

SUPERIOR COURT OF NEW JERSEY  
APPELLATE DIVISION

DOCKET NO. A-4207-19T2

STATE OF NEW JERSEY,

Plaintiff-Respondent,

vs.

COREY PICKETT,

Defendant-Movant.

Criminal Action

On Leave Granted to  
Appeal an Interlocutory  
Order of the Superior  
Court of New Jersey, Law  
Division, Hudson County.

Ind. No.: 17-07-470-I

**Sat below:** Hon. Patrick J. Arre, J.S.C.

---

**BRIEF OF *AMICUS CURIAE* THE LEGAL AID SOCIETY IN SUPPORT OF  
DEFENDANT-MOVANT**

---

J. David Pollock (No. 022452012)  
The Legal Aid Society  
199 Water St., 5th Floor  
New York, NY  
(212) 577-3300

TABLE OF CONTENTS

	<u>Page</u>
TABLE OF AUTHORITIES .....	i
IDENTITY AND INTEREST OF <i>AMICUS CURIAE</i> .....	1
PRELIMINARY STATEMENT .....	1
STATEMENT OF FACTS, PROCEDURAL HISTORY AND TABLE OF JUDGMENTS, ORDERS, AND RULINGS .....	3
ARGUMENT .....	3
I. LEGAL STANDARD .....	3
II. ACCESS TO SOURCE CODE IS REQUIRED TO ASSESS THE RELIABILITY OF PROBABILISTIC GENOTYPING SOFTWARE .....	3
A. Probabilistic Genotyping Software is Complex and Inherently Unverifiable .....	3
B. The Now-Discredited Forensic Statistical Tool (FST) Should Serve as a Cautionary Tale for Software Developed Within a Black Box and Shielded from Adversarial Scrutiny. ....	7
C. This Court Should Not Rely on Flawed Precedents From Other Jurisdictions in Deciding Whether to Grant Reasonable Access to TrueAllele’s Source Code. ....	10
CONCLUSION .....	13

TABLE OF AUTHORITIES

	Page(s)
<b>Cases</b>	
<i>Frye v. U.S.</i> , 293 F. 1013 (D.C. Cir. 1923).....	3
<i>People v. Williams</i> , 35 NY3d 24 (2020).....	12
PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, <i>Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods</i> 79 (Sept. 2016), <a href="https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf">https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf</a> .....	4, 9
<i>State v. Doriguzzi</i> , 334 N.J. Super. 530 (App. Div. 2000).....	3, 11
<i>State v. Harvey</i> , 151 N.J. 117 (1997).....	3
<i>State v. Henderson</i> , 208 N.J. 208 (2011).....	3
<i>State v. Moore</i> , 188 N.J. 182 (2006).....	3
<b>Other Authorities</b>	
Lauren Kirchner, <i>Traces of Crime: How New York’s DNA Techniques Became Tainted</i> , N.Y. TIMES (Sep. 4, 2017), <a href="https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html">https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html</a> .....	9, 10

**INDEX TO APPENDIX**

FORENSIC DNA EVIDENCE INTERPRETATION 13-35 (John S. Buckleton, Jo-Anne Bright & Duncan Taylor eds., CRC Press 2016)	Aa1-25
Affidavit of Eugene Lien (July 17, 2017), from <i>New York v. Johnnie Jackson</i> , Ind. No. 727/201 (N.Y. Sup. Ct.)	Aa26-28
Denial of Motion to Quash, ECF No. 67 (July 6, 2016), from <i>United States v. Kevin Johnson</i> , 15-Cr-00565-VEC (S.D.N.Y.)	Aa29-30
Letter from The Legal Aid Society and Federal Defenders of New York to the New York State Inspector General in Reply to the OCME (December 7, 2017)	Aa31-37
Jeanna Matthews et al., <i>The Right To Confront Your Accusers: Opening the Black Box of Forensic DNA Software</i> , PROCEEDINGS OF THE 2019 AAAI/ACM CONFERENCE ON A.I., ETHICS, AND SOCIETY 321-327 (2019)	Aa38-44
Julia Gasston, et al., <i>An examination of aspects of the probabilistic genotyping tool: Forensic Statistical Tool</i> , 1362 WILEY INTERDISCIPLINARY REVIEWS: FORENSIC SCI. 1-6 (2019)	Aa45-50
Declaration of Nathaniel Adams, ECF No. 153-1 (Jan. 9, 2017), from <i>United States v. Kevin Johnson</i> , 15-Cr-00565-VEC (S.D.N.Y.)	Aa51-101

**IDENTITY AND INTEREST OF AMICUS CURIAE**

The Legal Aid Society (LAS) is the nation's largest and oldest provider of legal services to low-income families and individuals and is the primary public defender in New York City. In 2013, LAS created a DNA Unit to assist the nearly 700 LAS staff attorneys representing clients citywide in defending clients in cases involving DNA evidence.

Increasingly, LAS clients are prosecuted based upon evidence generated by complex DNA software programs. From 2012 to 2017, LAS litigated the admissibility of a proprietary probabilistic genotyping program, the Forensic Statistical Tool (FST), including seeking access to its source code. Through the efforts of attorneys at LAS and other defender organizations, the Office of Chief Medical Examiner (OCME) was ordered to provide the source code to a defense expert for examination in 2016. Significant flaws were discovered in the program, and OCME ceased using FST in 2017. Because of both the breadth of its representation and its experience advocating around this issue, LAS is well-positioned to address why defense access to source code is critically important to the *Frye* inquiry.

**PRELIMINARY STATEMENT**

The rapidly expanding field of probabilistic genotyping is extraordinarily complex. On that, all parties can agree. Probabilistic genotyping software programs exist at the

intersection of DNA genotyping, complex statistical modeling, and superhuman processing power. Purveyors of such programs – including TrueAllele, which is at issue in this case – routinely market their ability to wring results from tiny amounts of DNA and provide “reliable answers on previously unsolvable DNA evidence.” *TrueAllele® Technology: Computer interpretation of DNA evidence* (last accessed Oct. 15, 2020), [https://www.cybgen.com/solutions/brochures/lab\\_brochure.pdf](https://www.cybgen.com/solutions/brochures/lab_brochure.pdf).

However, from great power comes great responsibility. In few other legal arenas do the consequences of a mistake weigh more heavily, as the average juror, lawyer, and judge plainly do not possess the specialized knowledge that would be required to independently assess a report from a program like TrueAllele. For that reason alone, significant and detailed expert review must take place prior to any admission in court. Such review, as discussed in this brief, requires disclosure of the program’s underlying computer source code.

Moreover, *Amicus* submits, the argument that the State repeatedly falls back on in this appeal – that TrueAllele should be deemed reliable because numerous other states have concluded it is reliable, and *none* required an examination of its source code – is one that this Court should approach with a healthy skepticism.

**STATEMENT OF FACTS, PROCEDURAL HISTORY AND TABLE OF JUDGMENTS,  
ORDERS, AND RULINGS**

*Amicus curiae* adopt and incorporate by reference the statement of facts, procedural history and table of judgments, orders, and rulings as set forth in the Defendant-Movant's Brief.

**ARGUMENT**

**I. LEGAL STANDARD**

New Jersey courts employ the *Frye* test to determine the admission of expert testimony in criminal cases. *State v. Harvey*, 151 N.J. 117, 169-170 (1997). The *Frye* test requires courts to evaluate whether the science underlying expert testimony is generally accepted in the relevant scientific community. *Frye v. U.S.*, 293 F. 1013, 1014 (D.C. Cir. 1923); see also *State v. Doriguzzi*, 334 N.J. Super. 530, 539 (App. Div. 2000). Scientific theories are accepted as reliable when they are "based on a sound, adequately-founded scientific methodology involving data and information of the type reasonably relied on by experts in the scientific field." *State v. Henderson*, 208 N.J. 208, 247-48 (2011) (citing *State v. Moore*, 188 N.J. 182, 206 (2006)).

**II. ACCESS TO SOURCE CODE IS REQUIRED TO ASSESS THE  
RELIABILITY OF PROBABILISTIC GENOTYPING SOFTWARE.**

**A. Probabilistic Genotyping Software is Complex and  
Inherently Unverifiable.**

The differences between gold-standard forensic DNA testing and probabilistic genotyping software, of which TrueAllele is but

one variant, are well-developed in Defendant-Movant's brief. Db12-17. However, it bears emphasis that the interpretation stage of DNA testing is not always straightforward, particularly when low levels of DNA or complex mixtures are involved. "Stochastic," or random, effects and artifacts present major interpretational challenges for scientists conducting forensic DNA testing because they make it much more difficult to figure out the genotypes, or DNA profiles, of the individuals in the sample: Alleles not belonging to the true donors appear; alleles belonging to the true contributors disappear or are distorted; and artifacts masquerade as real alleles. Aa1-25 (FORENSIC DNA EVIDENCE INTERPRETATION 13-35 (John S. Buckleton, Jo-Anne Bright & Duncan Taylor eds., CRC Press 2016)).

Probabilistic genotyping software was developed to attempt to address these interpretational challenges. As the President's Council of Advisers on Science and Technology (PCAST) concluded, there can be little doubt that the "objective" approach of probabilistic genotyping represents an improvement on a purely subjective approach. See PRESIDENT'S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* 79 (Sept. 2016), [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast\\_forensic\\_science\\_report\\_final.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf) ("PCAST Report"). However, that does not mean that probabilistic



genotyping software can solve every DNA mixture, or even that it is appropriate for every DNA case.

Thus, while acknowledging its promise, PCAST emphasized that probabilistic genotyping is in its infancy and must be subject to "careful scrutiny." *Id.* In particular, PCAST found, probabilistic genotyping programs should be evaluated to determine "(1) whether the methods are scientifically valid . . . and (2) whether the software correctly implements the methods." *Id.* Critically, this evaluation process should consist of studies conducted by multiple different testing entities, including entities "not associated with the software developers." *Id.*

Most technical procedures or scientific tests in a lab can be verified because there exists a true amount or concentration to be measured. However, probabilistic genotyping programs do not have a measurable 'ground truth' and thus cannot be objectively validated. A practical example of two 'ground truths' in the metric system of measurement are water's freezing point at zero degrees Celsius and its boiling point at 100 degrees. The accuracy of any thermometer can be tested against water's boiling and freezing points. In contrast, there is no objectively "correct" likelihood ratio, so different programs looking at the same evidence will often arrive at different results and lead to strikingly different conclusions.

Moreover, probabilistic genotyping software programs are tested and validated on pristine, lab-controlled samples and then applied to inherently messier and more complex real-world samples. Samples created in a laboratory, in other words, simply lack the complexity of mixed DNA left at a crime scene. There is no way for a criminalist to know, for instance, how many people contributed to a DNA mixture recovered from an alleged crime scene or the degree to which the DNA in such a sample has degraded because of conditions in the environment. As a result, it is difficult, if not impossible, for the creators of probabilistic genotyping software to understand and account for that uncertainty. To be sure, probabilistic genotyping programs employ biological and mathematical models to account for stochastic effects and assign statistical weight to the likelihood of such effects in a particular sample. However, the inner workings of those models, and the data and assumptions built into the program, add layer upon layer of complexity to the software and are, accordingly, centrally important in evaluating the program's reliability.

In the end, given the complexity of the methods and their inherently unverifiable results, the only way to determine whether a program gets it wrong in a given case is to look directly at how the program performs its calculations. And to be sure those

calculations are performed correctly requires access to the program's source code.

**B. The Now-Discredited Forensic Statistical Tool (FST) Should Serve as a Cautionary Tale for Software Developed Within a Black Box and Shielded from Adversarial Scrutiny.**

The Forensic Statistical Tool (FST) provides an apt and troubling case study in the need for access to source code. A probabilistic genotyping program developed and deployed exclusively by the New York City Office of Chief Medical Examiner (OCME), FST was used in thousands of criminal prosecutions while its creators fought tooth and nail against the demands of *Amicus Curiae* and others for access to its source code.

From the day in 2011 when FST was first utilized in real casework until 2016, when a federal district court denied OCME's motion to quash a subpoena for the source code, FST operated as a "black box." Indeed, during that time, *OCME alone* was privy to the calculations it performed or the source code performing those calculations.

OCME's sole access to the source code meant that the laboratory alone also had exclusive knowledge of an explosive error: the program, as written, didn't actually work. In 2011, shortly after the laboratory transitioned from using FST on only lab-created samples to using it on real-life ones, OCME analysts realized that FST simply did not work as promised. Specifically,

when FST was applied to real casework samples, it would, at times, produce a *negative* likelihood ratio. Aa26 (Affidavit of Eugene Lien ¶ 6 (July 17, 2017), from *New York v. Johnnie Jackson*, Ind. No. 727/201 (N.Y. Sup. Ct.) ("Lien Aff.")). This was unquestionably an error: A negative likelihood ratio is an incoherent result because likelihoods can never be negative, only positive.

The OCME failed to disclose this error to anyone, and likely would never have done so had a federal district court judge not forced the OCME's hand by ordering the release of the FST source code in 2016. Aa29-30 (Denial of Motion to Quash, ECF No. 67 (July 6, 2016), from *United States v. Kevin Johnson*, 15-Cr-00565-VEC (S.D.N.Y.)). To this day, the OCME has never provided a satisfying explanation as to *why* the negative likelihood ratios occurred. See Aa31-37 (Letter from The Legal Aid Society and Federal Defenders of New York to the New York State Inspector General in Reply to the OCME (December 7, 2017)).

Whatever the cause of the flaw, after it was discovered, FST temporarily was taken offline. Aa26 (Lien Aff. ¶ 6). OCME proceeded in secret to alter the FST software in 2011, ostensibly to correct for the production of negative likelihood ratios. *Id.* ¶¶ 6-8. But rather than remedy the root cause of the error, in its haste to re-launch FST the OCME instead added a crudely designed function named "CheckFrequencyForRemoval" or "CheckFreq." Aa38-44 (Jeanna Matthews et al., *The Right To Confront Your Accusers: Opening the*

*Black Box of Forensic DNA Software*, PROCEEDINGS OF THE 2019 AAAI/ACM CONFERENCE ON A.I., ETHICS, AND SOCIETY 321-327 (2019)). Rather than addressing the underlying source code error, the OCME implemented CheckFreq to simply discard certain DNA information from being included in FST's calculation if there was a possibility that it could result in a negative likelihood ratio. *Id.* In studies of the impact of this function on FST's results, Gasston et al. conclude that the data-dropping behavior could, in some circumstances, skew the results of the likelihood ratio in favor of the party arguing that DNA from the "person of interest" is present (often the prosecutor). Aa45-50 (Julia Gasston, et al., *An examination of aspects of the probabilistic genotyping tool: Forensic Statistical Tool*, 1362 WILEY INTERDISCIPLINARY REVIEWS: FORENSIC SCI. 1-6 (2019)).

Of course, nobody outside the OCME was aware of these events as they happened because FST's source code remained a closely-held secret until 2016; all the while, FST was used to analyze evidence in at least 1,350 criminal cases over that span. Lauren Kirchner, *Traces of Crime: How New York's DNA Techniques Became Tainted*, N.Y. TIMES (Sep. 4, 2017), <https://www.nytimes.com/2017/09/04/nyregion/dna-analysis-evidence-new-york-disputed-techniques.html>.

When the FST source code was finally subject to close examination, the defense expert in *Johnson* concluded that he "did not leave with the impression that FST was developed by an

experienced software development team,” and that, pending more rigorous testing, “the correctness of the behavior of FST software should be seriously questioned.” Aa72 (Declaration of Nathaniel Adams, ECF No. 153-1, at 21 (Jan. 9, 2017), from *United States v. Kevin Johnson*, 15-Cr-00565-VEC (S.D.N.Y.)). Not coincidentally, soon after the disclosure of FST’s source code the OCME announced that it would begin phasing out FST as of January 1, 2017. See *Traces of Crimes, supra*.

**C. This Court Should Not Rely on Flawed Precedents From Other Jurisdictions in Deciding Whether to Grant Reasonable Access to TrueAllele’s Source Code.**

One of the State’s central arguments for the proposition that the reliability of TrueAllele is clear without resort to its source code – the litany of out-of-jurisdiction decisions that have concluded that TrueAllele is reliable *without permitting its source code to be inspected* – is circular reasoning at its finest.<sup>1</sup> This Court should see through that smokescreen.

Relying on the purported strong foundation created by “numerous nationwide admissibility rulings” that admitted

---

<sup>1</sup> The State’s other arguments are likewise flawed, but in light of the numerous other *Amici* vying for the Court’s attention (with which The Legal Aid Society joins), they are beyond the scope of this brief. Of particular relevance here, however, the State fails to respond to – much less counter – Defendant-Movant’s cogent argument that protective orders regularly shield trade secrets from being illegally stolen in civil patent litigation, and are quite plainly able to play the same role here.

TrueAllele without source code disclosure, the State trumpets that “[s]ource code was not needed in any of these cases.” Rb17-18. On closer inspection, however, those cases have instead created a house of cards, one on top of another with little to keep everything standing. To cite an example, one early case that is discussed in the State’s brief, *Commonwealth of Pennsylvania v. Kevin Foley* (Ra15-33), is relied upon in at least ten subsequent cases that appear in the State’s Appendix. See Ra7, Ra37, Ra46, Ra74, Ra82, Ra144, Ra170, Ra186, Ra193, Ra213. Indeed, many of the cases in the State’s Appendix include a variation on the same laundry list of admissibility rulings, but few (if any) consider the source code issue independent of this seemingly insurmountable wall of authority.

While the admissibility rulings of courts in other states are no doubt compelling, the precedential value of such cases ultimately lies not in their number or the fact that each permitted the admission of evidence derived from TrueAllele, but rather from the power of the reasoning and the opinions of unbiased scientists underlying it. See *State v. Doriguzzi*, 334 N.J. Super. 530, 546 (App. Div. 2000) (“General acceptance within the relevant scientific community consists of more than just counting up how many cases go in a certain direction.”).

A recent decision of the New York Court of Appeals, *People v. Cadman Williams*, exemplifies this conundrum. Writing with the

benefit of hindsight regarding the checkered history of FST, *supra*, and a related methodology, LCN DNA testing, the New York Court of Appeals held that the lower court erred in failing to grant a *Frye* hearing as to either. *People v. Williams*, 35 NY3d 24 (2020). The Court's identified two major errors in the lower court's reasoning. First, as to LCN, "the motion court relied heavily on [a] prior judicial opinion" that "did not adequately assess whether OCME's LCN testing was generally accepted within the relevant scientific community." *Id.* at 39. "The repetition of a single, questionable judicial determination does not strengthen or add validity to such ruling, and it defies logic that an error, because it is oft-repeated, somehow is made right." *Id.*

Second, the Court found fault with the lower court's reliance on the State's representation that FST was built on established mathematical formulas that had achieved acceptance among scientists. The lower court "did not adequately account for either the proprietary nature of the FST or the relatively narrow subsection of the relevant scientific community able to examine and endorse that tool." *Id.* at 42. Because the FST was a "proprietary program exclusively developed and controlled by" the OCME, which also acted as its "sole developer and the sole user," the Court concluded that, for the purposes of *Frye*, the lower courts were bound to ensure that it was "supported by those with no professional interest in its acceptance." *Id.* at 41-42.



While the issue here is not whether to hold a *Frye* hearing, but whether the lower court was correct to constrain its scope, the holding in *Williams* is no less resonant. There is no doubt that TrueAllele and its source code are the exclusive domain of Cybergenetics, and the litany of admissibility decisions upon which the State relies must be seen in that light. Like *Williams*, this Court should not rely on a jurisprudential echo chamber alone. Instead, it should follow the lessons of FST and require source code disclosure in order to allow the parties to make reasoned, informed, and complete arguments.

#### CONCLUSION

For the foregoing reasons, *Amicus Curiae* The Legal Aid Society respectfully submit that this Court should heed the lessons of FST, reverse the lower court's order and provide the Defendant-Movant access to TrueAllele's source code under an appropriate protective order.

Respectfully Submitted,

    /s/ J. David Pollock  
J. David Pollock (No. 022452012)  
The Legal Aid Society  
199 Water St., 5th Floor  
New York, NY  
Dated: October 15, 2020