IN THE SUPERIOR COURT OF WASHINGTON
IN AND FOR THE COUNTY OF KING

STATE OF WASHINGTON, **Plaintiff,** **v.** **EMANUEL FAIR,** **Defendant.**	**No. 10-1-09274-5 SEA** **MOTION TO COMPEL CYBERGENETICS' TRUEALLELE CASEWORK SOURCE CODE**

MOTION

The defendant, Emanuel Fair, moves this court to compel the State and their expert, Dr. Mark Perlin to disclose to defense experts the source code of Cybergenetics' TrueAllele casework software. Dr. Perlin created TrueAllele Casework which is a unique artificial intelligence software system that claims deconvolute complex DNA mixtures. The State has retained Dr. Perlin and has submitted several evidence samples to TrueAllele for analysis. This motion is based upon Mr. Fair's United States and Washington State constitutional rights to due process and confrontation, CrR 4.7, *Brady v. Maryland*, 373 U.S. 83 (1963), *Kyles v. Whitley*, 514 U.S. 419 (1995),and on material provided below.

DATED THIS 7th DAY OF March, 2016

Respectfully submitted,

s/ Benjamin Goldsmith
WSBA #35017
Attorney for Defendant
King County Department of Public Defense
The Defender Association Division
810 3rd Ave. Ste 800
Seattle, WA 98104
Telephone: (206) 477-8700 x78734
Fax: (206) 447-2349
Email: Benjamin.Goldsmith@kincounty.gov

s/ Paul Vernon
WSBA #40641
Attorney for Defendant
King County Department of Public Defense
The Defender Association Division
810 3rd Ave. Ste 800
Seattle, WA 98104
Telephone: (206) 477-8700 x78792
Fax: (206) 447-2349
Email: paul.vernon@kingcounty.gov

DECLARATION OF COUNSEL

I, Benjamin Goldsmith, declare that the following is true and correct to the best of my knowledge:

1. Paul Vernon and I are the attorneys appointed to represent Mr. Fair in the above-entitled action.

2. Mr. Fair is charged by Information with one count of Murder in the First Degree with sexual motivation. As charged, Mr. Fair faces an indeterminate life sentence with a mandatory minimum sentence of 20 years in prison. The State indicated at the time of filing that they would seek a sentence of 45 years to life.

3. Much of State's case against Mr. Fair involves DNA evidence. Since 2008 when this crime occurred, the State has used the Washington State Patrol Crime Laboratory DNA Section, and three private labs, Bode Technologies, Sorenson Forensics, and most recently Cybergenetics to analyze DNA evidence in this case.

4. Cybergenetics uses propriety software called TrueAllele Casework. Dr. Mark Perlin, the CEO of Cybergenetics and creator of TrueAllele Casework, claims that the software deconvolutes DNA mixtures and performs probabilistic genotyping, which will be described in more depth below.

5. Probabilistic genotyping is a novel and evolving means of interpreting complex DNA mixtures.

6. TrueAllele Casework has been the subject of admissibility challenges in New York, California, Maryland, Ohio, and Virginia.

7. Cybergenetics is family owned business, the shares of which are entirely owned by Dr. Perlin, his wife, and a trust for their children.

8. Cybergenetics (i.e. Dr. Perlin) has declined in all cases to disclose the source-code which explains how his program functions.

9. As far as undersigned counsel is aware, this case will be the first time the State has attempted to use TrueAllele in court in this State. The defense has retained experts who are competent to assess whether TrueAllele Casework is a reliable method of DNA analysis.

10. One of the defense experts, Dr. Ranajit Chakraborty is the Director of the Center for Computational Genomics of the Institute of Applied Genetics and a Professor at the Department of Molecular and Medical Genetics of the University of North Texas Health Science Center in Fort Worth, Texas. His declaration and CV are attached at Appendix A. Dr. Chakraborty has served as a member of the U.S. National DNA Advisory Board (1995-2000) and is a frequent faculty member of the Scientific Working Group on DNA Analysis Methods (SWGDAM) since 1989. He has also testified as an expert for the prosecution in seminal DNA cases in this state including *State v. Copeland*, 130 Wash.2d 244 (1996), *State v. Gore*, 143 Wash.2d 288 (2001), and *State v. Jones*, 130 Wash.2d 302 (1996).

11. As part of his work, Dr. Chakraborty has personally written and supervised writing of several computer software source codes to conduct data analysis, some of which are routinely used in DNA forensics and relationship testing. *Id.* at 1.

12. Dr. Chakraborty is familiar with TrueAllele Casework both from his time serving at the New York DNA Subcommittee and in connection with work on several cases, including this one. *Id.* at 2.

13. Dr. Chakraborty has reviewed the validation studies published by Dr. Perlin. *Id.* However, the information that is published in the validation studies by Dr. Perlin is very generic and does not give details of several critical features of complex DNA mixtures such as the ones analyzed in this case. Without the software source codes of the system, it is impossible for Dr. Chakraborty to verify whether the underlying mathematical models of the system are accurately translated in the source code, or implemented accurately in computations. *Id.* at 3.

14. At least two law enforcement entities, The New South Wales Police Force, and the California Department of Justice have identified limitations in TrueAllele's functioning. *Id.* at 4-5.

15. Without the source code, Dr. Chakraborty, or any expert will be unable to verify the DNA interpretation that TrueAllele claims to conduct. *Id.* at 3.

16. Dr. Dan E. Krane is another expert the defense has retained to analyze TrueAllele. Dr. Krane is a Professor of Biological Sciences (with a courtesy appointment in the Department of Computer Science) at Wright State University in Dayton Ohio. Dr. Krane has published more than 50 scholarly papers and was a founding member and two-time gubernatorial appointee to the Commonwealth of Virginia's Scientific Advisory Committee, a 12-member panel established by statute to provide oversight and guidance to the Virginia Department of Forensic Science (the crime laboratory for the Commonwealth of Virginia). Dr. Krane's declaration and CV are attached at Appendix B.

17. Dr. Krane has read Dr. Perlin's publications regarding TrueAllele. *Id.* at 2. However, Dr. Perlin has not published any validation studies or third-party reviews of the

hundreds of variables or their submodels and their associated uncertainties, boundaries, and interrelationships that constitute the underlying probability model of TrueAllele. *Id.* at 3

18. It is Dr. Krane's expert opinion that the single best way to evaluate TrueAllele's probability model is through a review of its underlying source code. Source code is a code written in a high-level or assembly language, which is converted into object code by a compiler, assembler, or interpreter; a program in a source language. Source code is the precise yet human-readable description of the sequence, branches, and loops of computer instructions that constitute a computer program. While peer-reviewed articles are important parts of demonstrating scientific *concepts*, source code serves as the *implementation* of those concepts. *Id.* at 3.

19. Dr. Krane explains that computer software can be validated at several different levels. In the most superficial sense, software can be tested by evaluating the results it generates when provided with certain inputs. More substantive validation studies evaluate the input and output of individual components of a program'ssubcomponents or of modules that contain multiple components. "Black box" and "white box" testing roughly describe these two different kinds of software testing approaches. Black box testing requires little or no knowledge of the internal components of a computer program. White box testing requires knowledge of the internal components of a computer program (typically a combination of software design documents and source code). Some software engineers advocate black box testing because it evaluates the correctness of the output with respect to only the input and not the working of internal components. However,

software (like TrueAllele) that produces likelihood ratios (LRs) cannot be validated with only black box testing because the *correct* answer cannot be known (and therefore cannot be compared to the results generated by the program). *Id.* at 4-5.

20. Dr. Krane has found that others who have critically evaluated TrueAllele have noted significant concerns about the model it uses and the implementation of that model.

21. A review of the TrueAllele source code used in the analysis in this case would allow the defense experts:

a. A determination of what computations were performed by TrueAllele.

b. A determination of the scientific accuracy of the results of these computations by:

i. Evaluating whether the computations performed by TrueAllele and TrueAllele's resulting conclusions in the case of State of Washington vs. Fair are consistent with the published claims of Dr. Perlin.

ii. Evaluating whether these computations and conclusions are consistent with generally accepted principles that are routinely employed by human experts during the course of forensic DNA testing.

c. A determination of whether these methods were properly translated from concept to source code and that no mistakes were made during

the writing of TrueAllele's code. This would allow it to be determined if translation from concept to source code included any ambiguously described concepts that, in turn, led to unfounded assumptions in the source code.

d. A determination of whether alternative explanations of the observed data could have produced similar results to those produced by TrueAllele in this case. *Id.* at 9-10.

22. In DNA analysis, human experts are expected to explain how they arrive at a conclusion using alternative approaches when their preferred means of analysis fail in the context of a specific case's data. This same expectation can and should apply to a computer program such as TrueAllele. *Id.* at 12.

23. Dr. Krane, like Dr. Chakraborty, explains that the disclosures made by Dr. Perlin in his publications as well as in response to defense discovery demands do not sufficiently explain how TrueAllele arrives at its conclusions.

24. Like Dr. Chakraborty, Dr. Krane asserts that it is not possible to assess or confront TrueAllele's conclusions without a particularized understanding of the analysis it performs and that cannot be accomplished without its source code. *Id.* at 12.

25. On December 23rd, 2015 the defense requested disclosure of the source code from the King County Prosecuting Attorney's Office and Cybergenetics. Appendix C.

26. On January 27th, 2016, the State responded in part to the defense discovery demand and indicated "Cybergenetics will not provide its source code for TrueAllele". Appendix D.

27. The defense has consulted with an attorney experienced in patent litigation. Brian Ferguson is a partner at the firm of Weil, Gotshal, and Manges LLP and co-chair of the firm's Patent Litigation Practice Group. His declaration, its appendices, and his CV are attached at Appendix E.

28. Mr. Ferguson has dealt with the issue of production of source code in response to discovery requests on numerous occasions, in fact, in his practice area "this issue comes up very frequently". *Id.* at 2. In patent cases involving patents related to computer functionality "the production of source code by the defendant in response to discovery requests is almost always required." *Id.* at 5.

29. The defense experts have indicated that they would abide by a protective ordered signed by the Court. Without TrueAllele's source code, counsel will be unable to meaningfully move to preclude or challenge this evidence against Mr. Fair and will be unable to provide the effective assistance of counsel which he is entitled under the Sixth and Fourteenth Amendments of the United States Constitution.

30. The defense moves the Court to compel disclosure of TrueAllele's source code which can be protected by an appropriate protective order.

DATED THIS 7th DAY OF MARCH, 2016

Respectfully submitted,

Ben Goldsmith, WSBA #35017
Attorney for Defendant

BRIEF STATEMENT OF THE CASE

1. Mr. Fair is charged with an incredibly brutal murder. Mr. Fair is also innocent of this horrific crime. On October 31st, 2008 several residents of the Valley View Apartments in Redmond threw a Halloween party. One of those residents was Arpana Jinaga, the decedent. During the party, which lasted from the early evening until approximately 3 am on November 1, 2008, the apartments were open to the other residents and invitees. Emanuel Fair was one of the attendees at the party. Attendees of the party, including Mr. Fair, Ms. Jinaga and many others danced, drank, and posed for pictures inside Ms. Jinaga's apartment and throughout the apartment complex. During the party, Mr. Fair was in Ms. Jinaga's apartment for extended periods of time. The decedent was last seen alive shortly after 3 am when she left one of the other apartments on the ground floor and returned to her own third floor unit.

2. According to the estimates of King County Medical Examiner's office, Ms. Jinaga was beaten, strangled with a ligature, and possibly sexually assaulted between 3:30 am and 9 am on November 1. Her body was not discovered until two days later when a family friend went to her apartment to check on her at the request of her family in India.

3. The decedent was found lying naked face down on the floor of her bedroom covered with a green sheet. Crime scene investigators found that the coffee table in the living room had been wiped down with a bleach-like cleaning material and that bleach had been poured on the carpet of the living room, hallway, and bedroom. In the bathtub, the police found the decedent's comforter which was soaked in bleach and water. Her body was covered in motor oil and it appeared that the killer made some effort to light the body on fire. Her hands were soaked in

bleach and parts of her mattress had been wiped down with a similar bleach product as was used on the coffee table. Several valuable items belonging to Ms. Jinaga, including a laptop computer and television, were found in the apartment. Ms. Jinaga's blackberry phone and her digital camera were missing and have not been recovered.

4. The bed itself was largely stripped of bedding. The decedent's sheets, a bathrobe, and a Halloween costume she wore earlier in the evening were later found in the apartment complex's dumpster. Additionally, in the dumpster the police found a plastic bag containing a canister of motor oil. There are no direct witnesses to the killing. There are also no witnesses who claim that Mr. Fair had any inappropriate contact with the decedent prior to the killing.

5. The State investigated this case for almost two years before deciding to charge Mr. Fair. The police discovered evidence that continues to point to a number of suspects other than Mr. Fair. In order for the Court to appreciate the importance of the disclosure of TrueAllele's source code, and the reliability of the DNA evidence in this case, it is necessary to describe some of the other evidence in this factually complicated case.

6. Prior to charging Mr. Fair, much of the State's investigative focus fell upon Cameron Johnson, Ms. Jinaga's neighbor. Mr. Johnson was questioned by the police on four different occasions prior to Mr. Fair being charged. During an interview on November 5, 2008, two days after the body was found, Mr. Johnson claimed that after the party he fell asleep on the couch in his livingroom, which shares a common wall with Ms. Jinaga's apartment. Appendix F, 1374. Mr. Johnson claimed that he fell asleep and then woke up to the what "sounded like she [Arpana] was having sex." *Id.* at 1376. Mr. Johnson claimed that after he awoke this sound, he texted a prior girlfriend of his to say "yo". Mr. Johnson initially claimed not to know why he texted his

prior girlfriend but later admitted it was to engage in sexual activity with her. *Id.* at 1376, 1378. Mr. Johnson then told police he "just went back to sleep" and did not call or attempt to contact anyone else. *Id.* at 1380. Mr. Johnson also claimed that he attempted to call Ms. Jinaga at 10 or 11 am on Saturday, November 1st to "see how she was doing"[1]. *Id.* at 1388. The detectives interviewing Mr. Johnson asked to see his phone during the interview and quickly determined that he had attempted to call Ms. Jinaga at 2:56 am and again at 3:02 am. *Id.* at 1390. Mr. Johnson claimed not to remember his calls to Ms. Jinaga but admitted that he was possibly looking to engage in sexual activity with her because "She looked really good…I hadn't seen her for months, I'd never thought about her like that [prior to the night of the Halloween party]". *Id.* at 1392-93. While Mr. Johnson persisted in claiming he did not remember attempting to contact Ms. Jinaga at 2:56 am and 3:02 am, he did claim to remember texting his prior girlfriend at 2:59 am, three minutes after he texted Arpana for the first time. *Id.* at 1396.

7. Detectives *Mirandized* Mr. Johnson and then continued questioning him. When asked whether Mr. Johnson went over to Arpana's apartment after the 3:02 am call, he answered "I don't think so." *Id.* at 1346. Mr. Johnson's 3:02 am call was the last call that Ms. Jinaga's phone received.

8. By examining Mr. Johnson's phone, detectives were able to determine that he had not called Ms. Jinaga at 10 or 11 am on November, 1, 2016 as he had claimed. Instead, on the morning of November 1st, after Ms. Jinaga had been murdered, but two days before her body was discovered, Mr. Johnson Googled the location of several pawnshops and, without any prior planning, drove to the Canadian border and attempted to leave the United States. What occurred

[1] The investigation of the King County Medical Examiner's Office and Redmond Police Department strongly suggests that Ms. Jinaga was already dead at this point.

at the border is to some degree unclear. In interviewing one of Cameron Johnson's cousins, Detective Sowers, one of the Redmond Police Department Detectives investigating the homicide, described Cameron's conduct at the border as follows: "he basically tried to blow the checkpoint, went around the stop, and they stopped him. And when they asked him, he said 'I don't know what I'm doing here". Appendix G, 653. Mr. Johnson was not permitted to enter Canada because he lacked the proper documentation and he "didn't really have a reason to go up [there]". Appendix F, 1381. According to Mr. Johnson, border officials "pulled me in and searched my car". *Id.*

9. After being denied entry to Canada, Mr. Johnson attended a party of a friend in Everett. During that party he claimed to sustain injuries to his elbow and knee while wrestling with another partygoer. *Id.* at 1383. Detectives noted that Mr. Johnson was noticeably limping when he was interviewed by police on both November 3rd and November 5th, 2008. Appendix H, 2932. During the course of the investigation, detectives also learned that Cameron stated to friends and family: "What if I did this [the murder] and I don't remember". Appendix I, 5008.

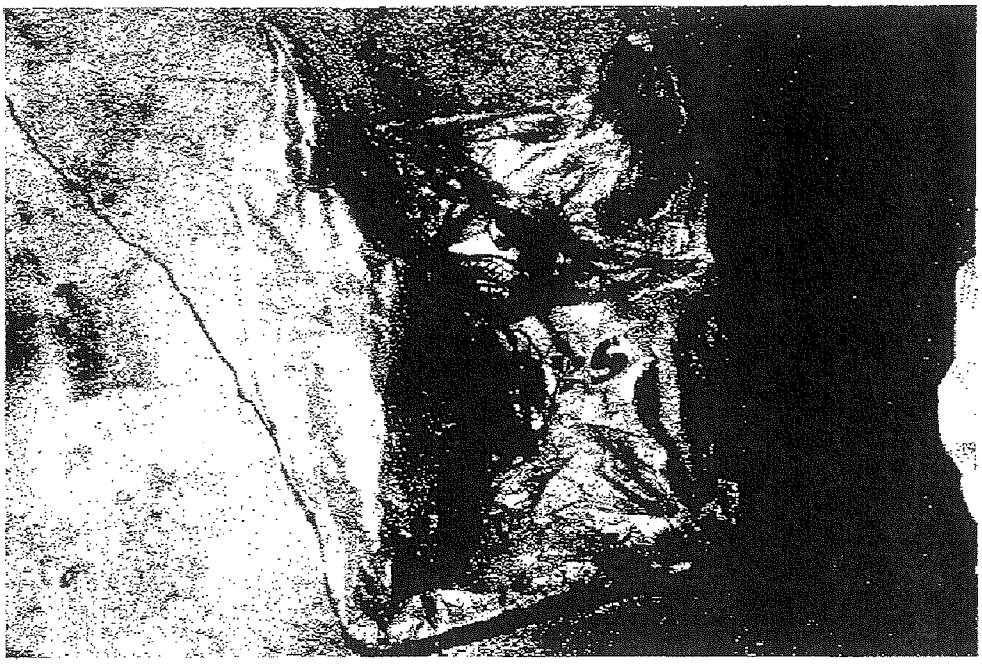
10. Much of the State's investigation focused on the collection and analysis of DNA evidence. In all, more than 50 items have been submitted by the State to four separate laboratories for a variety of methods of DNA analysis. The Washington State Patrol Crime Lab DNA section performed DNA analysis on items from this case beginning in late 2008 and have continued to do analysis as recently as December 2015[2]. In 2010, the State contracted with Bode Technologies, a DNA laboratory in Virginia to conduct Y-STR analysis. In 2010 and again in 2015 the State contracted with another private laboratory, Sorenson Forensics to conduct

[2] Some analysis done by the WSPCL in September 2015 was at the request of the defense.

additional Y-STR testing. In November 2015, the State retained Cybergenetics to use its TrueAllele Casework expert system to analyze several DNA samples.

11. Of the more than 50 items submitted for DNA analysis through the course of this investigation, the State submitted five to Cybergenetics: LDS-103 (Bottle of Castrol Syntec motor oil), LDS-101 (Robe-4), LDS 101(Robe-6), NKP-20 (tape-end), NKP-20 (tape end)[3].

12. When Ms. Jinaga's body was discovered by crime scene investigators, it was covered motor oil. A burnt paper towel was on the ground by her feet. Investigators found the bottle of Castrol Syntec motor oil in a plastic Kohl's shopping bag along with an oil-soaked shoelace in a dumpster outside the Valley View apartments (see below).



13. Investigators were able to determine that the oil from the Syntec

[3] As will become evident from the description below, although five separate samples were submitted to TrueAllele, the five samples came from 3 items.

container was the same as the oil which covered Ms. Jinaga's body.[4] Multiple rounds of DNA analysis were done on the oil container. In 2009, an analyst from the WSPCL found that DNA profile from the motor oil bottle was "mixed in origin and that Arpana Jinaga and Cameron Johnson are included as possible contributors two this profile." Appendix J, 3031. On December 7, 2015, the WSPCL issued an updated report which stated that the DNA profile from the motor oil bottle was a mixed profile of Ms. Jinaga and Mr. Johnson and that "It is 120 million times more likely that the observed DNA profile occurred as a result of a mixture of Arpana Jinaga and Cameron Johnson, than if it originated from Arpana Jinaga and an unknown unrelated individual selected at random from the U.S. population." Appendix K, 17537. In its December 17, 2015 report, Cybergenetics asserted that the match between the oil bottle and Mr. Johnson was "203 million times more probable than a coincidental match to an unrelated Caucasian person." Appendix L, 3.

14. After the homicide, Ms. Jinaga's bathrobe was among the items found in the apartment building's dumpster. Much of the robe was covered with oily residue and it tested positive in several spots for phenolphthalein, a presumptive test for the presence of blood. Several DNA swabs were taken from the robe. Robe-4 was a swab taken from the rear of the left shoulder of the robe. Several months later, the WSPCL retested the same area of the robe (the rear left shoulder). The second left rear shoulder sample is Robe-6.

15. On January 30, 2009, the WSPCL issued a report asserting that Robe-4 was a mixed DNA profile, originating from at least two individuals, that the major contributor to the mixture was Arpana Jinaga and that Mr. Fair was "included as a possible contributor to this profile" and

[4] This analysis was performed by BP, the corporation which manufactured Castrol Syntec.

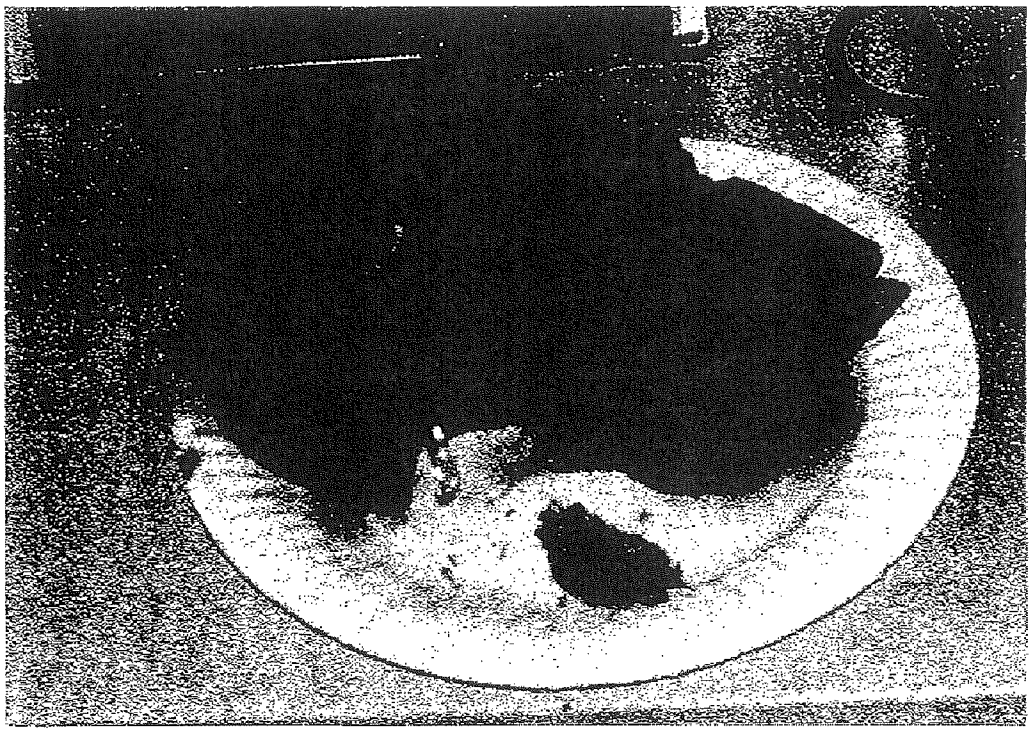
that "1 in 340 individuals is a potential contributor to this profile". Appendix J, 3032. Almost six years later, the WSPCL recalculated the likelihood of Mr. Fair being a potential contributor using an updated database and found that 1 in 339 individuals is a potential contributor. Appendix M, 16439. These numbers were calculated using a combined probability of inclusion (CPI). On December 7, 2015, at the request of the King County Prosecuting Attorney's Office, the WSPCL recalculated the data using a POPSTATS Likelihood Ratio[5] and found that the major component of Robe-4 matched Arpana Jinaga and that it was 120 times more likely that the DNA profile observed as a result of a mixture between Ms. Jinaga and Mr. Fair than if it originated from Ms. Jinaga and an unknown person. Appendix K, 17537. In its December 17th, 2015 report, Cybergenetics asserted that a match between Robe-4 and Emanuel Fair is "3.89 billion times more probable than a coincidental match to an unrelated African-American person." Appendix L, 3.

16. On August 18, 2009, the WSPCL issued a report stating that the DNA typing from Robe-6 was a mixture from "at least two individuals" with the major component belonging to Arpana Jinaga. Mr. Fair was "included as a possible contributor to the minor male component" with "1 in 6 individuals [as] a potential contributor". Appendix N, 3246. The sample was recalculated on October 15, 2015 using the updated database and returned the same result. Appendix M, 16443. In December 2015, the WSPCL recalculated the Robe-6 mixture using the POPSTATs likelihood ratio requested by the KCPAO and claimed that "it is 1000 times more likely that the observed DNA profile occurred as a result of a mixture of Arpana Jinaga and Emanuel Fair than if it originated from Arpana Jinaga and an unknown unrelated individual

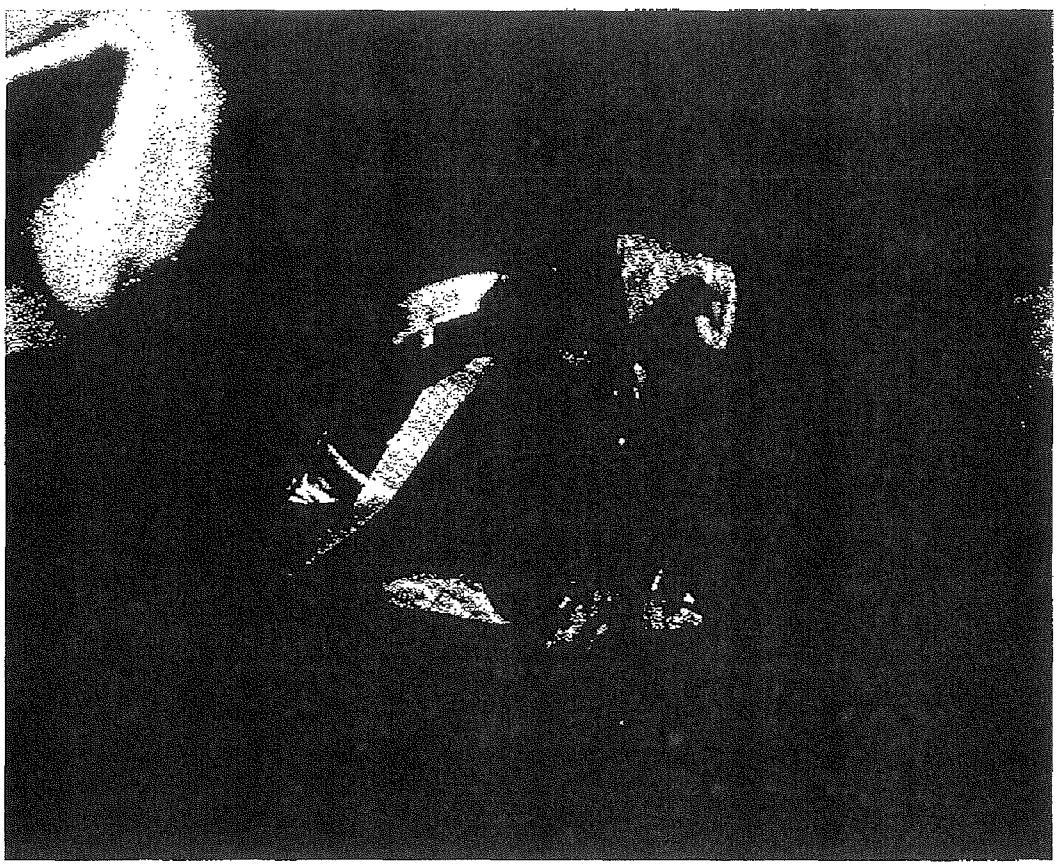
[5] The anticipated *Frye* hearings in this case will address numerous issues related to CPIs, Likelihood Ratios, and Probabilistic Genotyping. For the purposes of this motion, there will be a more limited discussion.

selected at random from the U.S. Population." Appendix K, 17538. Cybergenetics' December 17, 2015 report states that a match between Robe-6 and Emanuel Fair is "56.8 million times more probable than a coincidental match to an unrelated African-American person." Appendix L, 2.

17. At some point prior to her death, it appears that the killer gagged Ms. Jinaga. Crime scene investigators found a pair of brown underwear, that forensic evidence suggests was placed in Ms. Jinaga's mouth and secured by a roll of black electrical tape. The underwear was found on a paper plate in Ms. Jinaga's kitchen on top of half-eaten pizza crusts (see below)



and the black electrical tape was found still attached to the roll, on the back of the couch in Ms. Jinaga's livingroom (see below).



18. In March of 2009, DNA analysts from the WSPCL started to examine the roll of tape. The analysts cut a 1 cm piece from the extended end of the tape, swabbed the 1 cm piece for DNA and designed those swabs "tape end". Analysts then swabbed the edge of both sides of the tape roll and designated those swabs "tape side". On April 3, 2009, the WSPCL reported that the DNA profile from tape end was a mixed sample from at least two contributors, that the major component matched Ms. Jinaga and that Mr. Fair "is included as a possible contributor…[and that] it is estimated that 1 in 3.4 million individuals is a potential contributor to this profile." Appendix O, 3024. As to the tape side, the WSPCL found that the DNA profile was mixed sample from at least two contributors, that Ms. Jinaga's DNA matched the major donor but that

no meaningful comparisons could be made to the minor contributors. *Id.* The tape end was re-evaluated using the updated database in October 2015 and returned a probability of 1 in 3.5 million. Appendix M, 16446. When the WSPCL re-evaluated the sample using the POPSTATs likelihood ration in December 2015, the lab reported that "It is 320 billion times more likely that the observed DNA profile occurred as a result of a mixture of Arpana Jinaga and Emanuel Fair than if it originated from Arpana Jinaga and an unknown unrelated individual." Appendix K, 17538.

19. Cybergenetics' December 17th, 2015 report states that "a match between the tape end and Emanuel Fair is 45.7 trillion times more probable than a coincidental match to an unrelated African-American person." Appendix L, 3. The report further notes that a match between the tape sides and Cameron Johnson is 236 times more probable than a coincidental match to an unrelated Caucasian person. *Id.* at 4.

20. As the foregoing paragraphs illustrate, the results obtained by TrueAllele are radically different than those previously obtained by the Washington State Patrol Crime Laboratory.

ANALYSIS OF COMPLEX DNA MIXTURES

The collection and analysis of DNA samples comprises several steps. For the purposes of this motion, many of those steps are not controversial: an item from a crime scene is swabbed; the swab is preserved and brought to a laboratory where it is exposed to detergents designed to wash biological material from the head of the swab and break open the cells to free the genetic material inside; the genetic material is amplified or diluted to a concentration optimal for copying, a specific small section of the genome is preferentially targeted and copied millions of

times; these copies are exposed to a laser which causes them to illuminate. This illumination is captured on an electropherogram, which measures the illumination in relative fluorescence units (RFUs). What results from this process, under optimal circumstances, is information about the alleles present at a standard set of loci.

Loci are markers along an individual's DNA strands. STR DNA analysis process looks at a set of these loci, common to all individuals. At every locus, each individual has two alleles, one inherited from each biological parent. The overall combination of alleles at the standard loci are sufficiently different from person to person that it widely accepted that no two people, except for identical twins have the same DNA profile. *See State v. Bander*, 150 Wash.App.690, 699 (2009). An electropherogram will show, in a general sense, what alleles are present at which loci, and to what strength, in terms of RFU.

When there is a sufficient amount of high quality DNA in a sample and that sample is only of a single individual, this process can be relatively straightforward. However, a number of factors can complicate DNA analysis, especially when there are multiple contributors and smaller amount of DNA. Peaks can exist on an electropherogram that appear identical to alleles that are actually non-allelic events or "stutter". Stutter peaks are usually much smaller than a neighboring peak and are one allele number less. For example, in a sample where the DNA from more than one individual is present (a mixed sample) if the major donor is an allele 12 at locus FGA, and that 12 appears at peak height of 3000 RFU, than an 11 at a peak height of 300 RFU could either be stutter or the low level allele of a minor contributor to the mixture. This means a sample could appear to have biological information from another individual, when in fact, the peaks that appear to be alleles are actually just artifacts of the amplification process. Similarly, a

peak disregarded as stutter could actually be the biological information left by the actual perpetrator of the crime. In mixed samples, allelic drop-out can occur, meaning that alleles which actually are present, do not appear in the electropherogram or anywhere else in the data. Allelic drop-in can occur as well, where alleles that were not a part of the actual sample can contaminate the data. In samples where minor contributors may appear at lower RFUs, it can be difficult to distinguish the background "noise" of the electropherogram (i.e. non-allelic peaks) from low level alleles. In samples with multiple alleles present, alleles of minor donors can be masked by the presence of higher RFU alleles of the major donor. For example, an allele 12 at FGA which appears at 3000 RFU could represent (1) one allele of the major contributor, (2) both alleles of the major contributor if the major contributor was a homozygote at FGA[6], (3) one allele from the major contributor and a lower RFU 12 from a minor contributor that is masked by the higher RFU 12, or (4) a "12, 12" homozygote from the major contributor and a masked lower RFU 12 from a minor contributor. *See also* Appendix P, *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* (2010), Section 3.1.

In a complex DNA samples there is a considerable degree of uncertainty. Alleles can appear present that are in fact not even human DNA (stutter), alleles may be discounted as not being human DNA when in fact they truly are (stutter), alleles may be invisible but actually present (drop-out), present but not actually from the sample in question (drop-in), or present but invisible because they are masked by the allele of the major donor. This brief description does not begin to capture all of the uncertainty that exists around the analysis of complex DNA samples, but merely highlights some of the most common challenges. It is therefore exceedingly

[6] That a contributor is a homozygote means that he or she inherited the same allele number from both their mother and their father at a particular locus (e.g. 12, 12 at FGA)

difficult in complex samples to assign types (specific alleles at specific locations) or genotypes to particular individuals. The samples analyzed by Cybergenetics are all complex samples.

The Washington State Patrol Crime Laboratory and defense experts are able to analyze complex DNA samples. This is not to say that the method of analysis performed by the WSPCL is generally accepted or that their conclusions are correct, because the defense will assert in later motions that they are not. However, the work done by State experts other than TrueAllele is transparent, meaning that the defense can evaluate the work, discern the methods used, and mount appropriate challenges. The work done by the TrueAllele Casework is unique in many respects, one of them being that it is opaque: none of the defense experts, and it seems no one except for Dr. Perlin, knows how TrueAllele gets the results it claims to get.

TRUEALLELE CASEWORK

I. Introduction

According to its creator Dr. Perlin, TrueAllele Casework is an artificial intelligence system that replicates human expertise and makes inferences about the DNA evidence it claims to examine. Dr. Perlin claims that TrueAllele replicates human expertise by employing a complex mix of Bayesian mathematics and Markov Chain Monte Carlo sampling to "deconvolute," or untangle and separate complex mixtures of DNA.

Bayes' Theorem is a mathematical formula that addresses whether a particular theory (e.g., the hypothesis that a certain set of alleles is present at a certain locus) is true. In addition to using Bayesian mathematics, Markov Chain Monte Carlo methods allow the system to conduct random sampling of thousands or tens of thousands of possible answers to a particular problem. In a general sense, Dr. Perlin claims that TrueAllele looks at electropherogram peaks developed

by a particular crime lab and conducts extensive random sampling to determine what combination of alleles best describes the data that is shown. Dr. Perlin claims his software generates a list of the most probable alleles at each location from a piece of evidence. He asserts that his computer program essentially replaces the work done by a human analyst in "calling" the alleles found in a sample, and that his computer program is more specific and accurate than any human analyst in that it considers more data and more possible solutions than would be possible for a human being. No human being can replicate the process used by the computer to reach its conclusions and so no human can explain these results. Thus, without an adequate understanding of how this program works, the parties would be required to trust the results of a "black box" that provides highly inculpatory evidence without adequate explanation.

II. Flaws

Bayes' Theorem and Markov Chain Monte Carlo sampling are both longstanding mathematical concepts that are used in a variety of fields. Nonetheless, other scientists studying TrueAllele Casework have found that it is rife with potential flaws that can only be evaluated by disclosure of TrueAllele's source code.

Embedded within TrueAllele are a number of mathematical models which model the molecular behavior of DNA (e.g. stutter, drop-in, drop-out, masking, peak-height imbalance). Dr. Perlin has never published nor disclosed what exact models he uses to address these and other molecular behaviors. Appendix A and B. This prevents the defense from sufficiently analyzing TrueAllele for three separate reasons. First, unless defense experts know exactly what molecular models are used in TrueAllele, the defense will be unable to determine whether the models are correct. *Id.* Second, to incorporate the models into the software, the mathematical

formulas of the models need to be "translated" into the computer source code of TrueAllele. *Id.* Even if the underlying mathematical models which Dr. Perlin used are correct, if the models are incorrectly translated into the source code of the software than TrueAllele would not appropriately model molecular behavior. *Id.* Third, even if the model and translation of the model into the source code are correct, if the code causing the model to execute, or be applied, is incorrect than the correct mathematical formulas could be applied in incorrect ways. *Id.*

Dr. Perlin may claim that TrueAllele has been put through multiple "validation studies" with positive results and so no further examination of how the computer works is necessary. This is misleading. If TrueAllele were attempting to solve an algebra problem which had a single, correct answer, than such validation studies could be sufficient. The analysis of complex DNA samples entirely different, however. TrueAllele produces likelihood ratios (LRs) which are statements of probability. There is no immutable "correct" answer when analyzing complex DNA samples, only probabilities. Because there is no "correct" answer, TrueAllele cannot be validated merely be looking at its results. Appendix B, 4.

Recently, errors in coding of STRmix a probabilistic genotyping system were found to have affected likelihood ratios in criminal cases, so the concerns of coding errors in such programs are not abstract[7].

[7] *Queensland authorities confirm 'miscode' affects DNA evidence in criminal cases.* The Courier-Mail: March, 20 2015:www.couriermail.com.au/news/queensland/queensland-authorities-confirm-miscode-affects-dna-evidence-in-criminal-cases/news-story/833c580d3f1c59039efd1a2ef55af92b.

Furthermore, some of the studies that have attempted to examine TrueAllele, even without the benefit of the source code, identified significant problems that demonstrate the need for further investigation of the system's reliability.

A. New South Wales Police Force (2011)

In 2011, a review team from the New South Wales (NSW) Australia Police Force conducted an evaluation of TrueAllele and produced a 103 page report. Appendix Q.

The review team found two significant problems with TrueAllele's system. The first arose in the area of artifacts, or non-allelic events in the electropherogram. The review team noted that "TA [TrueAllele] will assess all peaks as potential contributors to the DNA genotypes and will not disregard or give less weight to apparent artifacts." *Id.* 19214. In one example, TrueAllele found it most probable that the alleles at locus FGA were 20.3, 24, in fact in the thousands of sampling runs that TrueAllele performed, it found the incorrect answer almost 60% of the time. *Id.* The NSW team knew that the actual alleles present at the FGA locus were 24, 24, so the result in this instance would not have been inculpatory to a defendant. However, as the NSW team pointed out, if the artifact peak overlapped a smaller actual peak, than the effect could have been to falsely include a suspect's genotype when it was actually not present. *Id.* The NSW team noted: "the inclusion of artefacts may affect the capability of the system to identify the most accurate genotype probability distributions at individual loci which will translate to an effect on the overall LR (likelihood ratio)". In other words, TrueAllele's systems sometimes identifies non-biological artifacts as alleles, this could cause the system to assert the presence of an individual's DNA when it is not actually supported by the evidence.

The NSW team raised even more significant concerns about TrueAllele's modelling of stutter (peaks which are indistinguishable from alleles, but can either be a non-genetic artefact of the amplification process or an actual allele). The report states "when minor components of a mixture are at similar peak heights to stutters, the genotype probabilities are not consistent with what would be expected given a reasonable consideration of stutter contribution." *Id.* at 215. In one particular locus in the sample, TrueAllele found 97.7% of the time over the course of its thousands of runs that the correct allele pairing was 15, 17 even though the 15 was stutter and not an actual allele. *Id.* at 216. The NSW team noted that "The performance of TA [with] respect to the determination of genotype probability distribution for minor contributors at levels in the stutter range was ***considered to be questionable*** and many examples were seen throughout this study." *Id.* at 216 (emphasis added.)

The NSW report also found a "lack of consistency to stutter modeling" especially "when the real peaks are of a similar peak height to stutter peaks." *Id.* at 217, 221. The NSW team suggested that "TrueAllele might assign more realistic probabilities to the correct genotypes if it used stutter modeling parameters determined by laboratory empirical testing." *Id.* NSW stated that Cybergenetics made "a commitment to restore calibration...to deal with stutter and other laboratory dependent parameters by the end of 2011", essentially to improve the system to address these flaws. *Id.* at 221. In discovery, Cybergenetics claims to have disclosed updates to the system. The meaning of these updates is unclear, but none of them appear to address the stutter flaws highlighted by New South Wales.

The NSW report concluded that "the handling of other parameters such as stutter, relative amplification of alleles at a locus, and DNA degradation are not disclosed. This makes it

difficult to determine how TA handles these issues and it has been noted that TA does not perform very well in relation to some of these on weak samples. Therefore TA has an element of the unknown". *Id.* at 234.

B. California Department of Justice 2014

In 2014, the California Department of Justice undertook a comprehensive comparison of TrueAllele and a different probabilistic genotyping software, called STRmix, to determine which system the DOJ would purchase and implement. The California DOJ chose STRmix over TrueAllele and highlighted significant concerns about TrueAllele.

Both STRmix and TrueAllele use Markov Chain Monte Carlo (MCMC) in their analysis. Appendix R, 4. MCMC is a method of randomly sampling data. Although in theory MCMC will find the most probable answer to the particular question, Dr. Perlin has admitted that since a Markov chain is a random experiment sometimes the experiment has a useful outcome and sometimes it does not. Dr. Perlin has also admitted that since MCMC is a random sampling method, there should be some variation among the answers it produces. An inherent part of this method, therefore is that when TrueAllele uses MCMC to determine how likely it is that a particular person's DNA is part of a particular mixture it gives a different answer every time.

In its comparison of TrueAllele to STRmix, the California DOJ found that in some situations running pairs of MCMC runs gave identical results. Appendix S, pg. 31. In most scientific experiments, identical results in duplicated experiments is an indication of success. However, because of the random nature of MCMC, one would almost never expect to see identical results in subsequent MCMC runs. The DOJ noted that the identical likelihood ratios "should not have occurred given the randomness of the MCMC process." *Id.* Furthermore, the

DOJ found that on some occasions different MCMC runs resulted in values that were 6.5 quadrillion times different than one another and "many deviations were observed that could lead to different conclusions about the strength of the evidence (e.g. LR<<1 in one interpretation becoming LR >>1 in another)." *Id.* The importance of this cannot be overstated. The California DOJ found that in some cases, when analyzing the same item of evidence on duplicate runs, TrueAllele produced radically different answers , including answers on one run which strongly included an individual as a donor to an item and another run which strongly excluding the individual as a donor, and vice versa.

The California Department of Justice identified another potential fundamental flaw with TrueAllele. TrueAllele uses the electronic data generated by state crime laboratories, in this case the WSPCL. However, upon entering the electronic data (e.g. the peaks seen in an electropherogram) into the TrueAllele system, the DOJ found that "heights and allele designations were seen to occasionally change from the values listed prior to upload to the server and those listed after processing in the server." Appendix R, 6. The DOJ analyst wrote that this "remains a concern." *Id.* This means that on occasion, TrueAllele transforms data that it receives rather than merely interpreting it. This transformation of data has the potential to fundamentally alter any results produced by the system. By changing the allele designation and strength of the peak, TrueAllele then changes the foundation for all of the assessments that the system makes. Any data about how likely the presence of a particular allele is, is dependent on the system reviewing information about the allele itself including the height of its peak and the fact that the allele itself is present. If the peak height is transformed, or the allele is misnamed

(the actual allele present is an 11.1 and TrueAllele transforms it to an 11) than any data that flows from such an analysis is fundamentally flawed.

The NSW and California DOJ reports demonstrate that law enforcement entities have found significant flaws in TrueAllele's ability to distinguish actual alleles from non-allelic events, generate reproducible results (including that TrueAllele returns results for identical samples which can vary between highly inculpatory to highly exculpatory for the same item), and that it sometimes transforms rather than interprets the data. These flaws, coupled the lack of transparency of how TrueAllele generates its results demonstrates the need for disclosure of the source code.

C. State v. Fair

The discovery provided by the State about TrueAllele's analysis in this case gives rise to many of the same concerns raised by the New South Wales Police and the California Department of Justice. Dr. Chakraborty and Dr. Krane are familiar with Dr. Perlin's publications, the results TrueAllele has generated in other cases, and the data produced in this case. Both experts have significant concerns about the reliability and accuracy of TrueAllele which can only be addressed through review and testing of the source code. *See* Appendix A, B.

One of the documents generated by TrueAllele when it examines an item of evidence is a Data Table. Appendix T, 18077-18083. The table claims to record the "quantitative peak data from STR experiments on biological evidence"; in other words, this is TrueAllele's assessment of the electronic information obtained by the WSPCL. The data table illustrates the presence of some of the concerns of the California Department of Justice about how TrueAllele transforms the peak height data from the laboratory. As an example, at locus D3S1358 (D3) on Robe-6

TrueAllele claims to identify 31 potential alleles just at the D3 locus. *Id.* at 18079. In contrast, the WSPCL identifies 5 alleles at D3. Appendix U, 4817. If all of those 31 potential alleles were from actual biological material, than it would mean that sample included biological material from at least 16 people. In assessing which alleles TrueAllele thinks form the most likely genotypes, TrueAllele only identifies 4 of the 31 as likely belonging to a contributor at the D3 locus. Appendix V, 18120. There are multiple steps (allelic smoothing, stutter, drop-in, drop-out, peak height analysis) that would have to occur from identifying 31 alleles to deciding that 27 of them do not represent biological material that are completely unexplained in either the discovery disclosed, or in any of Dr. Perlin's publications. As Dr. Krane notes: "it is important to know precisely why and how TrueAllele arrived at the results that it arrived at in the case of WA v. Fair. A careful evaluation of the computational steps taken by TrueAllele would allow it to be determined if the program: 1) reflects what is described by Dr. Perlin, 2) is consistent with the practices of the forensic DNA profiling community, 3) is free from bugs and errors, and 4) if TrueAllele can and does provide sufficient explanations for the observed data in this case." Appendix B, 10. Indeed, "Human experts are expected to explain how they arrive at a conclusion…this same expectation can and should apply to a computer program such as TrueAllele. *Id.* at 12. This is especially important in the case of TrueAllele which uses "all allele peak height information including those falling below a threshold level established by almost all of the DNA laboratories of the world…or by the manufacturers of the DNA kits." Appendix A, 4.

Another of the major concerns identified by the California Department of Justice is present in the data in this case as well. In some instances, the DOJ found that different MCMC

runs resulted in outcomes (likelihood ratios) that were 6.5 quadrillion times different than one another and "many deviations were observed that could lead to different conclusions about the strength of the evidence (e.g. LR<<1 in one interpretation becoming LR >>1 in another)." Appendix R. With regard to Robe-4, TrueAllele's 12.17.15 report states: "a match between robe-4 and Emanuel Fair is "3.89 billion times more probable than a coincidental match to an unrelated African-American person." A more careful review of the data disclosed to this point demonstrates that results TrueAllele obtained from Robe-4 ranged from 10^9 (1,000,000,000) to 10^3(1000), meaning that the report could have been written "a match between robe-4 and Emanuel Fair is "1000 times more probable than a coincidental match to an unrelated African-American person", which is much less compelling. On Robe-6 the results "matching" Mr. Fair to the sample ranged between 10^7 (10,000,000) to $10^{3.64}$(roughly 4,300), and on the Tape-end from 10^{14} (100,000,000,000,000) to $10^{4.67}$ (46,773).

One of the chief concerns of the Department of Justice in implementing TrueAllele throughout the State of California was that the vast variability in its results "could lead to different conclusions about the strength of the evidence." Appendix R, S. The Court should have the same concerns here. Without the source code, the defense will be unable to effectively evaluate TrueAllele to determine why its results vary in such a dramatic fashion and to counter any claim by Dr. Perlin as to which result should be viewed as more accurate.

The TrueAllele data in this case also gives rise to concerns separate from those identified either by the New South Wales Police Force or the California Department of Justice. In at least one instance, TrueAllele identified a possible genotype that appears to be entirely absent from the data. One of the possible genotypes for locus D18 in the Robe-6 sample is a possible 16, 23

allele pairing. Appendix V, 18119. TrueAllele found that the probability of this allele pairing to have occurred at locus D18 was just under 1%. In the context of TrueAllele's MCMC analysis of this sample, the computer accepted 50,000 different calculations of possible genotypes and accepted 16, 23 as the genotype at D18 490 times. The data table for D18, which lists all the possible alleles present from the electronic data, down to a height of 10 RFUs, does not report an allele 23 as a being present. Appendix T, 18078. In other words, Trueallele found 490 times that the most likely genotype at a particular locus included an allele that was not even present in the data.

The work of the NSW Police Force, the California Department of Justice, and the defense experts raise significant questions about the functioning and reliability of TrueAllele. A review of the data disclosed so far makes clear that the Court should have similar concerns. The defense cannot adequately evaluate the functioning, veracity, and reliability of TrueAllele without its source code. For Mr. Fair, his life may well be determined by the TrueAllele results. The Court should not permit this to happen without allowing the defense experts to examine the source code of the system.

MEMORANDUM OF LAW

I. CrR 4.7, and the Due Process Clause of the United States and Washington Constitutions Require Disclosure of the Source Code

Washington law provides for liberal discovery in criminal cases. CrR 4.7. This is evidenced by Washington's discovery rule, which exceeds the scope of discovery rules adopted in many other jurisdictions. *See State v. DeWilde*, 12 Wn.App. 255 (1974) (Washington rule

broader than ABA draft). "Persons who are on trial for the commission of a crime are entitled to all information contemplated under Criminal Rule 4.7." *State v. Coe*, 101 Wn. 2d 772, 784 (1984). The prosecutor is obligated, upon request, to seek out material information held by third parties. *See* CrR 4.7 (d). In addition, the Court can order the disclosure of materials outside of the required disclosures under CrR 4.7(a), (c), (d), upon the showing of materiality. CrR 4.7, *State v. Norby*, 122 Wn.2d 258 (1993). "The principles underlying CrR 4.7 require meaningful access to copies based on fairness and the right to adequate representation." *State v. Boyd*, 160 Wn.2d 424, 433, 158 P.3d 54, 59 (2007). "The discovery rules "are designed to enhance the search for truth" and their application by the trial court should 'insure a fair trial to all concerned, neither according to one party an unfair advantage nor placing the other at a disadvantage.'" *Id.* (quoting *State v. Boehme*, 71 Wn.2d 621, 632–33, 430 P.2d 527 (1967)). The right to due process in a criminal trial is, in essence, the right to a fair opportunity to defend against the State's accusations. *Chambers v. Mississippi*, 410 U.S. 284, 294, 93 S. Ct. 1038, 1045, 35 L. Ed. 2d 297 (1973).

As part of its obligations, the State must disclose evidence that is material and favorable to a defendant "to protect against surprise that might prejudice the defense." *State v. Barry*, 184 Wash. App. 790, 796 (2014). The suppression of evidence favorable to the defense violates due process, irrespective of the good faith or bad faith of the prosecution. *Brady v. Maryland*, 373 U.S. 83, 87 (1963). A defendant's constitutional due process right to disclosure relates to evidence which is favorable to the defendant and material to guilt or punishment. *Id. See also United States v. Agurs*, 427 U.S. 97 (1976). This duty of disclosure encompasses impeachment evidence, as well as, exculpatory evidence. *United States v. Bagley*, 473 U.S. 667, 676

(1985)(recognizing the significance of impeachment evidence in holding "...it is on such subtle factors as the possible interest of the witness in testifying falsely that a defendant's life or liberty may depend.").

Materiality is not addressed in the context of computer source code in Washington State case law. In the context of records sought by the defense, courts have found that "materiality" requires that the defendant "make a particularized factual showing that information useful to the defense is likely to be found in the records." *State v. Diemel*, 81 Wn. App. 464, 469 (1996). A showing of "materiality does not require demonstration by a preponderance that disclosure of the suppressed evidence would have resulted ultimately in the defendant's acquittal." *Kyles v. Whitley*, 514 U.S. 419, 434 (1995).

"The proper standard of materiality must reflect our overriding concern with the justice of the finding of guilt ... It necessarily follows that if the omitted evidence creates a reasonable doubt that did not otherwise exist, constitutional error has been committed. This means that the omission must be evaluated in the context of the entire record." *United States v. Agurs*, 427 U.S. 97, 112-113 (1976). Evidence is material "if there is a reasonable probability that it would impact the outcome of the trial." *State v. Gregory*, 158 Wn. 2d 759, 791, 147 P.3d 1201 (2006) (citing Ritchie, 480 U.S. at 57). In *Ritchie*, 480 U.S. at 57-58, for example, the court held that in camera review of the child protective services' file was warranted even though it was impossible to say in advance whether any information in the records would actually support the defendant's arguments.

The prosecution is obliged to disclose the source code pursuant to *CrR* 4.7(d). The State anticipates calling Dr. Perlin as an expert at trial, Dr. Perlin possesses the source code

of TrueAllele and the substance of his testimony will largely consist of explaining the results generated by the software. Washington courts have joined an number of other jurisdictions in finding that before computer generated models or simulations are deemed admissible, the party offering such evidence must demonstrate (1) the computer is functioning properly; (2) the input and underlying equations are sufficiently complete and accurate (and disclosed to the opposing party so that they can be challenged); and (3) the program is generally accepted by the appropriate community of scientists for use in the particular situation at hand. *State v. Sipin*, 130 Wash.App. 403, 414 (2005).

The defense has made a sufficient showing that potential errors exist in the software. The results of those errors could either result in the creation of erroneously inculpatory evidence or the masking of exculpatory evidence. The defense has also demonstrated that the materials disclosed by Dr. Perlin are insufficient to analyze his software and that such analysis can only be conducted upon the source code itself. In addition, the defense has made a sufficient showing that the Court should order disclosure of the source code under *CrR 4.7(e)* as the declarations of Dr. Krane and Dr. Chakraborty, as well as the forgoing motion have demonstrated the materiality of the source code.

II. Disclosure of Source Code, When Required to Resolve Contested Issues, is Both Common and Justified Under Federal Case Law

The paucity of state law on the issue of disclosure of source code is largely due to such disputes almost exclusively occurring in federal court. The defense consulted with an expert in patent litigation, an area where the disclosure of source code is routinely addressed. Brian Ferguson is a partner at the firm of Weil, Gotshal, and Manges LLP and co-chair of the

firm's Patent Litigation Practice Group. Appendix E, 1. Mr. Ferguson has practiced law for 24 years has represented clients including Apple Inc., General Electric Company, and National Semiconductor Corporation (a subsidiary of Texas Instruments) in patent disputes in federal court and before the International Trade Commission. *Id.* Mr. Ferguson has dealt with the issue of production of source code in response to discovery requests on numerous occasions, indeed "this issue comes up very frequently in patent cases". *Id.* at 2. In patent cases involving patents related to computer functionality "the production of source code by the defendant in response to discovery requests is almost always required." *Id.* at 5. Mr. Ferguson explains: "the reason why the production of source code is required is simple: it is only the source code that determines the instructions present in software that a computer will execute in order to carry out a particular function...the source code is therefore frequently the single most critical piece of evidence in patent cases". *Id.* at 5, 6.

In essence, patent cases cannot be decided without the disclosure of source code so as a result "Courts routinely order production of source code". *Id.* at 6. This has occurred in numerous recorded cases including those listed in Mr. Ferguson's declaration. *Id.* Indeed, production of source code has become "so commonplace in patent litigation that several jurisdictions have adopted local rules expressly requiring it as part of routine discovery." *Id.* at 7. Mr. Ferguson has represented parties in litigation when disclosure of source code was ordered by the court. *Id.* at 8. In such cases, district courts issue protective orders to protect any such disclosures. During Mr. Ferguson's 24 years of practice focusing on this issue he is not "personally aware of any instance wherein a party's claim to trade secret protection for source

code was lost a result of either an intentional or inadvertent violation of a protective order." *Id.* at 8-9.

In *U.S. v. Budziak*, the defendant was charged with distribution of child pornography and as part of his defense sought discovery of the software the FBI used in its investigation into his online file sharing activities. *U.S. v. Budziak*, 697 F.3d 1105 (2012). In its investigation, the FBI used a program called EP2P which "purportedly allow[ed] the FBI to view all files that a particular user on the file-sharing network is making available for download." *Budziak*, 697 F.3d 1105, 1107. Mr. Budziak moved for discovery related to the software, which was denied. *Id.* at 1111. Specifically Budziak argued that discovery of the EP2P software was material to his defense. *Id.* The Ninth Circuit agreed and found that access to the EP2P software was crucial to Budziak's ability to assess the program and the testimony of the FBI agents who used it to build the case against him. *Id.* at 1112. The Court specifically noted that Buziak's computer experts stated that they could discern useful information through discovery of the software. In such cases, "criminal defendants should not have to rely solely on the government's word that further discovery is unnecessary. This is especially so where, as here, a charge against the defendant is predicated largely on computer software functioning in the manner described by the government, and the government is the only party with access to that software." *Id.* at 1113.

Local federal court decisions in the civil cases contemplate the disclosure of source code under a protective order. In *Avocent Redmond Corp. v. Rose Electronics, Inc*, 242 F.R.D. 574 (W.D. Wash. 2007), the parties, appearing before Judge Pechman in a patent dispute involving Keyboard-Video-Mouse switch technology, agreed to protective orders of information including source code. In *Telebuyer, LLC v. Amazon.com, Inc.*, No. 13-CV-1677, 2014 WL 5804334, at 1

(W.D. Wash. July 7, 2014), Amazon moved for a protective order to cover the disclosure and handling of confidential source code. In determining whether to issue a source code, the Court analyzed "whether a particularized harm will result from disclosure of information to the public, then it balances the public and private interests to decide whether a protective order is necessary." *Id.* at 1. While the Court adopted most of Amazon's stipulations with regard to a protective order, the source code was nonetheless disclosed.

The use of protective orders in Superior Court to protective sensitive information other than source code is widespread, especially when sexually explicit images of children are potential evidence in a case. In King County Superior Court, when one of the parties retains a neuropsychologist, propriety testing information is routinely transferred between experts for the parties, often without the use of a protective order. In *State v. Grenning*, 169 Wash. 2d 47 (2010), the defendant was charged with 72 counts of child sex crimes. The State seized computers which contained sexually explicit pictures of the defendant's child victims from his home as part of its investigation and only allowed him limited access to the hard drives such that he was unable to obtain an expert willing to examine the hard drives. *Id.* at 49. Specifically, the trial court, in the original protective order required that the images be examined only in the County building during the hours of 8:30am-4:30 pm Monday through Friday. *Id.* The court recognized that denying the defense sufficient access to critical evidence had constitutional implications and that "[c]ourts have long recognized that effective assistance of counsel, access to evidence, and in some circumstances, expert witnesses, are crucial elements of due process and the right to a fair trial." *Id.* at 54-55. Ultimately, the court held that the defendant was entitled to a mirror image copy of his computer's hard drive. The court noted that in objecting to

disclosure, the State failed to "offer[] any more than mere allegations that the evidence might be improperly disseminated by the defense team." *Id.* at 54.

Much of the court's decision rested upon the earlier decision in *State v. Boyd*, 160 Wash.2d 424 (2007), a case involving "tens of thousands of 'commercial' images of unidentified minors engaged in sexually explicit conduct." The trial court denied a defense motion to compel the State to provide a mirror image of the hard drive to enable independent testing by a defense expert. *Id.* The State Supreme Court reversed, requiring disclosure of the material with a protective order, and explained that:

> Effective assistance of counsel, access to evidence, and in some circumstances, expert witnesses, are crucial elements of due process and the right to a fair trial. The Fifth Amendment to the United States requires that prosecutors make available evidence "favorable to an accused ... where the evidence is material either to guilt or to punishment." *Brady v. Maryland,* 373 U.S. 83, 87, 83 S.Ct. 1194, 10 L.Ed.2d 215 (1963). The Sixth Amendment right to effective assistance of counsel advances the Fifth Amendment's right to a fair trial. That right to effective assistance includes a "reasonable investigation" by defense counsel. *See Strickland v. Washington,* 466 U.S. 668, 684, 691, 104 S.Ct. 2052, 80 L.Ed.2d 674 (1984); *In re Pers. Restraint of Brett,* 142 Wash.2d 868, 873, 16 P.3d 601 (2001). It also guarantees expert assistance if necessary to an adequate defense. *State v. Punsalan,* 156 Wash.2d 875, 878, 133 P.3d 934 (2006). Supporting the right to effective representation, CrR 4.7(h)(4) provides that notwithstanding protective orders, the evidence must be disclosed "in time to permit ... beneficial use."

Id. at 434-35.

State courts in other jurisdictions have ordered propriety software source code to be disclosed for testing by the defense. *See, e.g., State v. Chun,* 191 N.J. 308 (2007) (ordering Draeger Safely Diagnostics Inc. to disclose source code for the purpose of independent software testing); *Commonwealth v. Camblin*, 471 Mass. 639 (2015) (noting a district court order mandating disclosure of Alcotest's source code); *State v. Underdahl* 767 N.W.2d 677 (2009)

(reversing the Court of Appeals and ordering production of complete computer source code for Intoxilyzer 5000EN pursuant to defense request).

As is evident from Mr. Ferguson's declaration and the case law cited above, if Cybergenetics was involved in a patent lawsuit in federal court related to their software, there is a great likelihood that the source code would be ordered disclosed, pursuant to a protective order. Similarly, information of a much sensitive and devastating nature, child pornography, is routinely disclosed to the defense under the terms of protective orders in this court. It would be an absurd result to grant Cybergenetics more protection for its source code than it would be afforded in a civil lawsuit, where no individual's liberty is at stake. It would be equally absurd to privilege Cybergenetics' source code in a way that child pornography is not protected against disclosure.

The core question before this court is one of materiality. Based on the declaration of the defense experts, the NSW report, the California Department of Justice report, and a review of some of the discovery disclosed in this case, there are evident flaws and irregularities in the functioning of TrueAllele Casework. The reliability of the system and the veracity of the results cannot be effectively evaluated without the source code in this case. Accordingly, the defense requests that the Court compel disclosure of the TrueAllele Casework Source Code under an appropriate protective order.

III. The Confrontation Clause Requires Disclosure of the Source Code

The Sixth Amendment of the United State Constitution provides: "[i]n all criminal prosecutions, the accused shall enjoy the right ... to be confronted with the witnesses against him.

Melendez-Diaz v. Massachusetts, 557 U.S. 305, 309, 129 S. Ct. 2527, 2531, 174 L. Ed. 2d 314 (2009). A witness's testimony against a defendant is thus inadmissible unless the witness appears at trial or, if the witness is unavailable, the defendant had a prior opportunity for cross-examination. *Id.* at 309.

In *Melendez-Diaz*, the Court considered whether certificates of analysis were "testimonial" and thus implicated the confrontation clause. *Id.* at 310. In its analysis, the Court noted that not only were the affidavits made under circumstances which would lead an objective witness to reasonably believe the statement would later be used in trial, but under State law, the *sole purpose* of the affidavits was to provide "prima facie evidence of the composition, quality, and the net weight" of the analyzed substance. *Id.* Consequently, the Court found the certificates of analysis to be testimonial and barred their introduction absent a showing that the analysts were unavailable to testify and that the petitioner had a prior opportunity to cross-examine. *Id.* at 311

The Washington Court considered the application of *Melendez-Diaz* in *State v. Lui*, 179 Wash.2d. 457 (2014). In *Lui*, the Washington State Patrol Crime Laboratory sent DNA samples to a private laboratory for analysis. *Lui*, 179 Wash.2d. 466. Gina Pineda, the private laboratory's associate director and technical leader testified about the testing of the samples at trial although she did not personally participate in or observe the tests. *Id.* Ms. Pineda did use the electronic data produced during the testing process to create a DNA profile that reflected "[her] own interpretation and [her] own conclusions. *Id.*

The Court found that, if DNA evidence, or other scientific or technical evidence, is used against a defendant in court, the confrontation clause is implicated and that an appropriate

witness must be must be subject to cross-examination. The witness required by the confrontation clause is the person who has made the final comparison that is used against the defendant . *Id.* at 484. The Court distinguished the application of the confrontation clause from other contexts by explaining "the DNA testing process does not become inculpatory and invoke the confrontation clause until the final step, where a human analyst must use his or her expertise to interpret the machine readings and create a profile." *Id.* 486. In *Lui,* because Pineda had created the profile that was used to incriminate the defendant, she was the appropriate witness to testify and Mr. Lui's confrontation clause rights were not violated. *Id.* at 489. Accordingly, the Court found that the only "witness against" the defendant in the course of the DNA testing process is the final analyst who examines the machine-generated data, creates a DNA profile, and makes a determination that the defendant's profile matches some other profile." *Id.*

According to Dr. Perlin, the TrueAllele Casework system replaces human expertise, evaluates data, makes inferences about the DNA evidence it claims to examine, and then generates a "match" to potential reference samples. Dr. Perlin gives voice to the results of the software, but he does not interpret data that the software generates, he merely reports the results. In *Lui,* the Court found that the defendant's right to confrontation was not violated because the expert who drew the inferences about the DNA data (e.g. whether it matched the reference sample of a particular individual) was Dr. Pineda, the witness who testified and was cross-examined. TrueAllele draws all of the inferences generated in the TrueAllele report. If TrueAllele actually does what Dr. Perlin claims, than it is, in fact, the expert, and Dr. Perlin is merely the mouthpiece reporting the results.

The FBI Quality Assurance Standards for United States laboratories that conduct DNA testing and participate in the FBI's Combined DNA Index System (CODIS) define a DNA analyst as an individual who "conducts and/or directs the analysis of forensic samples, *interprets data and reaches conclusions*" (FBI Quality Assurance Standards for Forensic DNA Testing Laboratories 2, 2009, emphasis added). If TrueAllele performs as Dr. Perlin claims that it does, than Mr. Fair's confrontation rights will only be satisfied by disclosure of the source code, as it is the operation of TrueAllele, rather than the expertise of Dr. Perlin that must be challenged.

CONCLUSION

Based on the above, Mr. Fair respectfully requests that the court grant his motion for disclosure of the TrueAllele source code under a protective order.

DATED THIS 7th DAY OF March, 2016

Respectfully submitted,

s/ Benjamin Goldsmith
WSBA #35017
Attorney for Defendant
King County Department of Public Defense
The Defender Association Division
810 3rd Ave. Ste 800
Seattle, WA 98104
Telephone: (206) 477-8700 x78734
Fax: (206) 447-2349
Email: Benjamin.Goldsmith@kincounty.gov

s/ Paul Vernon
WSBA #40641
Attorney for Defendant
King County Department of Public Defense
The Defender Association Division
810 3rd Ave. Ste 800
Seattle, WA 98104
Telephone: (206) 477-8700 x78792
Fax: (206) 447-2349
Email: paul.vernon@kingcounty.gov