© natali_mis | AdobeStock

When DNA Is Not a Gold Standard: Failing to Interpret Mixture Evidence

Forensic science connects evidence through shared characteristics. Markings on a bullet can appear to match grooves in the barrel of a gun. Latent fingerprints left at a crime scene may be similar to ridge patterns on a suspect's hand. Tracks in the mud may mirror the treads of a shoe or tire. Police gather forensic evidence to help build a case, and police dramas on television convey the myth of forensic infallibility through the "CSI" effect.[1]

In 2009, the National Academy of Sciences (NAS) published its seminal report titled *Strengthening Forensic Science in the United States*.[2] The NAS report reviewed many forensic modalities and questioned their scientific validity. The interpretation of forensic data is often unreliable. Match statistics are needed to gauge the strength of match between items, relative to coincidence. But forensic statistics are typically absent or incorrect. Human bias can skew answers by unconsciously selecting favorable data, using knowledge about defendant characteristics, or by trying to please stakeholders who have a desired criminal justice outcome.

Deoxyribonucleic acid (DNA) evidence seems immune to such criticism, long serving as a gold standard for other forensic disciplines. Abundant DNA from one person produces pristine data signals. Interpreting these clear signals yields an unambiguous genetic type ("genotype"). Comparing definite genotypes, relative to a random person, yields a reliable match statistic that numerically conveys the probative force of DNA evidence. But most crime scene DNA is now a mixture of two or more people, with good data but less certain interpretation. As the NAS report noted, there may be problems with how the DNA was interpreted, such as when there are mixed samples.

Simplistic interpretation of DNA mixture data often fails to produce an accurate match statistic or give any answer at all. While the limitations and liabilities of unscientific DNA mixture interpretation were recognized early on,[3] only recently has this profound forensic failure come to the fore. Crime laboratories in Austin, Texas, and Washington, D.C., have been shuttered in large part because of failed DNA mixture interpretation.[4] Virginia re-evaluated DNA match statistics for mixture evidence in hundreds of cases.[5] Texas is reviewing 24,000 criminal cases for flawed interpretation of DNA mixture evidence.[6] The New York State Police (NYSP) has suppressed reliable DNA mixture interpretation methods that could expose its crime laboratory's mistakes in thousands of cases.[7] These numbers extrapolate to hundreds of thousands of mixture items throughout the United States, and the national press has taken notice.[8]

This failure of forensic DNA interpretation is of broad concern. Pervasive errors in DNA match statistics undermine public trust in science and erode confidence in government agencies that misuse science to obtain convictions. A failed DNA gold standard portends little hope for fledgling forensic fields. Perhaps the greatest loss is true justice in a free society. Misinterpreting DNA evidence causes injustice for defendants denied potentially exculpa-

BY MARK W. PERLIN, PH.D., M.D., PH.D.

tory evidence, injustice for victims whose cases are lost when inculpatory evidence is unreported, and injustice for innocents victimized by crime that DNA could have prevented.

This article reviews the history of failed DNA mixture interpretation. It begins in 1985, at the start of the genomics revolution, discussing the origins of modern DNA testing. Proceeding in five-year increments, it outlines the missed opportunities and policy failures that have resulted in the current situation. The article offers recommendations to help overcome long-standing DNA interpretation problems.

Biology

The human genome contains three billion DNA letters, a text written across 23 chromosomes in the nucleic acid alphabet A, C, G, and T. This textual information is used to operate, maintain, evolve, and grow human organisms. Part of the genome's power is the encoding of this biological operating system. Another aspect is the variation between people found in noncoding regions that scientists can use to trace ancestry, map disease, and distinguish between individuals.

Scattered throughout the human genome are genetic locations (loci) that have a short DNA word repeated in tandem. These short tandem repeats (STR) are a rich source of genetic variation. The number of repeated words at a locus varies between different people, and these STR length variants (alleles) can be used to identify individuals.

A cell nucleus has two complete genome copies of the 22 human autosomal chromosomes, one inherited from each parent. At a particular locus on a chromosome, there are two alleles – maternal and paternal. A person's pair of alleles at a genetic locus defines the person's genotype at that chromosome location.

An STR locus with many (for example, 15) allele variants yields very many genotype allele pair possibilities (for example, 100). Examining multiple independent STR loci multiplies those possibilities, allowing for a trillion trillion possible genotypes (24 powers of ten). Since there are fewer than 10 billion people alive today (10 powers of ten), there are far more STR genotypes than people, making DNA useful for identification.

1985: Revolution

Three technologies triggered the DNA revolution in the automated genetic analysis of minute biological samples. The first technology was polymerase chain reaction (PCR), which let scientists easily make millions of copies of small DNA quantities at a genetic locus.[9] The second technology was the automated DNA sequencer (now called a "genetic analyzer"), which used electrophoretic separation and laser detection to measure DNA fragment length and quantity.[10] Finally, cheap ubiquitous computing enabled automated analysis of genetic data.[11]

The STR genetic marker was an early beneficiary of this technological juxtaposition.[12] PCR amplification of an STR locus produced DNA fragments in detectable quantities. Separating fragments on a DNA sequencer showed data peaks, with longer alleles having greater length. Computer analysis of STR data could identify and size these peaks to indicate allele events and would eventually automate genotype determination.

1990: Threshold

The original STR genetic tests were done on DNA samples from a single source, not mixtures. The locus data had one or two tall peaks, corresponding to the one or two parental alleles in an individual's genotype. The testing was done for genetic diagnosis, genome mapping, and drug discovery.[13]

With simple single source data, the interpretation issue was separating the true alleles from background noise or data artifacts. This separation was accomplished by drawing a line that separated tall allele peaks from short non-allele peaks. A DNA sequencer manufacturer advised setting this threshold at around 100 relative fluorescent units (RFU). There was no statistical science involved, just a rule of thumb to help technicians interpret their allele data based on peak height.

1995: Variation

With abundant DNA from one person, and clean data signals, thresholds worked well to separate tall allele peaks (1000 to 2000 RFU) from baseline instrument noise (5 to 15 RFU). However, other data artifacts produced peaks over baseline, or subtler peak patterns.

PCR stutter is an error in the DNA copying mechanism.[14] When the polymerase enzyme copies a DNA region of STR text, it can lose its place and skip over one of the short repeated words. This deletion drops a repeated word, generating a DNA fragment one word shorter than the actual allele (e.g., a 10 allele with 10 repeated words can produce a fragment having only 9 repeated words). Such stutter alleles show around 5-15 percent of the true allele's peak height and reside adjacent to the allele peak. Stutter peaks can be identified and removed with single source DNA data, but complicate the interpretation of mixed or low-level DNA.

The original STR loci used in genetic testing had two letters in a repeated word.[15] These di-nucleotide repeats were popular with geneticists because their high genome density placed them near most genes. However, they gave complex stutter patterns with a long trail of fragments having from 5 to 10 dropped words. For that reason, forensic identification (which had to be explained to lay juries, and only needs a dozen loci) employed tetra-nucleotide repeats having four letters in a repeated word.[16] Their simpler stutter patterns usually show just one prominent stutter peak.

Automated computer analysis could mathematically separate stutter peaks from STR locus data.[17] Some genetic and forensic practitioners used this computerized approach,[18] but most technicians were more comfortable removing stutter visually.

Other random factors affect genotype data. These largely arise from the inherent random variation in PCR copying. Within a copying cycle, some DNA fragments will copy more efficiently than others. Given identical DNA input, this random copying process introduces variation in the data output, with each PCR experiment producing its own data pattern. This natural variation in DNA counting is well known to scientists, and it has been mathematically modeled.[19]

2000: Mixture

The Federal Bureau of Investigation (FBI) helps regulate forensic DNA analysis in the United States. The agency's Scientific Working Group on DNA Analysis Methods (SWGDAM) convenes twice a year to discuss policies of interest to the FBI laboratory. SWGDAM members are forensic practitioners, mainly government employees of crime laboratories or police organizations. They are not experts in modern statistical computing and its application to interpreting DNA data.

The FBI had developed a population statistics computer program (Popstats) for calculating DNA match statistics. This software was distributed free of charge to state and local crime laboratories that used the FBI's

COmbined DNA Index System (CODIS) database of DNA from convicted offenders, crime scenes, and other sources. Popstats could calculate match statistics for DNA mixtures using the combined probability of inclusion (CPI) method.

DNA mixture data contains considerable information. Since genotypes come in allele pairs, the peak heights in mixture data can be used to separate the data into the genotypes of each contributor.[20] For example, a data pattern of two tall peaks and two short peaks can be separated into a major genotype providing a large quantity of two alleles and a minor genotype contributing a small amount of two other alleles.

The FBI chose to simplify mixture interpretation by simplifying the data. Its approach made little use of the STR data's highly informative peak heights, patterns, and variation. Instead, the FBI applied a threshold to the data signal, separating peaks into two categories — present or absent. Since thresholds worked well with single source DNA data, the idea was to use them again for interpreting mixtures.

A probability of inclusion (PI) statistic was easily calculated at each locus by plugging the population frequency of each present allele into a simple inclusion probability formula (add the frequencies and square the sum). Since the loci are independent, multiplying the individual locus PI's together using the product rule calculates a combined CPI probability.[21]

The SWGDAM 2000 guidelines[22] for STR interpretation were promulgated and widely adopted. Most DNA mixture interpretation in the United States was then done using CPI or a related threshold-based method that considered a known victim reference (CLR). These simple methods let the FBI and crime labs across the country analyze mixtures of two or more people, the most common type of criminal DNA evidence.

There is no scientific basis, however, for this threshold approach to analyzing mixtures. Modifying observed data can introduce error or bias. Since peak variation scales with height in the PCR counting process, a "one size fits all" threshold cannot work with mixtures — major and minor contributors have different variations in allele quantity. The CPI method was scientifically unfounded.

The FBI and other labs did not empirically validate mixture thresholds or CPI for match statistic accuracy. Data interpretation followed SWGDAM rules, not rigorous science. Laboratory audits and accreditation stressed adherence to these rules, ensuring widespread usage. Unfortunately, thresholds often gave "inconclusive" results on informative data, and CPI usually gave an inaccurate DNA match statistic.[23]

2005: Notice

A wake-up call came in 2005 from the National Institute of Standards and Technology (NIST) in the U.S. Department of Commerce. NIST conducted a MIX05 inter-laboratory comparison study, sending the same two-person DNA mixture data to 69 participating laboratories for interpretation.[24] The results showed extreme variation in reported results. There were many "inconclusive" responses. The 29 labs that provided match statistics had numbers ranging from 31 thousand to 213 trillion, spanning 10 powers of ten on the same data.[25]

These results on mixture statistic variation were widely disseminated throughout the forensics DNA community and wider audiences.[26] NIST presented the mixture reporting discrepancies at conferences, workshops, and scientific meetings. However, the crime labs continued to use the FBI Popstats software, reporting inaccurate DNA mixture match statistics that were not validated or reproducible.

The CPI mixture interpretation failure extends to the CODIS database. Simple allele list comparison (based on set intersection, not identification information) has a high false positive match rate for DNA mixtures. To reduce false matches to the wrong suspect, most mixtures analyzed by crime labs are not uploaded to this database. Investigators cannot use CODIS for this DNA evidence to solve crimes or identify suspects.

2010: Crisis

Scientists and statisticians wrote about the DNA mixture failure. They contended that thresholds lacked a scientific foundation.[27] They found that CPI statistics for low-level mixtures with little DNA could be unfair to defendants.[28] They questioned whether CPI even made any sense as a match statistic.[29]

There was concern about human bias in the CPI method, and producing subjective results that were suspect-centric or pro-prosecution.[30] A human analyst first adjusts the data (applying thresholds, removing apparent stutter, etc.), and then looks at the defendant's genotype to decide if the person is included in the mixture. Only after first changing the data and assuming inclusion does the analyst then run CPI software to calculate a match statistic, a number often used in court to help establish guilt. Assuming guilt to establish guilt is circular reasoning.

There is bias when an analyst subjectively picks data by choosing loci after first looking at the defendant's genotype.[31] One report showed how analysts could justify including any "Tom, Dick, or Harry" who was not actually in the DNA evidence.[32] In another study, analysts who had the "potentially biasing context" that their corroborating DNA evidence "was essential to the prosecution" did not exclude a defendant from a mixture; however, without such context, only 1 of 17 other DNA examiners agreed, while 16 "reached a different and conflicting conclusion" (12 exclude, 4 inconclusive).[33] Most mixture interpretation software requires an analyst to prepare the input by first selecting a subset of the data.

In its oft cited "cartoon" paper,[34] the FBI proposed a solution: since one threshold failed, use two thresholds. The FBI introduced a second "stochastic threshold" at a higher level to discard data that might have too much variation. No statistical theory or empirical data supported this unfounded proposal — just cartoon drawings. No validation studies were done to establish accuracy. Sophisticated mathematics can model data variation, but applying another simplistic threshold simply discards more data.

Regardless, the FBI's SWGDAM 2010 guidelines[35] imposed stochastic thresholds on crime laboratories, making the cartoon paper *de facto* national policy. The labs compliantly determined these thresholds and applied them to mixture evidence. The second threshold greatly decreased their match statistics and increased inconclusive outcomes,[36] eliminating needed DNA information.

2010: Alternative

In 1999, Cybergenetics began developing a sophisticated statistical solution to DNA mixture interpretation.[37] After a decade of development, testing and refinement, the TrueAllele® technology become available. In 2009, TrueAllele was used in a Pennsylvania homicide trial.[38] A scientific study (the first of over 30 such validation studies, seven of them peer-reviewed) demonstrated a large "information gap" between the empirically tested TrueAllele system and the FBI's unvalidated CPI approach.[39]

Instead of discarding peak data, TrueAllele uses all the height information. Rather than applying thresholds, TrueAllele statistically determines PCR and instrument variation directly from the data. The statistical computation explores most conceivable solutions, objectively separating genotypes out of the mixture data. Only afterwards does comparison with another person's genotype calculate a match statistic. A human analyst cannot bias the outcome: all the data goes into the computer, with results based on scientific evidence, not a person's preconceptions.

However, some DNA workers preferred to "control their data" and would not cede that control to an accurate, objective forensic calculator. Others viewed TrueAllele as a threat to their jobs or status as "DNA experts." Moreover, TrueAllele automation opened a window into past interpretation failures that could expose potential liabilities. The subjective examination of DNA mixture data continued to produce inaccurate or inconclusive match statistics.

2015: Failure

A 2011 TrueAllele validation study conducted jointly with the NYSP DNA lab (Albany, New York) showed that CPI vastly underreported DNA's probative value.[40] Whenever the lab was able to report a CPI statistic, the number was (on average) a million times less than the true match statistic on the same data. CPI analysis removed considerable DNA information.

A 2013 NYSP validation study examined how human mixture analysis performed on data where the TrueAllele computer produced a match result.[41] TrueAllele's median match statistic was around a quadrillion. When TrueAllele gave a result, 70 percent of the time thresholds failed to report any match statistic. Human review was silent about most DNA evidence, incorrectly concluding that informative items were inconclusive.

In 2013, NIST conducted a MIX13 inter-laboratory study.[42] The hope was that the new stochastic threshold procedure had adequately addressed natural data variation. The hope went unrealized when a hundred participants examined a three-person mixture that did *not* contain a particular suspect. Seventy groups *incorrectly included* this suspect, whose DNA was not present in the mixture (70 percent false match rate), giving irrelevant DNA match statistics that ranged from 9 to 344,000. Twenty-four labs found the comparison inconclusive. Only six correctly excluded the suspect (6 percent accuracy rate), with one of them using TrueAllele.

A 2014 TrueAllele validation paper conducted on 72 Virginia mixture cases showed the extent of CPI's lost information.[43] On 100 DNA comparisons, the average TrueAllele match statistic of a hundred billion (10^{11}) dropped to only millions (10^{6}) when a threshold was applied and CPI calculated. Applying a second (stochastic) threshold to the same mixture data further reduced the modified CPI statistic to just hundreds (10^{2}). Moreover, the SWGDAM 2010 procedure did not eliminate all false matches.

In 2015, comparison of inclusion probability with TrueAllele match information showed that CPI was a one-sided random number generator, uncorrelated with identification information.[44] The subjective CPI statistic depends on the number of loci tested, not on the probative value of the DNA evidence. That is why (using all loci) CPI always gave the same answer — around a million — regardless of the data. After an analyst first decides that a defendant's DNA is in a mixture (viewed as guilt by a jury), CPI can afterwards provide an impressive statistic that only restates a human judgment.

Law

DNA holds considerable prejudicial sway over a jury. In a courtroom, the three letters can seem to abbreviate "Do Not Acquit." When DNA match statistics are routinely wrong or lack probative value, it is hard to justify introducing them in criminal trials. The Federal Rules of Evidence (FRE) provide legal mechanisms for excluding harmful DNA evidence from court.[45]

FRE 403 permits a court to "exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence." A CPI match statistic essentially counts up the number of loci deemed an "inclusion" by a human analyst.[46] The statistic is cumulative evidence that reframes an analyst's subjective conclusions as an objective-sounding match number that can mislead a jury. Since CPI is uncorrelated with identification information, it has little probative value. Mixture statistics that are more prejudicial than probative can be challenged in a pretrial hearing to keep the jury from hearing unfair DNA results.

FRE 702 guides who can testify as an expert witness to render a scientific opinion about DNA evidence. The expert's testimony must be based on reliably applying a reliable method to sufficient data. After a pretrial hearing, a judge can exercise her gatekeeper role to protect the jury from hearing unreliable scientific evidence. Unreliable DNA match statistics are susceptible to challenge. The judge may rule that inaccurate or insufficiently validated DNA mixture statistics are not admissible. Challenging unreliable DNA interpretation can keep out bad evidence, even when there is good underlying data.[47]

In *Brady v. Maryland*,[48] the U.S. Supreme Court held that the suppression by the prosecution of evidence favorable to an accused upon request violates due process where the evidence is material either to guilt or to punishment. This ruling applies irrespective of the good faith or bad faith of the prosecution because society wins not only when the guilty are convicted, but when criminal trials are fair.

When mixture interpretation fails, no DNA match statistic is reported. The absence of a report can hide potential exculpatory DNA evidence. But if a defendant requests all data from all laboratory testing, *Brady* requires the government to provide that data. Effective interpretation of the government's DNA data by an independent expert might exonerate the accused or implicate another person.

Conclusion

Unscientific, untested "statistical" analysis of DNA mixtures has led to incorrect results on hundreds of thousands of items of evidence. When thresholds give an "inconclusive" result on mixtures with data, that silent nonanswer is usually wrong. Likewise, when CPI match statistics are reported, the answer is usually wrong.

Innocent people remain in prison because informative DNA was not used in their defense. Defendants are wrongfully convicted when misinterpreted DNA cannot identify the true culprit. Perpetrators go free when DNA evidence is failed by forensic statistics. Freed criminals then commit more crime, which DNA should have prevented, needlessly harming innocent victims. This is not the fairest justice that DNA science can provide.

Modern genotyping programs use probability to help interpret DNA mixtures.[49] TrueAllele has a fully Bayesian model[50] that considers all data and all solutions. Less thorough programs remove data to simplify the problem

using thresholds, dropout parameters,[51] or peak filters.[52] Subjective programs let a human operator choose input data and parameters to overcome software limitations.[53] While crime labs have started adopting better match statistic software, validation studies are needed to determine their range of applicability.[54]

The following recommendations may help society move beyond mixture interpretation failure and enjoy consistently more reliable DNA evidence:

1. *Open DNA data to public scrutiny*. The crime labs have failed to produce reliable match statistics for over 15 years. The solution is open access to all DNA data so that impartial scientists can publicly reassess crime lab results in every case.

2. *Revisit all past DNA mixture cases*. Hundreds of thousands of DNA mixtures have been improperly interpreted. Only an unbiased, accurate software review of all this evidence can rectify the problem.

3. *Educate trial attorneys and judges*. Law attracts many who would rather not study science or mathematics. However, lawyers need to understand the evidence they attack or defend. Appropriate education is needed to teach them DNA statistics.

4. *Fully automate mixture interpretation*. Human analysts are trained to remove DNA data from the input to their interpretation software, which introduces bias and error. Automated computing can help eliminate such human decision-making.

5. *Extensively validate DNA interpretation*. Most mixture statistics have not been validated for their intended use. No method, whether done by man or machine, should ever be introduced as evidence without supporting validation.

6. *Keep methods within their limits.* Defense vigilance helps ensure that crime labs stay within the bounds of their validated interpretation methods. Without this DNA pressure, false positives may falsely identify or convict innocent people.

7. *Go beyond laboratory limits*. Better interpretation methods can solve DNA mixtures that crime labs cannot. Independent groups should interpret these data. Otherwise false negatives may fail to identify, withholding potentially exculpatory evidence.

Unfounded DNA statistics have inflicted considerable injustice on defendants, crime victims, and society. Every case that involved inconclusive DNA mixtures or unfounded match statistics should be revisited. It is time to rectify two decades of forensic failure with accurate, objective, and validated DNA interpretation.

© *Mark W. Perlin, 2016. All rights reserved*.

Notes

1. N.J. Schweitzer & M.J. Saks, *The CSI Effect: Popular Fiction About Forensic Science Affects the Public's Expectations About Real Forensic Science,* 47 Jurimetrics 357 (2007).

2. National Research Council, Strengthening Forensic Science in the United States: A Path Forward (2009).

3. J.M. Butler, Forensic DNA Typing: Biology and Technology Behind STR Markers (2000).

4. K.L. Alexander, *National Accreditation Board Suspends All DNA Testing at D.C. Crime Lab*, Wash. Post, May 26, 2015; T. Plohetski, *Austin Police DNA Lab Closed Amid Forensics Commission's Concerns*, Austin American-Statesman, June 10, 2016.

5. M.P. Flaherty, *Virginia Re-evaluates DNA Evidence in 375 Cases*, Wash. Post, July 16, 2011.

6. S. Augenstein, *DNA Mixtures Present Statistical Problem, Texas Labs Proactively Examining Thousands of Cases*, Forensic Magazine, October 2015.

7. Editorial Board, *A Mystery at the State Lab*, Times Union, 2016.

8. M. Shaer, *The False Promise of DNA Testing*, The Atlantic, June 2016.

9. K.B. Mullis, *The Unusual Origin of the Polymerase Chain Reaction*, Scientific American, April 1990.

10. L.M. Smith et al., *Fluorescence Detection in Automated DNA Sequence Analysis*, Nature, June 1986.

11. G. Williams, *The Apple Macintosh Computer*, Byte Magazine, 1984.

12. J. Weber & P. May, *Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction*, 44 Am. J. Hum. Genet. 388 (1989).

13. L.S. Schwartz et al., *Fluorescent Multiplex Linkage Analysis and Carrier Detection for Duchenne/Becker Muscular Dystrophy*, 51 Am. J. Hum. Genet. 721 (1992).

14. X.Y. Hauge & M. Litt, *A Study of the Origin of 'Shadow Bands' Seen When Typing Dinucleotide Repeat Polymorphisms by the*

PCR, 2(4) Hum. Molec. Genet. 411 (1993).

15. J. Weber & P. May, *Abundant Class of Human DNA Polymorphisms Which Can Be Typed Using the Polymerase Chain Reaction*, 44 Am. J. Hum. Genet. 388 (1989).

16. A. Edwards et al., *DNA Typing and Genetic Mapping with Trimeric and Tetrameric Tandem Repeats*, 49 Am. J. Hum. Genet. 746 (1991).

17. M.W. Perlin et al., *Toward Fully Automated Genotyping: Genotyping Microsatellite Markers by Deconvolution*, 57(5) Am. J. Hum. Genet. 1199 (1995).

18. S.Y. Hill et al., *A Genome Wide Search for Alcoholism Susceptibility Genes*, Am. J. Med. Genetics Part B: Neuropsychiatric Genetics, 128B(1):102-13 (2004).

19. G. Stolovitzky & G. Cecchi, *Efficiency of DNA Replication in the Polymerase Chain Reaction*, 93(23) Proc. Nat'l Acad. Sci. USA 12947 (1996).

20. M.W. Perlin & B. Szabady, *Linear Mixture Analysis: A Mathematical Approach to Resolving Mixed DNA Samples*, 46(6) J. Forensic Sci. 1372 (2001).

21. For a CPI example with data and calculations, *see* M.W. Perlin, *Inclusion Probability for DNA Mixtures Is a Subjective One-Sided Match Statistic Unrelated to Identification Information*, 6(1) J. Pathol. Inform. 59 (2015).

22. Scientific Working Group on DNA Analysis Methods (SWGDAM), Short Tandem Repeat (STR) Interpretation Guidelines, 2(3) Forensic Science Communications (2000).

23. M.W. Perlin, *Inclusion Probability for DNA Mixtures Is a Subjective One-Sided Match Statistic Unrelated to Identification Information*, 6(1) J. Pathol. Inform. 59 (2015).

24. J.M. Butler & M.C. Kline, NIST Mixture Interpretation Interlaboratory Study (2005) (MIX05) Poster #56, Promega's Sixteenth International Symposium on Human Identification; Grapevine, Texas (2005).

25. J.M. Butler, Standardizing Mixture Interpretation: An Open Discussion, Accommodating the Demands of Increasing Volume: A Workshop for Public and Private DNA Profiling Laboratories, Grapevine, Texas, Human Identity Trade Association (2005).

26. L. Geddes, *How DNA Evidence Creates Victims of Chance*, New Scientist, August 2010, https://www.newscientist.com/article/mg20727743-300-how-dna-evidence-creates-victims-of-chance (last visited June 17, 2018).

27. P. Gill & J. Buckleton, Commentary on: B. Budowle et al., *Mixture Interpretation: Defining the Relevant Features for Guidelines for the Assessment of Mixed DNA Profiles in Forensic Casework*, 54(4) J. Forensic Sci. 810 (2009); 55(1) J. Forensic Sci. 265 (2010).

28. D.J. Balding & J. Buckleton, *Interpreting Low Template DNA Profiles*, 4(1) Forensic Sci. Int. Genet. 1 (2009).

29. C.H. Brenner, The Mythical 'Exclusion' Method for Analyzing DNA Mixtures — Does It Make Any Sense at All?, AAFS 63rd Annual Scientific Meeting (Chicago, Illinois): American Academy of Forensic Sciences at 79 (2011).

30. National Research Council, Strengthening Forensic Science in the United States: A Path Forward (2009).

31. J.M. Curran & J. Buckleton, *Inclusion Probabilities and Dropout*, 55(5) J. Forensic Sci. 1171 (2010).

32. W.C. Thompson, *Painting the Target Around the Matching Profile: The Texas Sharpshooter Fallacy in Forensic DNA Interpretation*, 8(3) Law, Probability and Risk 257 (2009).

33. I.E. Dror & G. Hampikian, *Subjectivity and Bias in Forensic DNA Mixture Interpretation*, 51(4) Science & Justice 204 (2011).

34. B. Budowle et al., *Mixture Interpretation: Defining the Relevant Features for Guidelines for the Assessment of Mixed DNA Profiles in Forensic Casework*, 54(4) J. Forensic Sci. 810 (2009).

35. Scientific Working Group on DNA Analysis Methods (SWGDAM), Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories (2010), http://www.fbi.gov/about-us/lab/codis/swgdam-interpretation-guidelines.

36. M.W. Perlin et al., TrueAllele® Casework on Virginia DNA Mixture Evidence: Computer and Manual Interpretation in 72 Reported Criminal Cases, 9(3) PLoS ONE e92837 (2014).

37. M.W. Perlin & B. Szabady, *Linear Mixture Analysis: A Mathematical Approach to Resolving Mixed DNA Samples*, 46(6) J. Forensic Sci. 1372 (2001).

38. M.W. Perlin, *The Blairsville Slaying and the Dawn of DNA Computing*, in A. Niapas (ed.), Death Needs Answers: The Cold-Blooded Murder of Dr. John Yelenic (2013).

39. M.W. Perlin & A. Sinelnikov, *An Information Gap in DNA Evidence Interpretation*, 4(12) PLoS ONE e8327 (2009), doi: 10.1371/journal.pone.0008327.

40. M.W. Perlin et al., *Validating TrueAllele® DNA Mixture Interpretation*, 56(6) J. Forensic Sci. 1430 (2011).

41. M.W. Perlin, J.L. Belrose & B.W. Duceman, *New York State TrueAllele® Casework Validation Study*, 58(6) J. Forensic Sci. 1458 (2013).

42. M.D. Coble et al., NIST Inter-Laboratory Studies for DNA Mixture Interpretation (MIX13), 66th Annual Meeting of the American Academy of Forensic Sciences, Seattle, Washington (2014).

43. M.W. Perlin et al., TrueAllele® Casework on Virginia DNA Mixture Evidence: Computer and Manual Interpretation in 72 Reported Criminal Cases, 9(3) PLoS ONE e92837 (2014).

44. M.W. Perlin, *Inclusion Probability for DNA Mixtures Is a Subjective One-Sided Match Statistic Unrelated to Identification Information*, 6(1) J. Pathol. Inform. 59 (2015).

45. G. Fisher, Evidence (3d. ed.) (2012).

46. M.W. Perlin, *Inclusion Probability for DNA Mixtures Is a Subjective One-Sided Match Statistic Unrelated to Identification Information*, 6(1) J. Pathol. Inform. 59 (2015).

47. M.W. Perlin, *When Good DNA Goes Bad*, J. Forensic Res. (2013), S11:003. doi: 10.4172/2157-7145.S11-003.

48. *Brady v. Maryland*, 373 U.S. 83 (1963).

49. J. Ropero-Miller et al., *Landscape Study of DNA Mixture Interpretation Software*, National Institute of Justice (2015).

50. A. O'Hagan & J. Forster, Bayesian Inference (2d ed. 2004).

51. D.J. Balding, *Evaluation of Mixed-Source, Low-Template DNA Profiles in Forensic Science*, 110(30) Proc. Nat'l Acad. Sci. USA 12241 (2013).

52. R.G. Cowell et al., *Analysis of DNA Mixtures with Artefacts*, 64(1) J. Royal Stat. Soc'y Series C Applied Stat. 1 (2013).

53. M.R. Wilson, In response to: Todd W. Bille et al., *Comparison of the Performance of Different Models for the Interpretation of Low Level Mixed DNA Profiles*, 35 Electrophoresis 3125 (2014); 36 Electrophoresis 489 (2015).

54. Scientific Working Group on DNA Analysis Methods (SWGDAM), Guidelines for the Validation of Probabilistic Genotyping Systems (2015), http://media.wix.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf.

About the Author

Dr. Mark Perlin is the Chief Scientific Officer at Cybergenetics.

Mark W. Perlin, Ph.D., M.D., Ph.D.
Cybergenetics
Pittsburgh, Pennsylvania
412-683-3004
Email perlin@cybgen.com