SUPERIOR COURT OF WASHINGTON FOR KING COUNTY

STATE OF WASHINGTON, Plaintiff, vs. EMANUEL FAIR, Defendant.	No. 10-1-09274-5 SEA DECLARATION OF JOHN DONAHUE

I, John Donahue, hereby declare as follows:

1. I am over 18 years of age and I am competent to make this declaration.
2. I hold the following academic degrees: a Bachelor of Science from the University of Tennessee with a major in zoology and a Master of Arts from Indiana University with a major in microbiology.
3. I am currently employed as the DNA Technical Leader at the Beaufort County Sheriff's Office Forensic Services Laboratory in Beaufort, South Carolina.
4. I am familiar with Cybergenetics, and its TrueAllele software. TrueAllele is a probabilistic genotyping computer system that interprets DNA evidence using a statistical

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955

model. I am familiar with the many TrueAllele validation studies that have established the reliability of the method and software.

5. I (my lab) have used TrueAllele for approximately three years. We purchased the software in 2013, performed a validation study over the course of approximately two years, and implemented it into casework in January 2016.

6. We have never requested the source code for the TrueAllele software. I do not believe the source code is necessary for determining the reliability of TrueAllele because our validation study demonstrated to us that TrueAllele generated the expected results when examining single source samples containing DNA from one contributor as well as laboratory-prepared mixtures containing from two- to five- contributors.

7. In one part of this validation study we tested DNA extracted from single source blood samples with TrueAllele and generated match statistics for each individual sample. We then calculated match statistics for each sample with Popstats. Popstats is a forensic DNA statistics calculator program that is provided with the FBI's Combined DNA Index System (CODIS) software and is commonly used among forensic DNA laboratories across the United States. We found that the calculations for single source samples were essentially the same, with some very slight differences attributable to rounding and how each program handles calculations for alleles that were not observed in the FBI's allele frequency data.

8. We also tested 60 amplifications of mixed DNA to determine if TrueAllele could identify the contributors to the samples while also eliminating non-contributors. These mixed samples contained anywhere from two to five known contributors. The mixtures were created in different combinations with different mixture weights; for example, Mixture

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955

M1 was a mixture of DNA from samples 21 and 31 at a calculated mixture weight of 50:50. The weights of each contributor were calculated from the original concentration data so we were able to predict what the mixture weights should be. This was important for our study because as part of the mixture deconvolution process, TrueAllele calculates a mixture weight for each component of the mixture.

9. When we processed our mixed DNA samples we found was that TrueAllele was accurate in determining the mixture weights in comparison to the predicted mixture weights for most of our samples. In some mixtures we found that TrueAllele calculated a mixture weight that was different from what we had estimated when preparing the sample. In these instances we examined the original data as generated by the DNA sequencer and then calculated the mixture weight from the sequencer data by hand. When we did this we found that TrueAllele's calculations correlated with our hand calculations; that is, the original predicted mixture weight as calculated from the DNA sample concentration data was incorrect. It appears that we had introduced error into the mixture samples at some point in the preparation process, either through an inaccurate measure of the sample concentration or through inaccurate mixing of the samples. As a result, the actual mixture weights in these samples diverged from what we had predicted; however, this divergence was identified by TrueAllele and re-calculation by hand confirmed it.

10. Reproducibility was another factor we assessed in our validation study. If a scientific process is valid, it should be able to be reproduced. As part of the assessment of our mixtures we found that the TrueAllele process was reproducible between independent replicate runs. TrueAllele was able to produce similar mixture weights for the mixture components in different runs, while at the same time excluding non-contributors each

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955

time. We also determined through the validation study that the goal for a casework analysis was to determine if reproducible results could be produced for the mixture sample being tested, with reproducibility determined by multiple processing outputs other than the final likelihood ratio (LR). In other words, if all processing criteria are reproducible then the match statistic will also be reproducible. We did not need the source code to determine this.

11. I have reviewed the Declaration of Dan Crane in which he opines that "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). I disagree with that opinion because we tested all aspects of the TrueAllele program against known samples and known mixtures and found that TrueAllele produced the expected results. In the validation study referenced above we knew the DNA profile of every single contributor to every sample that we produced. We also predicted the approximate mixture weight/ratio of each contributor to every mixture, and on those occasions when TrueAllele calculated a different mixture weight, we re-examined the data and found that TrueAllele's calculation was representative of the data and that our predicted mixture weights were wrong.

12. In a validation study one can know what the correct contributor genotypes are and one can make an accurate estimate of what the mixture weight should be based upon the data. Our results made us more confident in TrueAllele because not only did the TrueAllele results correlate to most of our predictions, TrueAllele also identified for us those samples where the data showed us that our original predictions were incorrect.

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955

13. As an additional part of our validation study we assessed the specificity of TrueAllele by comparing the inferred mixture genotypes from the two- to five- contributor mixture samples referenced above to a random database of 10,000 single source profiles. In this specificity study we performed a total of 3.6 million comparisons. From these 3.6 million comparisons, TrueAllele identified 4,609 false positives, or matches in which the likelihood ratio was between 1 and 44,000 for samples from the random database that could not have contributed to the mixtures. Of these 4,609, only 14 had a LR of 1,000 or greater. Of these 14, 11 had resulted from highly complex five-contributor mixtures. The remaining 3 originated from complex four-contributor mixtures, with a maximum LR of approximately 1,200. We did not detect any false matches from two –contributor or three-contributor mixture samples.

14. This specificity study allowed us to determine a potential false positive region for TrueAllele analysis by which we can determine that for any four-contributor mixture, a likelihood ratio of 10,000 or less is in the range where a false positive might be expected, and therefore we report any sample with a likelihood ratio of less than 10,000 as statistically inconclusive. We consider this specificity calculator to be a valuable feature because we know that a LR in this range may be suspect. There may be other programs that allow for calculation of a false positive rate but I am not aware of them. We were able to derive all of these protocols without knowledge of the source code because we knew from these mixed samples exactly the results that TrueAllele should return.

15. For a casework mixture sample, I agree with the notion that a computer program cannot determine the "correct" answer because it is impossible to know what the "correct" mixture weights are for each component of a casework sample. Casework samples by

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955

definition are unknown: we cannot be 100% certain of the identity of the contributors, the number of contributors, the proportion of each contributor to the mixture, or any other conditions that may affect our ability to recover a true representation of the data. However, our laboratory does not presume that TrueAllele will show us the "correct" answer when we examine a casework sample. We do feel confident based on our training and our validation study that TrueAllele can provide us an accurate and reproducible prediction about the nature of the mixture, including the ability to identify potential contributors as well as to exclude non-contributors. At no point did we ever require the TrueAllele source code to make this assessment.

Under penalty of perjury under the laws of the State of Washington, I certify that the foregoing is true and correct to the best of my knowledge and belief.

Signed and dated by me this 2nd day of April, 2016, at Beaufort, South Carolina.

Name

Daniel T. Satterberg, Prosecuting Attorney
W554 King County Courthouse
516 Third Avenue
Seattle, Washington 98104
(206) 296-9000, FAX (206) 296-0955