214.
3. Andersen MM, Curran J, de Zoete J, Taylor D, Buckleton J. 2018. Modelling the dependence structure of Y-STR haplotypes using graphical models Forensic Science International: Genetics 37:29-36.

4. Andersen MM, Curran J, de Zoete J, Taylor D, Buckleton J. 2019. Modelling the dependence structure of Y-STR haplotypes using graphical models (vol 37, pg 29, 2018). Forensic Science International: Genetics 41:E3-E30.

5. Bille T, Weitz S, Buckleton JS, Bright JA. 2019. Interpreting a major component from a mixed DNA profile with an unknown number of minor contributors. Forensic Science International: Genetics 40:150-159.

6. Bright JA, Cheng K, Kerr Z, McGovern C, Kelly H, Moretti TR, Smith MA, Bieber FR, Budowle B, Coble MD, Alghafrr R, Allen PS, Barbee A, Beamer ABV, Buettner C, Russell M, Gehrig C, Hicks T, Charak J, Cheong-Wing K, Ciecko A, Davis CT, Donley M, Pedersen N, Gartside B, Granger D, Greer-Ritzheimer M, Reisinger E, Kennedy J, Grammer E, Kaplan M, Hansen D, Larsen HJ, Laureano A, Li C, Lien E, Lindberg E, Kelly C, Mallinder B, Malsom S, Yacovone-Margetts A, McWhorter A, Prajapati SM, Powell T, Shutler G, Stevenson K, Stonehous, AR, Smith L, Murakami J, Halsing E, Wright D, Clarks L, Taylor DA, Buckleton J. 2019. STRmix^{TM}collaborative exercise on DNA mixture interpretation. Forensic Science International: Genetics 40:1-8.

7. Bright JA, Jones Dukes M, Pugh SN, Evett IW, Buckleton JS. 2019. Applying calibration to LRs produced by a DNA interpretation software. Australian Journal of Forensic Sciences: Early access OCT 2019.

8. Bright JA, Kelly H, Kerr Z, McGovern C, Taylor D, Buckleton JS. 2019. The interpretation of forensic DNA profiles: an historical perspective. Journal of the Royal Society of New Zealand: early access NOV 2019.

9. Bright JA, Richards R, Kruijver M, Kelly H, McGovern C, Magee A, McWhorter A, Ciecko A, Peck B, Baumgartner C, Buettner C, McWilliams S, McKenna C, Gallacher C, Mallinder B, Wright D, Johnson D, Catella D, Lien E, O'Connor C, Duncan G, Bundy J, Echard J, Lowe J, Stewart J, Corrado K, Gentile S, Kaplan M, Hassler M, McDonald N, Hulme P, Oefelein RH, Montpetit S, Strong M, Noel S, Malsom S, Myers S, Welti S, Moretti T, McMahon T, Grill T, Kalafut T, Greer-Ritzheimer M, Beamer V, Taylor DA, Buckleton JS. 2018. Internal validation of STRmix^{TM}- A multi laboratory response to PCAST. Forensic Science International: Genetics 34:11-24.

10. Bright JA, Taylor D, Kerr Z, Buckleton J, Kruijver M. 2019. The efficacy of DNA mixture to mixture matching. Forensic Science International: Genetics 41:64-71.

11. Buckleton JS, Bright JA, Cheng K, Budowle B, Coble MD. 2018. NIST interlaboratory studies involving DNA mixtures (MIX13): A modern analysis. Forensic Science International: Genetics 37:172-179.

12. Buckleton JS, Bright JA, Cheng K, Kelly H, Taylor DA. 2019. The effect of varying the number of contributors in the prosecution and alternate propositions. Forensic Science International: Genetics 38:225-231.

13. Buckleton JS, Bright JA, Ciecko A, Kruijver M, Mallinder B, et al. 2020. Response to: Commentary on: Bright et al. (2018) Internal validation of SSTRmix^{TM}- A multi laboratory

response to PCAST, Forensic Science International: Genetics 34: 11-24. Forensic Science International: Genetics 44:Article Number: UNSP 102198. DOI: 10.1016/j.fsigen.2019.102198

14. Buckleton JS, Bright JA, Gittelson S, Moretti TR, Onorato AJ, Bieber FR, Budowle B, Taylor DA. 2019. The Probabilistic Genotyping software DSTRmix^{TM}: Utility and evidence for its validity. Journal of Forensic Sciences 64:393-405.

15. Buckleton JS, Lohmueller KE, Inman K, Cheng K, Curran JM, Pugh SN, Bright JA, Taylor DA. 2019. Testing whether stutter and low-level DNA peaks are additive. Forensic Science International: Genetics 43: Article 102166.

16. Buckleton JS, Pugh SN, Bright JA, Taylor DA, Curran JM, Kruijver M, Gill P, Budowle B, Cheng K. 2020. Are low LRs reliable? Forensic Science International: Genetics 49:Article 102350.DOI: 10.1016/j.fsigen.2020.102350.

17. Buckleton J, Robertson B, Curran J, Berger C, Taylor D, et al. 2020. A review of likelihood ratios in forensic science based on a critique of Stiffelman "No longer the Gold standard: Probabilistic genotyping is changing the nature of DNA evidence in criminal trials." Forensic Science International 310: Article Number: 110251.

18. Caliebe A, Andersen MM, Curran J. 2020. Reducing complex dependency structure by graphical models – with an application to Y-chromosomal haplotypes. Genetic Epidemiology 44:472-473.

19. Cheng K, Bright JA, Kerr Z, Taylor D, Ciecko A, et al. 2020. Examining the additivity of peak heights in forensic DNA profiles. Australian Journal of Forensic Sciences DOI: 10.1080/00450618.2019.1704060. Early Access: JAN 2020

20. Cheng K, Skillman J, Hickey S, Just R, Moreno L, Bright JA, Kelly H, Lin MH, Curran JM, Buckleton J. 2020. Variability and additivity of read counts for aSTRs in NGS DNA profiles. Forensic Science International: Genetics 48: Article Number 102351. DOI: 10.1016/j.fsigen.2020.102351.

21. Coble MD, Bright JA. 2019. Probabilistic genotyping software: An overview. Forensic Science International: Genetics 38:219-224.

22. Gittelson S, Berger CEH, Jackson G, Evett IW, Champod C, Robertson B, Curran JM, Taylor D, Weir BS, Coble MD, Buckleton JS. 2018. A response to "Likelihood ratio as weight of evidence: A closer look' by Lund and Iyer. Forensic Science International 288:E15-E19.

23. Kalafut T, Schuerman C, Sutton J, Faris T, Armogida L, Bright JA, Buckletonc J, Taylor D. 2018. Implementation and validation of an improved allele specific stutter filtering method for electropherogram interpretation. Forensic Science International: Genetics 35:50-56.

24. Kelly H, Bright JA, Kruijver M, Cooper S, Taylor D, (Strong M, Beamer V, Buettner C, Buckleton J. 2018. A sensitivity analysis to determine the robustness of STRmix (TM) with respect to laboratory calibration. Forensic Science International: Genetics 35:113-122.

25. Kruijver M, Bright JA, Kelly H, Buckleton J. 2019. Exploring the probative value of mixed DNA profiles. Forensic Science International: Genetics 41:1-10.

26. Kruijver M, Bright JA, Kelly H. 2019. Exploring the DNA mixture deconvolution through simulation. Australian Journal of Forensic Sciences 51:514-517.

27. Lin MH, Bright JA, Pugh SN, Buckleton JS. 2020a. The interpretation of mixed DNA profiles from a mother, father, and child trio. Forensic Science International: Genetics 44: Article Number: UNSP 102175. DOI: 10.1016/j.fsigen.2019.102175.

28. Lin MH, Bright JA, Pugh SN, Buckleton JS. 2020b. The interpretation of mixed DNA profiles from a mother, father, and child trio (vol 49, 102175, 2020). Forensic Science International: Genetics 49: Article Number: 102321. DOI: 10.1016/j.fsigen.2020.102321.

29. McGovern C, Cheng K, Kelly H, Ciecko A, Taylor D, Buckleton JS, Bright JA. 2020. Performance of a method for weighting a range in the number of contributors in probabilistic genotyping. Forensic Science International: Genetics 48: Article Number 102352. DOI: 10.1016/j.fsigen.2020.102352.

30. Rodriguez JJRB, Bright JA, Salvador JM, Laude RP, De Ungria MCA. 2019. Data on likelihood ratios of two-person DNA mixtures interpreted using semi- and fully continuous systems. Data in Brief 26: Article UNSP 104455.

31. Schuerman C, Kalafut T, Buchanan C, Sutton J, Bright JA. 2020. Using the nondonor distribution to improve communication and inform decision making for low LRs from minor contributors in mixed DNA profiles. Journal of Forensic Sciences 65:1072-1084. DOI: 10.1111/1556-4029.14306.

32. Taylor DA, Buckleton JS, Bright JA. 2019. Comment on "DNA mixtures interpretation - A proof-of-concept multi-software comparison highlighting different probabilistic methods' performances on challenging samples" by Alladio et al. Forensic Science International: Genetics 40:E248-E251.

33. Taylor D, Curran J, Buckleton J. 2018. Likelihood ratio development for mixed Y-STR profiles. Forensic Science International: Genetics 35:82-96.

34. Taylor D, Rowe E, Kruijver M, Abarno D, Bright JA, Buckleton J. 2019. Inter-sample contamination detection using mixture deconvolution comparison. Forensic Science International: Genetics 40:160-167.

35. Vintiner SK, Veth JS, Bright JA. 2020. A review of DNA profiling success for laser microdissected forensic casework samples. Australian Journal of Forensic Sciences 52:282-292. DOI: 10.1080/00450618.2018.1510030

36. Wasser SK, Torkelson A, Winters M, Horeaux Y, Tucker S, Otiende MY, Sitam FAT, Buckleton J, Weir BS. 2018. Combating transnational organized crime by linking multiple large ivory seizures to the same dealer. Science Advances 4(9): article eaat0625.