SUPERIOR COURT OF THE STATE OF NEW YORK

COUNTY OF NIAGARA: CRIMINAL TERM

THE PEOPLE OF THE STATE OF NEW YORK

Indictment 2015-041

VS.

DAVID SMITH

NOTICE OF MOTION

Defendant

SIRS/MADAMES:

PLEASE TAKE NOTICE, that upon the annexed affirmation of **Dominic Saraceno**, ESQ., Counsel for the defendant, and upon all the proceedings heretofore had herein, the undersigned will move this Court at a date to be determined by the Court, or as soon thereafter as Counsel can be heard for an Order granting:

1. Requesting a Frye Hearing.

2. Reservation of Rights to Make Further Motions

And for such other and further relief as to this Court may deem just and proper.

SUPERIOR COURT OF THE STATE OF NEW YORK

COUNTY OF NIAGARA: CRIMINAL TERM

_____

THE PEOPLE OF THE STATE OF NEW YORK

VS. **AFFIRMATION**

DAVID SMITH

Defendant

_____

STATE OF NEW YORK) SS:

COUNTY OF NIAGARA)

I, **Dominic Saraceno**, ESQ., hereby submit and affirm the following to be true under the penalties of perjury that:

1. I am an attorney admitted to practice in the State of New York and represent the defendant.

2. I am familiar with all prior proceedings pertaining to the instant indictment.

3. This affirmation is made in support of an application for various forms of relief set forth in the annexed Notice of Motion.

4. This affirmation is made upon information and belief, the sources thereof being official Court papers, conversations with the defendant and Assistant District Attorney, and the files maintained in my office.

I. MOTION REQUESTING FRYE HEARING

Background

The people, after obtaining a buccal swab from the defendant, compared the defendant's DNA to DNA found on some items believed to be used in the commission of a crime. The results were initially inconclusive. The people indicated that the crime lab would be re-testing the items using different thresholds. A day before jury selection, the DNA testing documents were turned over to defense counsel and it became apparent that the DNA was not tested using different thresholds, but rather an entirely different DNA test was performed using a new forensic software called STRmix.

STRmix has never been accepted in a New York Court, so it is by its nature novel scientific evidence. To be admissible in New York Courts, it must first pass the Frye test as formulated in Frye v. United States, 293 F. 1013 (1923) and subsequently adopted by the New York Court of Appeals in People v. Middleton, 54 N.Y.2d 42.

That protocol requires that expert testimony be based on a scientific principle or proceeding which has been "sufficiently established to have gained general acceptance in the particular field in which it belongs" (Frye, at 1014).

The People assert that STRmix operates in essentially the same way that TrueAllele does, and because the court in Wakefield found TrueAllele to pass the Frye test, the Court here should find STRmix to be admissible without the need for a Frye test.

STRmix is not TrueAllele

TrueAllele is a generally accepted technology for automated and accurate interpretation of DNA mixture evidence. Starting in 1999, TrueAllele was carefully tested and improved over a 10-year development period before being used in criminal cases. TrueAllele results have been reported in a dozen criminal cases in New York, with expert testimony given in six trials (Chemung, Monroe, Onondaga and Schenectady counties) and a successful Frye hearing in People v. Wakefield.

STRmix is recent foreign copycat software that purports to share some of TrueAllele's DNA analysis capabilities. The Erie County crime laboratory used STRmix to analyze DNA mixture short tandem repeat (STR) data in this case. The People contend that STRmix is just like TrueAllele, and so a Frye hearing is not required. However, that position is incorrect.

There are similarities between the two software programs. Both use quantitative peak data derived from amplification of a DNA mixture. Both consider genotype combinations

based on mixture weights, and derive a data variance. Both determine probabilities of data to calculate a likelihood ratio (LR) match statistic to help identify a suspect.

However, there are many differences between these two "probabilistic genotyping" programs. Some differences involve the applicability of the software, and how well they work in practice on actual evidence. Below is a description of some of these differences, and their impact on the reliability and relevance of the LR results.

Difference Number 1 - Subjectivity

When TrueAllele solves a problem and infers genotypes, it does not know the defendant's genotype. An objectively determined TrueAllele evidence genotype can be compared later on with one, ten or a thousand genotypes.

STRmix, however, requires the defendant's genotype as part of its operation. The software knows the answer that the prosecutor wants. And without knowing that desired answer, the New Zealand software cannot give the prosecutor any answer at all.

Difference Number 2 – Modeling/Assumptions

TrueAllele uses all the quantitative STR data, applies its model of how such data is formed from mixed genotypes, and reaches objective unbiased conclusions. STRmix, unlike TrueAllele, needs to make unwarranted assumptions about data when it performs its more limited analysis. Some of these assumptions include:

Thresholds. STRmix imposes a "threshold" that discards peak data below a certain level. This unscientific data removal affects its conclusions.

Drop out. STRmix can mathematically conjure up chances for data that isn't there, using that imagined data to misidentify a defendant. Its "drop out" approach uses probability to insert absent evidence that the prosecution wishes it had.

Drop in. STRmix uses probability to justify otherwise inexplicable data peaks. So when potentially exclusionary data are present, the software massages it away.

Fixed variance. STRmix uses a preset number to explain the variability of STR data. Such a fixed value conflicts with statistical modeling practice, which standardly infers variance from data. The limitation can lead to inaccurate probabilities, undermining the whole point of probabilistic genotyping.

Difference Number 3 - Calibration

Modern statistical computing infers information about parameters and their variation directly from provided data. This is how TrueAllele operates.

STRmix, however, requires a "calibration" based on unrelated laboratory data. For example, PCR "stutter" is a PCR amplification artifact that introduces small shadow peaks. Unlike TrueAllele, STRmix must first determine the statistical parameters of this artifact, even though those parameters may not accurately describe the actual stutter data present in a later case.

Difference Number 4 - Validation

Validation studies conducted on actual data assess the reliability of a method. The usual axes are ***sensitivity*** (ability to include someone who has contributed their DNA to evidence), ***specificity*** (ability to exclude someone who has not contributed DNA to evidence), and ***reproducibility*** (getting the same results in multiple method applications). TrueAllele has undergone extensive testing on both laboratory synthesized and actual casework data, including 30 validation studies of which 7 are peer-reviewed.

STRmix, however, is not as extensively tested as TrueAllele. This paucity of studies is presumably because STRmix fails to give reasonable answers in many situations (e.g., too many contributors, very low DNA amounts, highly unbalanced genotype combinations, mixtures of related individuals).

Most worrisome is STRmix's high false inclusion rate. By STRmix's own estimate, false positives that erroneously include the wrong person occur in .01%-.05% of STRmix analyses. Falsely including an innocent man is contrary to a "reliable" or "generally accepted" forensic method (see attached exhibit).

Difference Number 5 - Operation

The TrueAllele process is a fully automated procedure. A human operator enters all the original quantitative data into the system. TrueAllele then automatically solves the problem, separating the data into the genotypes of each contributor to a mixture. Afterwards, TrueAllele compares evidence and reference genotypes to calculate match statistics.

STRmix, however, requires a human operator to spend considerable time manually preprocessing the data before entering it into the program. A person must visually examine every peak data event, determining what they believe to be "real" data versus artifacts. This subjective determination is unusual in modern statistical computing, and introduces opportunities for bias and error.

Difference Number 6 - Formulations

TrueAllele splits DNA mixture into genotypes. Afterwards, and only afterwards, it makes a comparison with a defendant. TrueAllele addresses one hypothesis – did a person contribute their DNA to the evidence.

STRmix, however, lumps genotypes together in its analysis. It requires a forensic laboratory analyst to enter a prosecutor's hypothesis about who contributed their DNA to a mixture. Moreover, STRmix requires that the analyst also enter into the program a defense hypothesis of an alternative explanation. Many different hypotheses can be conjectured and compared, each producing a different LR match statistic.

Difference Number 7 - Comprehensibility

TrueAllele splits mixture data into genotypes of individual contributors. Therefore its explanations are similar to those of standard approaches from 20 years ago for unmixed DNA from a single person. The match statement is "a match between this evidence and defendant is (some number) times greater than coincidence."

STRmix, however, lumps all the data together. The software can only consider complex explanations involving combinations of different contributors. The STRmix likelihood ratio statement reflects this complexity, and may be incomprehensible to the jury.

For example, under one (of many) set of assumed prosecutor and defense hypotheses, a STRmix user's LR statement could read: "the probability of observing this DNA evidence under the prosecution's hypothesis that the defendant, victim and an unknown person all contributed their DNA to the mixture, divided by the probability of observing the evidence under the defense hypothesis that three unknown people contributed their DNA, is (some number)."

What is a jury member supposed to make of "the probability of observing the data"? Will they really understand the different convoluted hypotheses, and the resulting conditional probabilities? Are juries actually comprised of statisticians who can understand these highly technical likelihood ratio statements?

Difference Number 8 - Errors

Unlike TrueAllele, errors in STRmix software have occurred. These STRmix software errors have necessitated post-conviction review of criminal cases in Australia that used the program (see attached exhibit).

Difference Number 9 - Acceptance

TrueAllele has successfully overcome Frye and Daubert challenges in seven states (California, Louisiana, New York, Ohio, Pennsylvania, South Carolina and Virginia), as well as in Australia and the United Kingdom. STRmix has yet to overcome a single admissibility challenge in the United States, and has not been accepted by any New York court. It is unclear whether the STRmix technology is sufficiently reliable to withstand such a challenge, which is why a Frye hearing is needed in this case.

Summary (Post Wakefield)

Supreme Court Justice Michael Coccoma accepted TrueAllele as reliable this year after a Frye challenge in Schenectady, NY in the case of People v. John Wakefield. The judge granted a Frye hearing "since Cybergenetics TrueAllele Casework has never been accepted in a New York Court, it is by nature novel scientific evidence." Applying the Judge's legal standard in Wakefield, STRmix must also be subjected to a Frye hearing since STRmix has never been accepted in a New York court. The People cannot conclude that it is similar to TrueAllele and therefore a Frye hearing is not needed. The People also cannot make a conclusory statement that STRmix is generally accepted in the scientific community, as that conclusion can only be reached after a hearing.

Judge Coccoma noted that TrueAllele offered three principal advantages: (1) *productivity* that eliminates human involvement, (2) *information* that dispenses with simplifying assumptions, and (3) *objectivity* that removes the suspect genotype from consideration. STRmix enjoys none of these TrueAllele advantages, since it involves people in the analysis process, requires simplifying assumptions, and considers the suspect in its deliberations. From the trial judge's perspective, STRmix is clearly not the same as TrueAllele.

The Wakefield ruling noted that validation studies found TrueAllele to be highly specific. This specificity is important for excluding innocent defendants. But STRmix has a relatively high false positive rate, which has not been as extensively studied and is critical for assessing the New Zealand's software potential abuse in criminal cases.

The ruling noted many validation studies that demonstrated that TrueAllele "thoroughly examined data, eliminated examiner bias, accurately preserved identification information, quantified match strength (whether positive or negative), and yielded reproducible results prior to its use thereof. These studies proved that Cybergenetics TrueAllele Casework is reliable, and that is part of the standard for admissibility." STRmix has not been shown to meet these reliability criteria, and so a Frye hearing is needed to determine whether the software is actually reliable as well as whether it has been accepted and endorsed in the scientific community.

The STRmix developers have written about what their user laboratories need to do in order to overcome an admissibility challenge. Amongst other factors, they "consider essential" (a) choosing suitable hypotheses for a case, knowledge of the (b) limits and uncertainties of an LR produced with STRmix, and (c) diagnosing poor performance.

The proper forum for arguing legal admissibility is not a scientific paper, but rather a court of law. The State would like to introduce DNA evidence with match statistics calculated using STRmix software, which is entirely distinct from TrueAllele. STRmix has not been accepted as reliable or generally accepted in the scientific community. Therefore the prosecution has the burden of proof to demonstrate STRmix reliability and acceptance, and a Frye hearing is required.

II. RESERVATION OF RIGHTS TO MAKE FURTHER MOTIONS

Subject to the resolution of the aforesaid motions, the defendant reserves the right to make any other motions or renew any motions already made upon the discovery of any new evidence or information.

The defendant further reserves the right to request an adjournment after the holding of any pre-trial hearings to investigate information obtained at said hearings pursuant to People v. Peacock, 31 NY2d 907 (1972).

WHEREFORE, your affiant respectfully requests this Court to grant the relief sought herein and for such other and further relief as to this Court may deem just and proper.

Dated:

Lockport, New York

_____
Dominic Saraceno, ESQ.