

**THE ALGORITHM SAYS YOU DID IT: THE USE OF BLACK
BOX ALGORITHMS TO ANALYZE COMPLEX DNA EVIDENCE**

*Katherine Kwong**

TABLE OF CONTENTS

I. INTRODUCTION.....	275
II. BACKGROUND.....	276
<i>A. The Science of Complex DNA Mixtures</i>	276
<i>B. The Unreliability of Previous Methods</i>	279
<i>C. The Development of Algorithmic Analytic Techniques</i>	281
<i>D. TrueAllele</i>	282
III. USAGE IN THE CRIMINAL JUSTICE SYSTEM	283
<i>A. The First Case: Commonwealth v. Foley</i>	283
<i>B. Subsequent Unsuccessful Challenges to the Use of DNA Analysis Algorithms</i>	286
<i>C. Use in Exoneration Cases</i>	287
IV. CONCERNS AND CRITICISMS	288
<i>A. Unestablished Scientific Validity</i>	288
<i>B. Lack of Transparency</i>	291
V. POTENTIAL RESPONSES	295

I. INTRODUCTION

DNA evidence has grown to be widely accepted as reliable proof of an individual's innocence or guilt.¹ Yet, despite the perception of DNA evidence as definitive proof, when DNA evidence involves complex mixtures of multiple individuals' DNA, the science is not as simple as it appears on television. Complex DNA samples are not as straightforward and objective to analyze as simple DNA samples, leaving substantial room for error and variability.² Commonly used techniques for analyzing and interpreting complex DNA mixtures have proven unreliable,

* J.D., Harvard Law School, 2017; MPH, Public Health Genetics, University of Washington, 2014. The author is grateful to Prof. Christopher Bavitz for his suggestions and support in developing this paper and to the editing team at JOLT for their input and hard work.

1. See EXEC. OFFICE OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., REPORT TO THE PRESIDENT: FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 2 (2016) [hereinafter PCAST REPORT].

2. See *id.* at 8.

creating concerns about the potential for improper prosecutions and convictions.³

To address the problems of unreliability associated with the subjective techniques typically used to interpret complex DNA mixture results, a number of companies and organizations are working to develop algorithmic systems to interpret the results of analyses of complex DNA mixtures.⁴ Unfortunately, these algorithmic programs have problems of their own. Multiple parties have raised concerns about the reliability and accuracy of the algorithmic programs, questioning their scientific validity, and the lack of transparency surrounding the algorithms and their use.⁵ The technologies that were intended to solve the problems associated with subjective interpretations of complex DNA mixture analyses have instead opened the door to a whole new set of problems that must be resolved.

Part II of this Note describes the science behind simple and complex DNA mixture analyses, the troubles with subjective analytic techniques, and the background of TrueAllele and related DNA analysis technologies. Part III explores how these technologies have been used in the criminal justice system for both exoneration and conviction, including how courts have ruled in response to challenges to their use. Part IV evaluates criticisms of the use of algorithmic DNA analysis technologies in the criminal justice system, including concerns about their scientific validity and the lack of transparency. Finally, Part V discusses potential responses to these criticisms.

II. BACKGROUND

A. The Science of Complex DNA Mixtures

DNA evidence has long been upheld as the “gold standard” for forensic science.⁶ Indeed, the public views DNA evidence as extremely reliable and accurate. According to a 2005 poll conducted by Gallup, 85% of Americans consider DNA evidence to be “very or completely reliable.”⁷ In multiple studies, researchers have found that jurors believe DNA evidence is more than 90% accurate.⁸ Unfortunately, that perceived certainty glosses over much of the complexity surrounding some types of DNA evidence.

3. *See id.*

4. *See id.* at 78.

5. *See, e.g., id.* at 8.

6. *Id.* at 2.

7. Katie Worth, *The Surprisingly Imperfect Science of DNA Testing*, FRONTLINE, <http://stories.frontline.org/dna> [<https://perma.cc/3EWA-DNKN>].

8. *See id.*

The science surrounding the analysis and interpretation of DNA evidence has evolved and grown over time.⁹ Most DNA analysis for forensic purposes involves samples from only one or two individuals.¹⁰ Analyses of DNA samples that come from a single individual (single-source samples) or from a simple mixture of two individuals (simple-mixture samples) have been well studied and thoroughly tested, and, according to the President’s Council of Advisors on Science and Technology (“PCAST”), they are generally considered to be “objective method[s] in which the laboratory protocols are precisely defined and the interpretation involves little to no human judgment.”¹¹ Objective methods — defined by PCAST as “methods consisting of procedures that are each defined with enough standardized and quantifiable detail that they can be performed by either an automated system or human examiners exercising little or no judgment” — are considered more scientifically valid and reliable than subjective methods, which “involve significant human judgment.”¹² Single-source and simple-mixture sample analyses are considered highly reliable because each of the steps involved in the analysis is “repeatable, reproducible, and accurate.”¹³ This trio of requirements is referred to as “foundational validity,” a concept that shows that a method can, *in principle*, be reliable.¹⁴ Foundational validity also requires estimates of accuracy — that is, empirical measurements of how frequently a method reaches an incorrect conclusion.¹⁵ This scientific concept “correspond[s] to the legal requirement . . . of ‘reliable principles and methods.’”¹⁶ Errors are considered unlikely so long as quality assurance standards are followed to prevent human errors arising from sample contamination, mislabeled samples, incorrect interpretation, or improper result reporting.¹⁷ This ability to reliably apply a method in practice (“validity as applied”) corresponds to the legal requirement that an expert “has reliably applied the principles and methods to the facts of the case.”¹⁸

In recent years, investigators have sought to use more complicated sources of DNA evidence, such as those that contain mixtures of multiple unknown persons’ DNA in unknown proportions.¹⁹ Such samples, known as complex mixtures, can come from sources like mixed blood stains, rape kits from gang rape cases, or surfaces where multiple indi-

9. See PCAST REPORT, *supra* note 1, at 2.

10. See *id.* at 7.

11. *Id.*

12. *Id.* at 5 n.3.

13. *Id.* at 7.

14. *Id.* at 4–5.

15. See *id.* at 5.

16. *Id.*

17. See *id.* at 7.

18. See *id.* at 5.

19. See *id.* at 7.

viduals have left behind minute amounts of DNA.²⁰ Improvements in methods for extracting DNA from evidentiary samples, laboratory analysis techniques, and other techniques to increase detection sensitivity have enabled forensic laboratories to analyze more complex mixtures than was previously possible.²¹

Complex mixtures, unlike single-source samples or simple mixtures, do not result in purely objective test results.²² While the laboratory processing of complex-mixture samples is the same as that for single-source and simple-mixture samples, the interpretation of the results of that processing differs significantly.²³ The DNA profile produced from complex samples contains all of the individual DNA profiles superimposed atop one another, which always requires some level of interpretation in order to determine which portions of the results may (or may not) have come from a suspect.²⁴

The required intervention of human judgment means this type of interpretation always involves some level of subjectivity. In single-source and simple-mixture analyses, every step can be objectively determined in a way that will be repeatable and reproducible, independent of judgment calls or subjective decisions. With complex DNA analysis, decisions must be made between different interpretations that might be equally or similarly valid — and those decisions may have significant impacts on the ultimate results of the analysis.

It is frequently impossible to tell how many individuals' DNA is present within a complex mixture, much less accurately distinguish each person's unique DNA profile from the overall mixture.²⁵ One study estimated that 3% of all three-person mixtures could be mistaken as containing the DNA of only two people, while 76% of all four-person mixtures could be mistaken as containing the DNA of either two or three people.²⁶ These challenges are frequently exacerbated by samples that have degraded or which originally contained only a small amount of DNA.²⁷

20. *See id.*

21. *See* Frederick Bieber et al., *Evaluation of Forensic DNA Mixture Evidence: Protocol for Evaluation, Interpretation, and Statistical Calculations Using the Combined Probability of Inclusion*, 17 *BMC GENETICS* 125, 126 (2016).

22. *See* PCAST REPORT, *supra* note 1, at 8.

23. *See id.* at 75.

24. *See id.* at 8.

25. *See id.*

26. *See* David R. Paoletti et al., *Empirical Analysis of the STR Profiles Resulting from Conceptual Mixtures*, 50 *J. FORENSIC SCI.* 1, 4 (2005).

27. *See* Bieber et al., *supra* note 21, at 129.

B. The Unreliability of Previous Methods

Forensic scientists have historically used a combination of subjective judgment and rudimentary calculations to interpret complex DNA mixture analyses.²⁸ Because the subjective choices made by the person analyzing the DNA test can significantly affect the result, there is a high risk of human error and/or bias being introduced into the interpretation.²⁹ Therefore, according to a report by PCAST on the scientific validity of various forensic science methods, “subjective analysis of complex DNA mixtures has not been established to be foundationally valid and is not a reliable methodology.”³⁰

In some ways, DNA evidence’s seemingly-certain success at proving the identity of individuals in the past has undermined its validity in the present, as investigators and labs seek to “push[] the envelope” with samples involving more complex mixtures and lower amounts of DNA available to analyze.³¹ Scientists from the National Institute of Standards and Technology (“NIST”) have raised concerns that, although lab methodologies for the analysis of DNA evidence have improved, the statistical interpretation techniques used to evaluate the laboratory test results have not improved at the same pace, undermining the quality of the final analytic interpretations.³²

The result is that many forensic laboratories’ current methods of interpreting analyses of samples with three or more individuals’ DNA or with low levels of DNA may be extremely unreliable.³³ In 2013, NIST asked 108 forensic labs to evaluate a three-person mixture to determine whether a suspect’s DNA was present in the mixture and received wildly varying conclusions — variation NIST ascribed to flaws in analytic methods.³⁴ Other studies have found similarly wide variations in conclusions, depending on who is analyzing the sample.³⁵

Human error or bias can also play a significant role in analyses, particularly when complex mixtures are involved. Because many labs do not blind interpreters from knowing details about cases, the technicians evaluating the results may be biased towards finding false positive

28. See PCAST REPORT, *supra* note 1, at 8.

29. *See id.*

30. *Id.*

31. See Michael D. Coble & John M. Butler, *DNA Mixture Interpretation: State of the Art*, Nat’l Inst. Standards and Tech., Presentation to the American Society of Crime Laboratory Directors/Laboratory Accreditation Board 54, (Jan. 8, 2015), http://www.cstl.nist.gov/strbase/pub_pres/ASCLD-LAB-Jan2015-CobleButler.pdf [<https://perma.cc/AS8J-5Y8C>].

32. *See id.*

33. *See id.* at 50–51.

34. *See id.* at 10–16.

35. See generally Linda Geddes, *Fallible DNA Evidence Can Mean Prison or Freedom*, NEW SCIENTIST (Aug. 11, 2010), <https://www.newscientist.com/article/mg20727733.500-fallible-dna-evidence-can-mean-prison-or-freedom/> [<https://perma.cc/WNF8-DA9D>].

matches.³⁶ Research has shown that when laboratory experts receive contextualizing information about the samples they are analyzing, their interpretations change, demonstrating both the risk of bias and the subjective nature of DNA mixture interpretation.³⁷ In one famous study, researchers used DNA evidence from a real-world gang rape case to demonstrate how contextualizing information might bias experts and improperly influence their results.³⁸

The two experts in the original case were aware of testimony against the suspect and were aware prosecutors were eager to use DNA evidence to corroborate the testimony.³⁹ Their analysis, which relied on a certain level of subjective judgment, concluded the suspect could not be excluded as a contributor to the DNA mixture.⁴⁰ Yet when the researchers presented the evidence to seventeen other experts without the contextualizing information, only one of the seventeen agreed with the original experts.⁴¹ Twelve of the seventeen experts went so far as to exclude the suspect as a possible contributor, reaching the opposite conclusion from the original experts who may have been swayed by extraneous evidence.⁴²

Bias is not the only potential risk involved with subjective analysis. Relatively minor flaws in protocols for calculating the probability of a match can have dramatic consequences. After Texas found a small number of errors in the FBI database it used for DNA match statistics, it offered to retest samples upon request, assuming the necessary tweaks to its calculations would result in minor changes to their results.⁴³ Instead, some probabilities shifted by orders of magnitude upon retesting.⁴⁴ In response, Texas began a massive effort to revisit old cases involving DNA mixture interpretation, potentially affecting thousands of cases going back more than 15 years.⁴⁵

Texas found that the massive discrepancies in probabilities between previous testing and retesting were not caused by the minor corrections

36. See Laurie Meyers, *The Problem with DNA*, 38 *MONITOR PSYCHOL.* 52 (2007).

37. See Itiel Dror & Greg Hampikian, *Subjectivity and Bias in Forensic DNA Mixture Interpretation*, 51 *SCI. & JUST.* 204, 204 (2011).

38. See *id.* at 205.

39. See *id.*

40. See *id.*

41. *Id.*

42. *Id.*

43. See Memorandum from Vincent J.M. Di Maio, MD, Presiding Officer, Tex. Forensic Sci. Comm'n, to Members of the Texas Criminal Justice Community (2015), <http://www.fsc.texas.gov/sites/default/files/documents/Unintended%20Effects%20of%20FBI%20Database%20Corrections%20on%20Assessment%20of%20DNA%20Mixture%20Interpretation%20in%20Texas%20NOTICE.pdf> [<https://perma.cc/ZWF7-WMBF>].

44. See Martin Kaste, *'Great Pause' Among Prosecutors as DNA Proves Fallible*, NPR (Oct. 9, 2015), <http://www.npr.org/2015/10/09/447202433/-great-pause-among-forensic-scientists-as-dna-proves-fallible> [<https://perma.cc/X3NC-7HNH>].

45. See *id.*

in the FBI database.⁴⁶ Instead, they were caused by flaws in the protocols for calculating statistical probabilities that an individual was present in a given complex-mixture sample.⁴⁷ The protocols for calculating such probabilities failed to adequately constrain subjective decisions and did not clearly state the limitations of the technique.⁴⁸

In response to these findings, Texas convened a group of experts to write a scientific protocol for a standardized approach to calculating probabilities related to complex-mixture samples.⁴⁹ While the proposed rules may help define a more objectively valid method, concerns remain that the subjectivity inherent in human interpretation may still cause problems despite the newly proposed rules' attempts to specify protocols.

C. The Development of Algorithmic Analytic Techniques

A number of companies have sought to address the issues with subjective analyses of complex mixtures by developing computer programs to consistently apply algorithmic decision making to complex mixture analysis.⁵⁰ PCAST states in its report that computerized algorithmic analysis programs “clearly represent a major improvement over purely subjective interpretation,” such as is typically practiced in a number of jurisdictions, but cautions that such programs must still be scrutinized to determine their reliability and validity.⁵¹ Concerns about the scientific validity of algorithmic analysis programs for complex DNA mixtures are discussed in greater detail in Part IV, *infra*.

Computerized algorithmic analytic programs, which rely on a technique called “probabilistic genotyping,” use mathematical algorithms to interpret complex DNA mixtures.⁵² Probabilistic genotyping uses mathematical models and simulations to estimate the likelihood that a particular individual’s DNA is part of the mixture present in the sample.⁵³

Probabilistic genotyping programs have incited significant excitement among law enforcement due to their potential to speed up analysis, remove the potential for human error, and permit analysis of samples

46. See PCAST REPORT, *supra* note 1, at 77–78.

47. See *id.* at 78.

48. See *id.*

49. See generally Bieber et al, *supra* note 21.

50. PCAST REPORT, *supra* note 1, at 8.

51. *Id.*

52. *Id.* at 78–79.

53. See EXEC. OFFICE OF THE PRESIDENT, PRESIDENT’S COUNCIL OF ADVISORS ON SCI. & TECH., AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS 8 (2017) [hereinafter PCAST ADDENDUM].

earlier techniques could not successfully analyze.⁵⁴ One sheriff described probabilistic genotyping as “this amorphous, magical unicorn thing,” and stated, “Everybody is either in the process of purchasing or plans to purchase and will purchase [probabilistic genotyping programs] in the future.”⁵⁵ Improved speed and accuracy are certainly laudable goals worth pursuing if a new technology is able to provide them.

This rush of enthusiasm is present among potential program developers as well. According to the PCAST report, “As of March 2014, at least 8 probabilistic genotyping software programs had been developed (LRmix, Lab Retriever, likeLTD, FST, Armed Xpert, TrueAllele, STRmix, and DNA View Mixture Solution), with some being open source software and some being commercial products.”⁵⁶

This Note primarily focuses on TrueAllele, one of the most prominent of these programs due to ongoing legal disputes related to its use and its founder’s proactive advocacy; however, many of the issues discussed in relation to TrueAllele may also apply to these other algorithmic interpretation programs as well.

D. TrueAllele

TrueAllele is a product of the company Cybergenetics, which was founded in 1994.⁵⁷ It utilizes probabilistic genotyping to analyze complex DNA mixtures.⁵⁸ Users use their own typical laboratory procedures for DNA extraction, amplification, and processing, and then upload the data files to TrueAllele’s servers for analysis and visualization.⁵⁹ TrueAllele then interprets the DNA analysis files uploaded by users and utilizes its proprietary algorithms to provide users with likelihood ratios of sample matches.⁶⁰ These likelihood ratios are commonly used in criminal trials to attempt to persuade the jury the defendant was certainly the culprit.⁶¹

Cybergenetics and its CEO, Mark Perlin, make a point of vociferously criticizing existing interpretation techniques. Dr. Perlin published

54. See Tracy Clark-Flory, *A Revolutionary Algorithm to Clear Up Rape Kit Backlogs*, VOACTIV (Sept. 10, 2016), <http://www.vocativ.com/332592/a-revolutionary-algorithm-to-clear-up-rape-kit-backlogs/> [https://perma.cc/Y88C-KM82].

55. *Id.*

56. PCAST REPORT, *supra* note 1, at 78–79.

57. *History*, CYBERGENETICS, <https://www.cybgen.com/company/history.shtml> [https://perma.cc/W7SN-CHW9].

58. *Casework*, CYBERGENETICS, <https://www.cybgen.com/products/casework.shtml> [https://perma.cc/2DEU-UFY3].

59. TrueAllele, *TrueAllele Process Overview*, YOUTUBE (May 1, 2013), <https://www.youtube.com/watch?v=OU29b5sW88Y> (last visited Dec. 20, 2017).

60. See *Casework*, *supra* note 58.

61. See Matthew Shaer, *The False Promise of DNA Testing*, THE ATLANTIC, <https://www.theatlantic.com/magazine/archive/2016/06/a-reasonable-doubt/480747/> [https://perma.cc/X22C-DHF6].

an article that called the interpretation method used by most laboratories a “random generator”⁶² and argues in interviews that his company’s technology produces better probability measurements.⁶³ The company promotes its technology as a way of eliminating the potential for human error: TrueAllele can enable the “complete removal of the human being from doing any subjective decision making,” according to Dr. Perlin.⁶⁴

TrueAllele claims to be able to distinguish between two, three, or even four individuals in a DNA mixture;⁶⁵ some of its materials claim its technique has been validated on mixtures of up to 10 individuals’ DNA.⁶⁶ The company heavily promotes its ability to create results with “previously unsolvable DNA evidence” and claims it can overcome numerous potential issues, including samples with low total amounts of DNA and mixtures with low amounts of a particular individual’s DNA.⁶⁷

According to the company, as of December 2016, TrueAllele-generated evidence had been admitted to courts after a *Frye* or *Daubert* challenge in California, Indiana, Louisiana, Massachusetts, New York, Ohio, Pennsylvania, South Carolina, Virginia, Australia, and the United Kingdom.⁶⁸ Also as of December 2016, crime labs in California, Louisiana, Maryland, South Carolina, and Virginia were using TrueAllele software to analyze DNA samples in a total of 500 criminal cases in 35 states for both prosecution and defense purposes.⁶⁹

III. USAGE IN THE CRIMINAL JUSTICE SYSTEM

A. The First Case: *Commonwealth v. Foley*

According to Cybergenetics, the first case in which TrueAllele (or, the company claims, *any* “advanced statistical computing method for

62. Mark Perlin, *Inclusion Probability For DNA Mixtures Is A Subjective One-Sided Match Statistic Unrelated To Identification Information*, 6 J. PATHOLOGY INFORMATICS 59, 59 (2015).

63. Seth Augenstein, *DNA Mixture Calculation Method Just “Random Number Generator,” Says New Study*, FORENSIC MAGAZINE (Nov. 16, 2015), <https://www.forensicmag.com/article/2015/11/dna-mixture-calculation-method-just-%E2%80%98random-number-generator%E2%80%99-says-new-study> [<https://perma.cc/7RMS-SLPJ>].

64. Shaer, *supra* note 61.

65. TrueAllele, *TrueAllele Process Overview*, YOUTUBE (May 1, 2013), <https://www.youtube.com/watch?v=OU29b5sW88Y>.

66. CYBERGENETICS, TRUEALLELE TECHNOLOGY: COMPUTER INTERPRETATION OF DNA EVIDENCE 4, https://www.cybgen.com/solutions/brochures/lab_brochure.pdf [<https://perma.cc/HY3Z-R5XF>].

67. *Casework*, CYBERGENETICS <https://www.cybgen.com/products/casework.shtml> [<https://perma.cc/V9GD-EPTS>].

68. TrueAllele, *Science: Indiana v. Forest*, YOUTUBE (Dec. 21, 2016), <https://www.youtube.com/watch?v=RNF6JtYikiE> (last visited Dec. 20, 2017).

69. *Id.*

interpreting DNA mixtures”⁷⁰) results were used as evidence for a criminal case was in *Commonwealth v. Foley*,⁷¹ a 2009 first-degree murder case in Pennsylvania. Foley was tried for the murder of the estranged husband of the woman with whom he was living.⁷² DNA evidence from under the victim’s fingernail contained the DNA of two individuals, the victim and the person presumed to have murdered him.⁷³ The DNA sample was tested in a Federal Bureau of Investigation lab, and the data from the tests performed at the FBI lab were used by three different experts to develop testimony about the DNA’s significance.⁷⁴

The three experts — Mark Perlin, an FBI forensic scientist, and a third scientist — all agreed that Foley’s DNA profile was consistent with the DNA found in the sample, but each testified to radically different probabilities that someone other than Foley would match the DNA found in the sample.⁷⁵ The FBI forensic scientist testified that the probability of another person contributing that portion of the DNA sample was 1 in 13,000; the other scientist testified that the probability was 1 in 23 million; and Dr. Perlin testified that the odds were 1 in 189 billion.⁷⁶

Foley argued that Dr. Perlin’s testimony should be ruled inadmissible for failing the *Frye* test.⁷⁷ Under Pennsylvania’s formulation of the *Frye* test, “novel scientific evidence is admissible if the methodology that underlies the evidence has general acceptance in the relevant scientific community.”⁷⁸ To oppose the admission of evidence, Foley needed to show that the scientific evidence being introduced was novel by demonstrating “that there is a legitimate dispute regarding the reliability of the expert’s conclusions.”⁷⁹ If he had successfully established the evidence as novel, the prosecution team would have had the burden of showing that “‘the expert’s methodology has general acceptance in the relevant scientific community’ despite the legitimate dispute.”⁸⁰

Unfortunately for Foley, the trial court ruled that the technique used by Dr. Perlin was a refined application of the previously accepted “product rule” method for calculating probabilities in forensic DNA

70. MARK W. PERLIN, CYBERGENETICS, COMMONWEALTH OF PENNSYLVANIA V. KEVIN JAMES FOLEY, <https://www.cybgen.com/information/presentations/2012/ISHI/Perlin-Commonwealth-of-Pennsylvania-v-Kevin-James-Foley/poster.pdf> [<https://perma.cc/D493-SGPY>]

71. 38 A.3d 882 (Pa. Super. Ct. 2012).

72. *Id.* at 885.

73. *Id.* at 887.

74. *Id.*

75. *See id.*

76. *Id.*

77. *Id.* at 888.

78. *Id.* (quoting *Betz v. Pneumo Abex LLC*, 998 A.2d 962, 972 (Pa. Super. Ct. 2010) (en banc)).

79. *Id.*

80. *Id.* (quoting *Betz v. Pneumo Abex LLC*, 998 A.2d 962, 972 (Pa. Super. Ct. 2010) (en banc))

analysis.⁸¹ Because the Supreme Court of Pennsylvania had previously upheld the admissibility of evidence based on the product rule, the trial court found that Dr. Perlin's method was generally accepted.⁸²

On appeal, the appellate court also found that Dr. Perlin's testimony was not novel and concluded, "the trial court did not abuse its discretion in admitting the testimony."⁸³ The appellate court rejected Foley's arguments that TrueAllele should be considered novel because it had never previously been used to analyze a mixed sample of DNA and because "no outside scientist can replicate or validate Dr. Perlin's methodology because his computer software is proprietary."⁸⁴

The court found it irrelevant whether TrueAllele had previously been used in court cases, arguing that whether a scientific method was disputed among scientists (and thus whether a method is novel) is not determined by whether a court has previously chosen to admit the evidence.⁸⁵ In any case, the court stated, TrueAllele was at the time being used for other purposes such as World Trade Center victim identification, as well as to build the United Kingdom's National DNA database, which undermined Foley's arguments the technology was not being used.⁸⁶

Regarding Foley's argument that TrueAllele's refusal to disclose its source code prevented its validation, the court stated, "scientists can validate the reliability of a computerized process even if the 'source code' underlying that process is not available to the public. TrueAllele is proprietary software; it would not be possible to market TrueAllele if it were available for free."⁸⁷ Furthermore, the court said, TrueAllele had been the subject of validation studies published in peer-reviewed journals, which indicated the contents of the validation studies had been "reviewed by other scholars in the field."⁸⁸ The court failed to note, however, that both of the studies it cited for this point had been authored by Perlin and his colleagues, which some critics argue undermines the studies' validity.⁸⁹ See Part IV, *infra*, for additional discussion of these issues surrounding the debatable significance of Dr. Perlin's involvement in TrueAllele's validation studies.

81. *Id.*

82. *Id.*

83. *Id.*

84. *Id.* at 888–89.

85. *See id.* at 889.

86. *See id.*

87. *Id.*

88. *Id.* at 889–90.

89. *See* PCAST REPORT, *supra* note 1, at 80; *see generally* Daniele Mandrioli et al., *Relationship Between Research Outcomes and Risk of Bias, Study Sponsorship, and Author Financial Conflicts of Interest in Reviews of the Effects of Artificially Sweetened Beverages on Weight Outcomes: A Systemic Review of Reviews*, PLOS ONE (Sept. 8, 2016) <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0162198> [<https://perma.cc/N4UA-2FY2>].

Ultimately, the court concluded Perlin's evidence was admissible in Foley's trial and Foley was convicted for first-degree murder.⁹⁰ The DNA evidence based on TrueAllele analyses that Perlin helped introduce was described as a trump over Foley's arguments that the jury was prejudiced.⁹¹ TrueAllele hailed the decision as one that "supports the legitimacy and advancement of the new computerized TrueAllele methodology for performing DNA match probability calculations."⁹²

B. Subsequent Unsuccessful Challenges to the Use of DNA Analysis Algorithms

The decision in *Commonwealth v. Foley* by no means ended arguments related to the admissibility of TrueAllele evidence. In Michael Robinson's Pennsylvania trial for a 2013 double homicide, his lawyers first argued Perlin's refusal to turn over the source code for TrueAllele violated Robinson's rights under the Confrontation Clause, saying that without knowing the details of how TrueAllele works, they would be unable to properly challenge the evidence.⁹³ Robinson's attorney, Ken Haber, argued, "The witness in this case is a computer You can't cross-examine a computer. The Constitution demands, and justice requires, we be permitted to find out what the computer is doing to come up with its answer."⁹⁴ But the judge denied the motion, ruling it could harm Cybergenetics if the company were required to disclose the source code for TrueAllele.⁹⁵

After the attempt to challenge the admissibility of TrueAllele evidence based on the Confrontation Clause failed all the way up the state Supreme Court, Robinson's lawyers filed a motion alleging TrueAllele's methodology, given the facts in his particular case, failed to meet the PCAST report's threshold for reliability and general acceptance.⁹⁶ This

90. *Id.* at 893.

91. See Sam Kusic, *Foley Sentenced to Life in Prison*, INDIANA GAZETTE, June 2, 2009, at 12.

92. *Pennsylvania Appeals Court Affirms Cybergenetics TrueAllele Admissibility For DNA Mixture Evidence In Foley Case*, CYBERGENETICS (Jan. 4, 2012), <https://www.cybgen.com/information/newsroom/2012/jan/Pennsylvania-appeals-court-affirms-Cybergenetics-TrueAllele-admissibility-for-DNA-mixture-evidence-in-Foley-case.shtml> [<https://perma.cc/249T-2FPW>].

93. See Paula Reed Ward, *Defense Tries New Tack to Fight DNA Evidence in Double Homicide Case*, PITTSBURG POST-GAZETTE (Nov. 21, 2016), <http://www.post-gazette.com/local/region/2016/11/21/Defense-tries-new-tack-to-fight-DNA-evidence-in-double-homicide-case/stories/201611210019> [<https://perma.cc/VN8Q-LJNK>].

94. Paula Reed Ward, *Legal Question: How Do You Cross-Examine a Computer?*, PITTSBURG POST-GAZETTE (Aug. 29, 2016), <http://www.post-gazette.com/news/science/2016/08/29/Legal-question-how-do-you-cross-examine-a-computer/stories/201608280021> [<https://perma.cc/N3VK-CESM>].

95. *Id.*

96. Ward, *supra* note 93.

attempt was also rejected, and the TrueAllele evidence was ultimately admitted as part of the case against Robinson at his trial.⁹⁷

This pattern has continued in a number of jurisdictions. As of January 2017, defendants in at least 7 states have sought access to TrueAllele's code for review as part of their trials and been denied access in the face of Cybergenetics's opposition.⁹⁸

C. Use in Exoneration Cases

Cybergenetics promotes TrueAllele as a tool for prosecutors and defense attorneys alike and has encouraged its use in exoneration cases.⁹⁹ In 2013, Mark Perlin agreed to provide free TrueAllele tests for Darryl Pinkins, who had been convicted of a gang rape.¹⁰⁰ After running the tests, Perlin stated he was "incredibly confident" that the results excluded Pinkins as a suspect based on the DNA evidence provided.¹⁰¹ This new evidence ultimately led to prosecutors admitting they were wrong and agreeing to overturn Pinkins' conviction.¹⁰²

When Cybergenetics has been challenged, Perlin has habitually invoked TrueAllele's usage in exoneration cases, arguing such cases help demonstrate TrueAllele's reliability.¹⁰³ Innocence Project leaders have also been among the more vocal defenders of Cybergenetics' refusal to release its code. As Greg Hampikian, the leader of the Idaho Innocence Project, sees it, "Microsoft Excel doesn't release its code either, but we can test it and see that it works, and that's what we care about," so full transparency on the part of Cybergenetics is unnecessary in his eyes.¹⁰⁴ Yet Microsoft Excel is not used in criminal cases to support claims that a specific individual did or did not commit a crime. Moreover, the type of

97. See Paula Ward & Tortsen Ove, *Jury Acquits Duquesne Man in Double Homicide Case*, PITTSBURG POST-GAZETTE (Feb. 7, 2017), <http://www.post-gazette.com/local/east/2017/02/07/Jury-gets-Duquesne-double-homicide-case/stories/201702070160> [https://perma.cc/ZY8R-KVML].

98. Lael Henterly, *The Troubling Trial of Emanuel Fair*, SEATTLE WEEKLY, <http://www.seattleweekly.com/news/the-troubling-trial-of-emanuel-fair/> [https://perma.cc/H74S-AWB9].

99. CYBERGENETICS, *supra* note 66, at 2.

100. 48Hours, *Guilty Until Proven Innocent*, CBS NEWS, <http://www.cbsnews.com/news/darryl-pinkins-roosevelt-glenn-convicted-in-1989-rape-guilty-until-proven-innocent/> [https://perma.cc/9HL6-WJDE].

101. *Id.*

102. *Id.*

103. See, e.g., Letter from Michael Perlin, Chief Sci. and Exec. Officer, Cybergenetics, to John Holdren, PCAST Co-Chair, Re: Report to the President on "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 2 (Sept. 16, 2016), <https://www.cyngen.com/information/newsroom/2016/sep/files/letter.pdf> [https://perma.cc/J9EP-Q8H7].

104. Lauren Kirchner, *Where Traditional DNA Testing Fails, Algorithms Take Over*, PROPUBLICA, <https://www.propublica.org/article/where-traditional-dna-testing-fails-algorithms-take-over> [https://perma.cc/QYX6-LET9].

mathematical calculations Microsoft Excel is used for are entirely free of human judgment and completely replicable by any other program — crucial differences from TrueAllele.

While TrueAllele has had a significant amount of success in the court system, helping to obtain both convictions and exonerations based on its DNA analyses, a number of critics have raised important concerns about TrueAllele and its use in criminal justice proceedings. Some of the concerns raised by defendants in the cases discussed in this Part have been picked up by broader commentary, and additional concerns have been raised by organizations such as PCAST. Part IV, below, describes these criticisms more fully.

IV. CONCERNS AND CRITICISMS

There are concerning potential issues with the use of algorithmic programs for complex DNA mixture analysis. The scientific validity of their methods remains unclear and openly questioned by some. This problem is compounded by the fact that widely used programs like TrueAllele refuse to disclose their methods, asking judges and juries to rely on their numbers with no way of knowing how TrueAllele reached those conclusions or challenging the methods used to interpret the results. This Part addresses each of these groups of concerns in turn.

A. Unestablished Scientific Validity

First and foremost are concerns about the scientific validity of probabilistic genotyping algorithms. PCAST's report on forensic science cautions that studies establishing the validity of complex mixture analysis remain scarce and states, at this time, it considers objective methods of analysis to have been established only under very limited circumstances (namely, "a three-person mixture in which the minor contributor constitutes at least 20 percent of the intact DNA in the mixture").¹⁰⁵ PCAST considered analyses under limited circumstances to be reliable based on specific published evidence in studies conducted on specific mixture types.¹⁰⁶ Mixtures with a different number of contributing individuals, different ratios of DNA mixtures, or low amounts of DNA have not been established to produce reliable results.¹⁰⁷ The PCAST report highlighted that the difficulty of reliably interpreting samples increases when the number of contributors increases, or when the proportion of the sample attributable to a minor contributor decreases, and specifically

105. PCAST REPORT, *supra* note 1, at 8.

106. *See id.* at 80.

107. *See id.* at 81.

states that scientific validity has only been established within the specific range for which experimental evidence of validity is available.¹⁰⁸

One of the problems with establishing valid results is the lack of third-party independent scientific studies. The PCAST report emphasizes the importance of third-party groups conducting validation studies, rather than relying on studies conducted by the developers of the algorithmic analysis programs.¹⁰⁹ TrueAllele's validation studies have been conducted primarily by individuals associated with the company in some way,¹¹⁰ making them questionable in the eyes of the broader scientific community.

Dr. Perlin of Cybergenetics disputes the need for third-party validation studies and has objected to the PCAST report's implied criticism of TrueAllele. In a letter protesting the PCAST report's call for independently authored reviews, he argued that peer review was sufficient to mitigate conflicts of interest.¹¹¹

Contrary to Perlin's arguments, and to the professed beliefs of the court in *Commonwealth v. Foley*, having internal validation studies published in peer-reviewed journals does not mean that the scientific community has debated and accepted the science involved; it merely indicates that the peer reviewers did not identify any disqualifying characteristics of the study as it was described by the paper, such as obvious methodological errors or inaccurate analysis of the results reported to the journal. Numerous problems exist with peer review, including selective reporting and publication, the frequent inability to reproduce the reported results,¹¹² and a lack of transparency surrounding the data and methods used to produce the results of peer-reviewed studies.¹¹³ Peer review of validation studies conducted by interested parties is not the equivalent of rigorous third-party evaluation studies for the purposes of general acceptance in the scientific community.¹¹⁴

Despite the lack of accepted scientific validation, Cybergenetics markets TrueAllele for mixtures with more than 3 individuals. In a New York state trial, TrueAllele was used as the sole evidence linking the 19-year-old defendant to the gun used in a crime.¹¹⁵ The gun had been han-

108. *See id.*

109. *Id.* at 79.

110. *See* PCAST REPORT, *supra* note 1, at 80; *see also, e.g.*, S.A. Greenspoon et al., *Establishing the Limits of TrueAllele Casework: A Validation Study*, 60 J. FORENSIC SCI. 1263 (2015).

111. *See* Perlin, *supra* note 103, at 1.

112. *See generally* Monya Baker, *1,500 Scientists Lift the Lid on Reproducibility*, 533 NATURE 452, 452–54 (2016).

113. *See* Jelte Wicherts, *Peer Review Quality and Transparency of the Peer-Review Process in Open Access and Subscription Journals*, PLOS ONE (Jan. 29, 2016) <https://doi.org/10.1371/journal.pone.0147913> [<https://perma.cc/ZQ4H-7N4L>].

114. *See* *People v. Collins*, 15 N.Y.S.3d 564, 581 (N.Y. Sup. Ct. 2015).

115. *Kirchner, supra* note 104.

dled by at least four people and potentially as many as five or six.¹¹⁶ TrueAllele was asked to analyze this “touch DNA” to link the defendant to the gun.¹¹⁷

Perlin testified at trial that a match between the DNA on the gun and the DNA of the defendant, who was black and Hispanic, was “1.78 trillion times more probable than a coincidental match to an unrelated African American person” and “892 billion times more probable than a coincidental match to an unrelated Hispanic person.”¹¹⁸ It’s unclear how these highly specific and extremely certain numbers can be reconciled with the documented difficulties with validly and reliably interpreting highly complex samples. Nevertheless, despite being the only physical evidence connecting the defendant with the gun, this testimony was persuasive enough to establish a connection and convict the defendant of criminal possession of a weapon, reckless endangerment, and menacing a police officer.¹¹⁹

Perlin has claimed that external empirical testing of TrueAllele is unnecessary because, he says, it is mathematically impossible for TrueAllele’s likelihood-ratio approach to produce a false positive.¹²⁰ It is unclear what (if any) support Perlin has for this extreme claim, as he does not appear to have provided proof of it to PCAST.¹²¹ PCAST responded by saying:

While likelihood ratios are a mathematically sound concept, their application requires making a set of assumptions about DNA profiles that require empirical testing. Errors in the assumptions can lead to errors in the results. To establish validity with a range of parameters, it is thus important to undertake empirical testing with a *variety* of samples in the relevant range.¹²²

This response highlights the concerns that remain about the subjective decisions embedded within TrueAllele and similar applications. Without external validation and rigorous examination of the underlying assumptions of the algorithms, the reliability of such methods cannot be firmly established. PCAST was not persuaded by Perlin’s argument and declined in its report addendum to change the view expressed in its orig-

116. *Id.*

117. *Id.*

118. *Id.*

119. *Id.*

120. See PCAST ADDENDUM, *supra* note 53, at 8.

121. See *id.*

122. *Id.* at 8–9 (internal citations omitted).

inal report that recommended additional independent scientific studies prior to acceptance of TrueAllele and related technologies.¹²³

Varying assumptions about DNA profiles can produce crucial differences in interpreting results. Different programs incorporate subtly different choices into their algorithms about how to interpret data, which can yield different results when analyzing the exact same complex mixture.¹²⁴ For example, an error in determining how many individuals' DNA were present within a mixture could have rippling effects through subsequent stages of analysis, affecting decisions about differentiating signal from noise and distinguishing which DNA came from which individual. Differences in assumptions about how to adjust for different quantities of DNA or how to evaluate whether the people in a mixture were related could also impact the ultimate results. Due to a lack of comparative studies, the exact nature of these potential differences based on programs' varying assumptions remains unclear.¹²⁵

Because of the effects these interpretive choices can have on the output of the algorithms, even analyses conducted using algorithmic probabilistic genotyping are not free of all subjectivity. Itiel Dror, who co-authored the famous 17-expert study demonstrating the variability of expert interpretation of DNA mixtures using subjective techniques, disagrees with Perlin's assertion that probabilistic genotyping algorithms' analyses of DNA mixtures is entirely objective.¹²⁶ "Using software doesn't solve the problem, because the human biases, assumptions, and discretions go into the software," he said in an interview.¹²⁷ "The software has human biases; to see what the biases are, we need to look at the software to see what it's doing."¹²⁸ Thus, the issues of establishing scientific validity and achieving transparency are intertwined.

B. Lack of Transparency

The lack of transparency about TrueAllele's exact methods and source code has raised concerns about whether TrueAllele's results may be biased or unreliable. Unreliability and accusations of bias have plagued similarly opaque proprietary algorithms purported to assess defendants' risk of future crime for use in sentencing, bond, and probation

123. *See id.* at 8.

124. *See* PCAST REPORT, *supra* note 1, at 79.

125. *Id.* at 80.

126. *See* Henterly, *supra* note 98.

127. *Id.*

128. *Id.*

decisions.¹²⁹ There is growing concern about the impacts such algorithms may have on individuals within the criminal justice system.¹³⁰

Algorithmic systems designed to interpret complex DNA mixture analyses are not exempt from the potential for errors in their code. STRmix, frequently considered TrueAllele's biggest rival, publicly acknowledged finding two errors in its source code.¹³¹ The coding error was only found after prosecutors in a trial sought to have faulty STRmix results admitted as evidence.¹³² STRmix's miscode only applied to a particular category of cases which make up a small percentage of all the samples tested; nonetheless, dozens of cases were affected and required the generation of a new set of likelihood ratios for those DNA interpretations.¹³³ TrueAllele might also have errors in its code that have not been detected yet due to Cybergenetics' lack of transparency. TrueAllele refuses to make its source code available to any third party, unlike STRmix, which makes its source code available for inspection by defense expert witnesses who sign a confidentiality agreement.¹³⁴

Cybergenetics has vigorously opposed any attempt to force it to disclose its algorithms or underlying code. Perlin argues that because Cybergenetics permits anyone who wishes to try out the software to have a free trial, anyone can verify TrueAllele's validity simply by running known samples through and seeing if TrueAllele produces the correct results. The former lab director of Kern County, California, supported Perlin's assertion that this was enough, saying, "I know that if I give it known samples, it works as expected, . . . so when I give it unknown samples, I have no reason to believe it wouldn't work the same way."¹³⁵

But running a limited number of samples — even the 40 known samples used by labs such as Kern County's to test TrueAllele for validity¹³⁶ — may not be enough to detect all types of errors. Small sample sizes lack the necessary statistical power to detect flaws in the code that

129. See Julia Angwin et al. *Machine Bias*. PROPUBLICA (May 23, 2016), <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [<https://perma.cc/5Z3H-MRA5>].

130. See *id.*

131. Ward, *supra* note 94.

132. Kirchner, *supra* note 104.

133. David Murray, *Queensland Authorities Confirm 'Miscode' Affects DNA Evidence in Criminal Case*, THE COURIER-MAIL (Mar. 20, 2015), <http://www.couriermail.com.au/news/queensland/queensland-authorities-confirm-miscode-affects-dna-evidence-in-criminal-cases/news-story/833c580d3f1c59039efd1a2ef55af92b?nk=4164b9d1e4e66d5dcf2bea812edd1d47-1494599443> [<https://perma.cc/9B8V-4HYJ>].

134. *Access to STRmix Software by Defence Legal Teams*, STRMIX (Apr. 2016), <https://strmix.esr.cri.nz/assets/Uploads/Defence-Access-to-STRmix-April-2016.pdf> [<https://perma.cc/8EWR-L5TE>].

135. Stephanie M. Lee, *People Are Going to Prison Thanks to DNA Software — But How It Works Is Secret*, BUZZFEED NEWS, <https://www.buzzfeed.com/stephaniemlee/dna-software-code> [<https://perma.cc/V6JJ-CBZV>].

136. *Id.*

might only affect samples with particular combinations of factors, for example. Critics of complex DNA analysis algorithms have pointed out that specialized populations — for example, the uniquely genetically insulated population of Hasidic Jews — may pose unique challenges for which the algorithms have not been validated.¹³⁷ The flaws in the code of STRmix were only discovered after it was used in thousands of cases in Australia and New Zealand¹³⁸ — 40 test samples could very conceivably miss subtle but important flaws in TrueAllele’s underlying code.

TrueAllele is patented,¹³⁹ which provides substantial protections. But Perlin argues Cybergenetics lacks the resources to litigate a patent dispute, and so must keep their code secret¹⁴⁰ due to the “highly competitive commercial environment” of probabilistic genotyping programs.¹⁴¹ He says that while Cybergenetics has not published TrueAllele’s source code or engineering details, it has published papers on the theory behind the program and the math involved.¹⁴² Furthermore, Perlin insists, “Source code is not used to assess forensic software reliability.”¹⁴³

TrueAllele’s critics disagree with Perlin’s assertion that the source code of TrueAllele is unimportant to disclose. In October 2015, the Electronic Privacy Information Center (EPIC) filed state public records requests in six states that use TrueAllele (California, Louisiana, New York, Ohio, Pennsylvania, and Virginia).¹⁴⁴ EPIC has thus far obtained validation study information from Virginia and contracts, technical specifications, and user manuals from Pennsylvania.¹⁴⁵ California, Louisiana, Pennsylvania, and Virginia have informed EPIC that they do not have access to the TrueAllele source code.¹⁴⁶ EPIC has announced that it will continue to seek TrueAllele’s source code because of the importance of algorithmic transparency and EPIC’s interest in open government and a fair criminal justice system.¹⁴⁷

Professor Erin Murphy has also argued that obtaining the code itself is crucial in order to fully evaluate TrueAllele. She writes, “Just as

137. Lauren Kirchner, *Traces of Crime: How New York’s DNA Techniques Became Tainted*, N.Y. TIMES, <https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html?mcubz=3> (last visited Dec. 20, 2017).

138. See Lee, *supra* note 135.

139. *Patents*, CYBERGENETICS, <https://www.cybgen.com/information/patents.shtml> [<https://perma.cc/R2MU-X85K>].

140. Lee, *supra* note 135.

141. David Kravets, *Secret Source Code Pronounces You Guilty as Charged*, ARS TECHNICA, <https://arstechnica.com/tech-policy/2015/10/secret-source-code-pronounces-you-guilty-as-charged/> [<https://perma.cc/L58L-PZQT>].

142. Shaer, *supra* note 64.

143. Kravets, *supra* note 141.

144. See *State FOIA: Secret DNA Forensic Source Code*, EPIC, <https://epic.org/state-policy/foia/dna-software/> [<https://perma.cc/DVK9-ES4J>].

145. *Id.*

146. *Id.*

147. See *id.*

courts would not accept opinions from witnesses not shown to have qualifications as an expert, so, too, should courts not accept opinions from digital ‘experts’ without probing the ‘qualifications’ of the technology.”¹⁴⁸ Without the disclosure of the code, argues Professor Murphy, courts and defense attorneys cannot fully consider the technology’s “qualifications” to decide whether it ought to be accepted by the court.¹⁴⁹

A lack of transparency surrounding analytic algorithms also gives prosecutors and law enforcement an opportunity to strategize to receive the answers they want, rather than the answers that would be independently determined to be most accurate and valid. Given the potential variations in interpretative choices made by various DNA profiles, without transparency there is an opportunity to shop around for desirable results from different programs. A CEO of a forensic consulting firm criticized the lack of transparency by saying, “The biggest issue is there is no truly independent assessment of TrueAllele or other programs They don’t work the same, and some are better at certain profiles and the community doesn’t know the benefits and weaknesses.”¹⁵⁰

These potential differences between programs can mean prosecutors and police might keep testing with different programs until they get the results that best support the case they are trying to make. In the case *New York v. Hillary*, after the state police and Cybergenetics were both unable to conclude that Hillary’s DNA was part of the DNA mixture found on the victim, the police asked STRmix’s developers to try to analyze the data as well.¹⁵¹ Hillary’s attorney characterized this move as follows: “(The DNA) came back that it was inconclusive and said it wasn’t Nick [Hillary] and they weren’t satisfied with that so they took it to New Zealand [to STRmix] to get what they wanted . . . something more to their liking.”¹⁵² After STRmix’s results were found to potentially implicate Hillary, his defense attorneys moved to have the evidence excluded.¹⁵³

The judge ultimately chose to exclude the evidence on the grounds that the New York state lab had not conducted internal validation studies on STRmix.¹⁵⁴ But this type of repeated testing in search of particular results raises significant concerns. Without access to the underlying

148. ERIN MURPHY, *INSIDE THE CELL: THE DARK SIDE OF FORENSIC DNA* 299 (2015).

149. *See id.*

150. Henterly, *supra* note 98.

151. Notice of Motion to Preclude at 8–9, *New York v. Hillary*, No. 2015-15, (N.Y. St. Lawrence Cty. Ct., May 31, 2016).

152. W.T. Eckert, *Hillary Defense Seeks to Keep DNA from Trial, Questions Testing Techniques*, WATERTOWN DAILY TIMES, <http://www.watertowndailytimes.com/news05/hillary-defense-seeks-to-keep-dna-from-trial-questions-testing-techniques--20160603> [https://perma.cc/K42D-AXN3].

153. *See id.*

154. Decision & Order at 10, *New York v. Hillary*, No. 2015-15, (N.Y. St. Lawrence Cty. Ct. Aug. 26, 2016).

source code for TrueAllele, it is impossible to analyze which differences in programming decisions may have resulted in differing results from the same samples. In contrast, STRmix provides its code to defense attorneys when requested,¹⁵⁵ although it remains a proprietary program otherwise.

The lack of independent studies establishing scientific validity for many uses of algorithmic DNA interpretation technologies and the lack of transparency about the subjective decisions embodied in these programs' codes exacerbate the issues caused by each. Without transparency, it is more difficult to rigorously evaluate scientific validity; without rigorous studies, it is more difficult to challenge specific issues caused by a lack of transparency. Because of the potential impacts these issues may have on defendants' outcomes, these issues should be resolved. Part V, below, discusses some options for how the criminal justice system might respond to these important criticisms and address these issues.

V. POTENTIAL RESPONSES

Given these significant problems, what can be done to address the potential issues related to the use of TrueAllele and similar technologies? Probabilistic genotyping algorithms offer too many potential benefits to simply be dismissed, despite the significant concerns associated with their use in the judicial system. Having grown accustomed to using DNA evidence, prosecutors and law enforcement officials are not going to stop trying to utilize such evidence in criminal cases. That leaves two apparent alternatives to the current problematic state of affairs: improve current non-probabilistic methods of interpretation or find a way to correct the current deficiencies in probabilistic genotyping algorithm programs.

Improving non-probabilistic methods of interpretation for complex DNA mixtures is a viable short-term option. The PCAST report acknowledged that the specific rules outlined in a recent Texas working group paper¹⁵⁶ could potentially address a number of problems with the existing method.¹⁵⁷ But this method is fundamentally flawed given the necessary subjective decisions and potential for human error. No matter the improvements in handling and processing the samples using non-probabilistic methods, the lingering subjectivity will ultimately result in unreliability. A more sustainable long-term solution would require the use of methods that utilize probabilistic genotyping.

Probabilistic genotyping algorithms should not be accepted for uses that have not been established as scientifically valid. PCAST recom-

155. Ward, *supra* note 94.

156. See Bieber et al, *supra* note 21, at 136.

157. PCAST REPORT, *supra* note 1, at 82.

mends that “[w]hen considering the admissibility of testimony about complex mixtures (or complex samples), judges should ascertain whether the published validation studies adequately address the nature of the sample being analyzed (e.g., DNA quantity and quality, number of contributors, and mixture proportion for the person of interest).”¹⁵⁸ In other words, just because TrueAllele has been potentially validated and accepted for three-person mixtures with a minority contribution of at least 20%, that does not mean that TrueAllele should be automatically accepted as accurate for samples with small amounts of DNA and more than three people (such as a gun touched by four, five, or six people).

Additional studies should be conducted by independent researchers to evaluate whether a given method is valid for particular types of samples. The PCAST report calls for both large-scale scientific studies on common sets of samples, to establish foundational validity of a method, as well as internal developmental validation studies by individual forensic laboratories to assess the as-applied validity in a particular setting.¹⁵⁹ With additional studies, many of the scientific validity concerns can potentially be resolved, either by establishing independent measures of validity and accuracy through empirical, external studies using a variety of samples or by firmly demonstrating the unsuitability of existing programs for particular types of samples and mixtures. In October 2017, the National Institute of Standards and Technology announced that it would conduct a “scientific foundation review” of forensic methods involving DNA analysis, including the analysis of mixtures.¹⁶⁰ This study may eventually lead to better understanding of the limits of reliability for samples containing mixtures of DNA or low quantities of DNA.¹⁶¹

This still leaves transparency concerns to be addressed. One option is to simply exclude evidence produced by black-box algorithms. Professor Erin Murphy argues, “[C]ourts should disallow statistical evidence generated by probabilistic software whose operators refuse to reveal their code.”¹⁶² While this may seem like an extreme position, there are strong moral and legal reasons for refusing to admit such evidence. As the Electronic Privacy Information Center has argued to Congress when advocating for greater transparency, “Secrecy of the algorithms used to determine guilt or innocence undermines faith in the criminal justice system.”¹⁶³ Policy principles are worth considering when crafting judi-

158. See PCAST ADDENDUM, *supra* note 53, at 9.

159. PCAST REPORT, *supra* note 1, at 82–83.

160. *NIST to Assess the Reliability of Forensic Methods for Analyzing DNA Mixtures*, NIST (Oct. 03, 2017) <https://www.nist.gov/news-events/news/2017/10/nist-assess-reliability-forensic-methods-analyzing-dna-mixtures> [<https://perma.cc/G5HG-CQNP>].

161. See *id.*

162. MURPHY, *supra* note 148, at 300.

163. Letter from Marc Rotenberg & Caitriona Fitzgerald, President and Policy Dir., Elec. Privacy Information Ctr., to the Honorable Trey Gowdy, Chair, House Comm. on the Judiciary, Subcomm. on Crime, Terrorism, Homeland Sec., & Investigations 3,

cial rules. Preserving faith in the criminal justice system is a principle worth protecting by creating fair and open processes for establishing whether the evidence truly indicates that a defendant is guilty (or not).

Furthermore, defendants' inability to meaningfully interrogate the algorithm being used against them carries implications for their Sixth Amendment right to be confronted with the witnesses against them¹⁶⁴ and the Fifth and Fourteenth Amendment rights to due process. Transparency is a crucial component of due process.¹⁶⁵ Protecting a private company's desire not to defend its property rights in court is a strange justification for permitting the inclusion of unverifiable, unchallengeable evidence. One expert witness who has been critical of TrueAllele's refusal to share its code commented that while both the right to confront one's accusers and the right to protect one's property are important, he believed the right to confront ought to outweigh Cybergenetics' desire to protect its property.¹⁶⁶

It is not unreasonable to believe courts might take this position. Unlike past motions, which sought to compel a private party to provide information about their trade secrets, prohibiting evidence unless the underlying algorithms are disclosed permits Cybergenetics to make a choice: do they choose to disclose their source code, which they claim would harm their business, or do they choose to be excluded from their primary market (the criminal justice system) because they do not meet the requirements?

Some courts have already been willing to exclude evidence produced by black-box DNA analysis. In *People v. Collins*,¹⁶⁷ the judge, in deciding to exclude evidence based on a forensic statistical tool ("FST") developed by the Office of the Chief Medical Examiner ("OCME"),¹⁶⁸ highlighted the fact that the tool was a "black box":

[T]he fact that the FST software is not open to the public, or to defense counsel, is the basis of a more general objection. This court understands the city's desire to control access to computer programming that was developed at great cost. But the FST is, as a result, truly a "black box" — a program that cannot be used by de-

<https://epic.org/testimony/congress/EPIC-HJC-ForensicEvidence-Mar2017.pdf>
[<https://perma.cc/4L5J-EFLR>].

164. U.S. CONST. amend. VI.

165. Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 20 (2014).

166. See Kirchner, *supra* note 104.

167. 15 N.Y.S.3d 564 (N.Y. Sup. Ct. 2015).

168. *Id.* at 578.

fense experts with theories of the case different from the prosecution's.¹⁶⁹

The judge also took note of a defense expert's testimony stating that OCME had failed to conduct a proper review with independent experts to establish FST's general acceptance within the scientific community:

Dr. Rosenberg stated that proper peer review of a software package like the FST requires the submission of that package to the independent experts — something not done by OCME with the FST. Dr. Rosenberg further opined that publication in peer-reviewed journals is not the equivalent of general acceptance in the relevant community. That must be judged from the results of publication.¹⁷⁰

Based on these concerns about the black box nature of FST and the lack of independent expert evaluation, the judge concluded, “the FST is not generally accepted in the DNA scientific community.”¹⁷¹ Accordingly, the evidence was excluded from the trial for failing to meet *Frye*'s requirements.¹⁷²

The source code of FST was finally revealed to an outside party for the first time in June 2016, when a federal judge in the Southern District of New York granted the defense team access.¹⁷³ Despite the government's argument that the source code was “proprietary and copyrighted,” Judge Valerie Caproni wrote in her order,

FST is a relatively new tool that has not been extensively examined or tested in federal court, and the results obtained from the use of FST on DNA samples recovered from crime scenes are potentially devastating to a criminal defendant. The fact that the results obtained from use of the FST can be devastating to a criminal defendant increases the need of the Court to be diligent about FST's reliability prior to admitting FST results into evidence.¹⁷⁴

169. *Id.* at 580.

170. *Id.* at 581.

171. *Id.* at 582.

172. *Id.* at 587.

173. See *United States v. Johnson*, No. 1:15-cr-00565 (S.D.N.Y. June 7, 2016) (order granting request for subpoena for disclosure of FST source code).

174. *Id.*

After the disclosure of the source code, a defense expert witness reportedly wrote in an affidavit that the program excluded potentially important data from its calculations, which could unpredictably affect the reported likelihood of a defendant's DNA being present in the analyzed mixture of DNA.¹⁷⁵ The two parties filed briefs on whether the FST analyses should be admitted under the *Daubert* standard, but no decision was reached on the subject because the prosecutors withdrew the evidence prior to the admissibility hearing.¹⁷⁶ ProPublica filed a motion in the case seeking permission to intervene in the case and requesting that Judge Caproni lift the protective order preventing disclosure of FST's source code.¹⁷⁷ The protective order was lifted and ProPublica posted the source code publicly.¹⁷⁸

FST's legal troubles continued in September 2017, when the Legal Aid Society and the Federal Defenders of New York wrote a letter to the state's inspector general alleging FST and a related technique were "unreliable" and based on "unsound statistical evidence."¹⁷⁹ The coalition of defense attorneys raised concerns that flaws in the testing may have led to wrongful convictions and innocent defendants choosing to plead guilty when told there was DNA evidence against them.¹⁸⁰ The letter specifically cited the secrecy surrounding the program and how it was developed and used when calling for investigation into the techniques used to analyze DNA evidence containing complex mixtures or low quantities of DNA.¹⁸¹ Although the New York medical examiner's office had already begun to transition away from FST to other tools, these lingering questions about the technology call into question thousands of cases.¹⁸²

Defenders of TrueAllele might argue that *Collins* and the problems with FST are distinguishable from cases involving TrueAllele because Cybergenetics permits defense attorneys to use a free trial of TrueAllele.¹⁸³ Admittedly, this may provide slightly more access than was available to the defense in *Collins*. But without the source code and details of the underlying algorithms, defense attorneys lack the information and capabilities necessary to fully explore potential alternative scenarios and investigate potential weaknesses in the program. Ultimately,

175. Kirchner, *supra* note 137.

176. See Docket, *United States v. Johnson*, No. 1:15-cr-00565 (S.D.N.Y. June 7, 2016).

177. See Memorandum in Support of Application by ProPublica for Leave to Intervene, Lift the Protective Order and Unseal Judicial Records at 1, *United States v. Johnson*, No. 1:15-cr-00565 (S.D.N.Y. June 7, 2016).

178. See *Forensic Statistical Tool Source Code*, GITHUB, <https://github.com/propublica/nyc-dna-software> [<https://perma.cc/GVR2-ZZ4K>]

179. Kirchner, *supra* note 137.

180. *Id.*

181. See *id.*

182. *Id.*

183. Kirchner, *supra* note 104.

TrueAllele's nominal provision of the ability to use the most superficial levels of its software is a distinction without a difference.

The *Collins* order's language on the importance of external review and validation is another sign of potential trouble for TrueAllele. A continued lack of independent expert review — and not just peer-reviewed articles — might lead a court to conclude that TrueAllele was not generally accepted by the scientific community. Although Michael Robinson was unsuccessful at convincing the court that the PCAST report cast sufficient doubt on TrueAllele to render it novel in the court's eye, another court without the *Foley* precedent might well be persuaded by the combination of the PCAST report and other criticisms of TrueAllele's lack of external validation studies.

But what if the judge in *Foley* was right and these programs cannot be commercially sustainable if they reveal their code? This is a rather hypothetical point, given that TrueAllele's primary competitor STRmix has begun to provide its code to defense teams, but one which a judge might still conceivably find persuasive. There are currently free open-source probabilistic genotyping software programs such as LRMix Studio that are available for use.¹⁸⁴ Therefore it is unlikely that this technology will be forced to disappear from courtrooms entirely. However, due to evidentiary rules and other challenges, it may be difficult to move from widespread acceptance of TrueAllele to conditioning inclusion on disclosure. A more immediately implementable solution might be to create judicially-enforced conditions for disclosure of source code during trials, such as a requirement that defense attorneys agree not to disclose or use the code for any purpose other than the immediate trial. This type of solution permits the rigorous interrogation of the algorithms and code by the defense team without implicating the types of competitive concerns Cybergenetics has thus far used to resist efforts to disclose its code.

The challenges posed by complex-mixture DNA samples mean that the criminal justice system needs to move to more objective analysis of this type of forensic evidence. Probabilistic genotyping algorithms may eventually provide objective, valid, reliable results for a variety of types of DNA evidence. But as of now, the lack of independently verified scientific validity evidence and the lack of transparency surrounding the subjective decisions embedded within the interpretive programs' codes undermine the use of algorithms to analyze complex DNA samples.

To address these issues, courts should rigorously examine whether a given algorithmic system has been validated for a particular type of evidence analysis and refuse to admit evidence that lacks demonstrated validity for a given mixture type. Courts should also consider adopting a rule barring the results of algorithmic analysis of complex mixtures un-

184. See, e.g., LRMIX STUDIO, <http://lrmixstudio.org/> [<https://perma.cc/B7WQ-X22U>].

less the source code and built-in assumptions behind the algorithmic process are disclosed to defense teams. Doing this will mitigate Confrontation Clause and Due Process concerns and preserve defendants' constitutional rights. Companies' trade secrets and property rights can additionally be protected by the adoption of rules prohibiting secondary disclosure or improper use of disclosed code. Adopting this rule will remove the "black box" aspect of black box probabilistic genotyping algorithms and encourage a more just criminal justice system.