IN THE INDIANA

COURT OF APPEALS

DARRYL PINKINS	)
	)
vs.	)
	)
STATE OF INDIANA	)

AFFIDAVIT AND AFFIRMATION OF MARK PERLIN

Mark W. Perlin, PhD, MD, PhD, being first duly sworn upon an oath, states as follows:

1. I founded Cybergenetics in 1994, and I am the Chief Executive Officer and Chief Scientific Officer for Cybergenetics. I created the TrueAllele Casework System, which is an automated short tandem repeat (STR) interpretation and analysis technology described in further detail herein. I hold eight patents in DNA mixtures and automated genotyping, and I am qualified to perform and interpret the DNA mixture, STR data in this case. The statements and information in this affidavit are based upon personal knowledge and professional experience in forensic DNA identification.

2. I hold advanced degrees in mathematics, medicine and computer science. I received my Bachelor of Arts in Chemistry in 1977 from the State University of New York at Binghamton. In 1982, I graduated from City University of New York Graduate Center with a PhD in Mathematics, and, in 1984, I received my medical degree from the University of Chicago, Pritzker School of Medicine. I graduated from Carnegie Mellon University in 1991 with a PhD in Computer Science. From 1986 to 1996, I was a Research Scientist and faculty member at Carnegie Mellon University where I directed projects in medical expert systems, artificial intelligence algorithms, and automated genome technology. For the past twenty years, from 1994 to the present day, I have worked as the Chief Scientific Officer and Chief Executive Officer for Cybergenetics, a company I founded in 1994. Cybergenetics manufactures the crime-solving TrueAllele® computer system for automated forensic DNA inference and match.

3. I have worked in research in many areas of medicine, science and technology. My early work was in MRI technology research, genomics, and cancer before moving to forensics. I have focused my research and technology innovation on automation and parallel processing efficiency. Together with colleagues, I constructed the first high resolution YAC contig map of Chromosome 11 in 1995 using our automated Inner Product Mapping technology. I was the first to develop a fully automated technology for interpreting and analyzing complex mixture DNA. I have collaborated with well-regarded crime laboratories in developing and testing technology for interpreting complex mixed crime scene DNA.

4. I have held academic positions as a research faculty member for ten years in the School of Computer Science at Carnegie Mellon University. I was also formerly adjunct faculty in the Department of Human Genetics at the University of Pittsburgh, and am currently adjunct faculty in the Forensic Science program at Duquesne University.

5. I am a member of the American Academy of Forensic Sciences, the American Society of Human Genetics, and the American Statistical Association professional organizations.

6. In 1973, I was awarded the Bausch and Lomb Science Award. I was inducted into Phi Beta Kappa in 1977, and in 1990, I was presented with the CUNY Alumnus Achievement Award. I served as a keynote speaker on five occasions. I gave the keynote speech at the SRI Biotechnology Conference in 1997, the Association for Computing Machinery Symposium on Software Reusability in 2001, the Duquesne University Summer Research Symposium in 2010, the International Conference on Forensic Research and Technology in 2012, and the International Conference on Forensic Inference and Statistics in 2014.

7. I hold eight patents related to DNA mixture analysis and automated genotyping. Patent numbers US5541067, US5580728, US5876933 and US6054268 pertain to patented processes that fully automate the analysis of STR genetic marker alleles. Patent number US6750011 is for a genotyping method. Patent number US6807490 pertains to a process for automatically analyzing nucleic acid samples. Lastly, patent numbers EP1229135 and US8898021 pertain to a process for analyzing mixtures of DNA. More specifically, these inventions are related to performing experiments that produce quantitative data, and then analyzing these data to characterize a DNA component.

8. I have regularly given nationwide and international presentations related to forensic DNA identification science and TrueAllele topics: M.W. Perlin, "Challenging DNA Evidence", Allegheny County Courthouse - Continuing Legal Education, Pittsburgh, PA, 27-Feb-2015; M.W. Perlin, "Solving cold cases by TrueAllele® analysis of DNA evidence", Finding Closure: The Science, Law and Politics of Cold Case Investigations, Cyril H. Wecht Institute of Forensic Science and Law 14th Annual Symposium, Pittsburgh, PA, 31-Oct-2014; J. Hornyak, W. Allan and M.W. Perlin, "Using TrueAllele® Casework to separate DNA mixtures of relatives", DNA Workshop, 124th California Association of Criminalists, San Francisco, CA, 20-Oct-2014; M.W. Perlin, "TrueAllele® Casework", Almost Everything You Wanted to Know About Probabilistic Software (But Were Afraid to Ask), Promega's Twenty Fifth International Symposium on Human Identification, Phoenix, AZ, 29-Sep-2014; M.W. Perlin, "TrueAllele® interpretation of DNA mixture evidence", Keynote talk, 9th International Conference on Forensic Inference and Statistics, Leiden University, The Netherlands, 20-Aug-2014; M.W. Perlin, "Solving crimes using MCMC to analyze previously unusable DNA evidence", American Statistical Association, Joint Statistical Meetings, Boston, MA, 3-Aug-2014; M.W. Perlin, "Computer interpretation of touch DNA mixtures", Seminar for Chiefs of Police in Western Pennsylvania, CSI Investigators Series, Pittsburgh, PA, 13-May-2014; M.W. Perlin, "Revolutionising DNA analysis in major crime investigations", The Investigator DNA Workshop, Aylesbury, Buckinghamshire, UK, 1-May-2014; M.W. Perlin, "Preventing rape in the military through effective DNA computing", Forensic Europe Expo, Olympia, London, 30-

Apr-2014; M.W. Perlin, "Cracking the DNA mixture code – computer analysis of UK crime cases", Forensic Europe Expo, Forensic Innovation Conference, Olympia, London, 29-Apr-2014; M.W. Perlin, "DNA-led investigation through computer interpretation of evidence", Pennsylvania State Police Training Seminar, Hershey, PA, 2-Apr-2014; M.W. Perlin and M.M. Legler, "Coding a safer society through computer interpretation of DNA evidence", MATLAB Virtual Conference, Europe and North America, 26-Mar-2014; M.W. Perlin, "TrueAllele® Computing: All the DNA, all the time", NSW Office of the Director of Public Prosecutions, Continuing Professional Development, Sydney, Australia, 19-Mar-2014; M.W. Perlin, "TrueAllele® interpretation of Allegheny County DNA mixtures", Allegheny County Courthouse, Pittsburgh, PA, 28-Feb-2014; M.W. Perlin, "Getting past first Bayes with DNA mixtures", American Academy of Forensic Sciences 66th Annual Meeting, Seattle, WA, 21-Feb-2014; M.W. Perlin, "No DNA Left Behind: When 'inconclusive' really means 'informative'", Schenectady County District Attorney's Office, Schenectady, NY, 31-Jan-2014; J. Butler, A. Mitchell, M.W. Perlin, A.M. Schubert, J. Friedman and J. Spriggs, "DNA mixture interpretations and statistics - to include or exclude", Fourth Annual Prescription for Criminal Justice Forensics, American Bar Association Criminal Justice Section, New York, NY, 7-Jun-2013; M.W. Perlin, "Death needs answers: the DNA evidence", The Cold-Blooded Murder of Dr. John Yelenic by Andrea Niapas, Christine Frechard Gallery, Pittsburgh, PA, 15-May-2013; M.W. Perlin, "Unleashing forensic DNA through computer intelligence", Forensic Europe Expo, Forensic Innovation Conference, London, United Kingdom, 24-Apr-2013; M.W. Perlin, "Revolutionising DNA Analysis in Major Crime Investigations", Workshop for Investigators, The Investigator Conferences, Rothley, Leicestershire, UK, 23-Apr-2013; M.W. Perlin, "Finding truth in DNA mixture evidence", Innocence Network Conference, Advanced DNA, Charlotte, NC, 20-Apr-2013; M.W. Perlin, "DNA mixture statistics", Virginia Spring Institute, Commonwealth's Attorneys' Services Council, Richmond, VA, 26-Mar-2013; G. Hampikian, V. Weedn, M.W. Perlin, A. Blumstein, J. Rangos, K. Mains, L. Irwin, A. Adepoju and W. Oliver, "Whose DNA is it anyway?", Duquesne University Forensic Fridays, Continuing Legal Education Program on DNA Access. Pittsburgh, PA, 15-Mar-2013; M.W. Perlin, K. Dormer, J. Hornyak, L. Schiermeier-Wood, and S. Greenspoon, "Virginia TrueAllele® validation study:

Casework comparison", American Academy of Forensic Sciences 65th Annual Meeting, Washington, DC, 22-Feb-2013; M.W. Perlin, K. Dormer, J. Hornyak, T. Meyers, and W. Lorenz, "How inclusion interpretation of DNA mixture evidence reduces identification information", American Academy of Forensic Sciences 65th Annual Meeting, Washington, DC, 22-Feb-2013; M.W. Perlin, "DNA mapping the crime scene: do computers dream of electric peaks?", Promega's Twenty Third International Symposium on Human Identification, Nashville, TN, 2012.

9. My work on the TrueAllele Casework system and genotyping has been published in respected peer-reviewed journals such as the *Journal of Forensic Sciences, American Journal of Human Genetics, PLoS ONE, Discrete Applied Math,* and *Genome Research*. I have authored about thirty scientific papers: Perlin, M.W., Hornyak, J., Sugimoto, G., and Miller, K. TrueAllele genotype identification on DNA mixtures containing up to five unknown contributors. Journal of Forensic Sciences, 2015 (in press); Perlin, M.W., Dormer, K., Hornyak, J., Schiermeier-Wood, L., and Greenspoon, S. TrueAllele Casework on Virginia DNA mixture evidence: computer and manual interpretation in 72 reported criminal cases. PLoS ONE, 9(3):e92837, 2014; Perlin, M.W., Belrose, J.L., and Duceman, B.W. New York State TrueAllele® Casework validation study. Journal of Forensic Sciences, 58(6):1458-1466, 2013; Ballantyne, J., Hanson, E.K., and Perlin, M.W. DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: combining quantitative data for greater identification information. Science & Justice, 53(2):103-114, 2013; Perlin, M.W. When good DNA goes bad. Journal of Forensic Research. S11:003, DOI 10.4172/2157-7145.S11-003, 2013; Perlin, M.W., Legler, M.M., Spencer, C.E., Smith, J.L., Allan, W.P., Belrose, J.L., and Duceman, B.W. Validating TrueAllele® DNA mixture interpretation. Journal of Forensic Sciences, 56(6):1430-47, 2011; Perlin, M.W. and Sinelnikov, A. An information gap in DNA evidence interpretation. PLoS ONE, 4(12):e8327, 2009; Perlin, M.W., Kadane, J.B., and Cotton, R.W. Match likelihood ratio for uncertain genotypes. Law, Probability and Risk, 8(3):289-302, 2009; Hill, S.Y., Shen, S., Zezza, N., Hoffman, E.K., Perlin, M., Allan, W. A genome-wide search for alcoholism susceptibility genes. American Journal

of Medical Genetics Part B (Neuropsychiatric Genetics), 128B: 102-113, 2004; Kadash, K., Kozlowski, B.E., Biega, L.A., and Duceman, B.W. Validation study of the TrueAllele® automated data review system. Journal of Forensic Sciences, 49(4):1-8, 2004; Perlin, M.W. and Szabady, B. Determining sequence length or content in zero, one, and two dimensions. Human Mutation,19(4), 2002; Perlin, M.W. and Szabady, B. Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. Journal of Forensic Sciences, 46(6), pp. 1372-77, 2001; Pálsson, B., Pálsson, F., Perlin, M., Gubjartsson, H., Stefánsson, K., and Gulcher, J. Using quality measures to facilitate allele calling in high-throughput genotyping. Genome Research, 9(10):1002-1012, 1999; Lancia, G., and Perlin, M. Genotyping of pooled microsatellite markers by combinatorial optimization techniques. Discrete Applied Math, 88(1-3):291-314, 1998; Perlin, M.W., Lancia, G., and Ng, S.-K. Toward fully automated genotyping: genotyping microsatellite markers by deconvolution. American Journal of Human Genetics, 57(5):1199-1210, 1995; Perlin, M.W., Burks, M.B., Hoop, R.C., and Hoffman, E.P. Toward fully automated genotyping: allele assignment, pedigree construction, phase determination, and recombination detection in Duchenne muscular dystrophy. American Journal of Human Genetics, 55(4):777-787, 1994.

10. I have written several book chapters on DNA identification and analysis science. These include: Perlin, M.W. The Blairsville Slaying and the Dawn of DNA Computing, In Death Needs Answers: The Cold-Blooded Murder of Dr. John Yelenic, A. Niapas, Ed., New Kensington, PA: Grelin Press, 2013; Perlin, M.W. DNA Identification Science. In Forensic Sciences, vol. 3, C.H. Wecht, Ed. Albany, NY: LexisNexis Matthew Bender; Chapter 37C, 2012; Perlin, M.W. Identifying human remains using TrueAllele® technology. Forensic Investigation and Management of Mass Disasters. M. I. Okoye and C. H. Wecht. Tucson, AZ, Lawyers & Judges Publishing Co: 31-38, 2007; Perlin, M.W. Mass casualty identification through DNA analysis: overview, problems and pitfalls. Forensic Investigation and Management of Mass Disasters. M. I. Okoye and C. H. Wecht. Tucson, AZ, Lawyers & Judges Publishing Co: 23-30, 2007.

11. In 1999, I started work on DNA mixture problems where two or more people contributed their DNA to a sample. I developed the TrueAllele Casework system to interpret these types of evidence mixtures and other less certain evidence. STR testing is a Polymerase Chain Reaction (PCR)-based methodology where the DNA of a biological sample is amplified through PCR. The DNA alleles that are amplified are STR loci that range in repeat length. The biological samples whose STR regions have been amplified are separated on a DNA sequencer to measure their STR length. Once the data is generated for these specifically amplified DNA markers, the mixture's data is processed through the TrueAllele system. TrueAllele uses computers and mathematics to analyze the STR DNA data, rather than human review. In other words, TrueAllele interprets the data by probabilistic genotyping rather than human interpretation. It should be noted that the data collection and signals are not different; the only difference between a human analyst and TrueAllele is the interpretation of the data. The computer looks at the data more thoroughly than human review, and the interpretation determines the nature of the genetic contributors that are present in the data. Thresholds used in human review are not used with a TrueAllele computer. Rather, all the data is examined statistically. The TrueAllele System considers approximately one hundred variables, but the exact number depends on the amount of data that is present.

12. I have testified in over twenty criminal trials and hearings, and I have worked on over two hundred cases and have filed as many reports, including one in Indiana. I have testified in criminal cases in front of the California, New York, Pennsylvania and Virginia courts, in Federal Court, and in a US Marine Corps court martial. I have also testified internationally in the United Kingdom and Australia. I am a qualified expert in both DNA evidence interpretation and the likelihood ratio. Additionally, I have testified in cases that have mixture evidence analyzed using the TrueAllele system.

13. The TrueAllele System uses a computer to assess the evidence objectively, without advance knowledge of the comparison reference. The computer assesses the data to separate the genotypes of each contributor to a mixture, and afterwards makes comparisons with appropriate references. TrueAllele

is based on two established scientific principles, Bayesian modeling and Markov chain Monte Carlo (MCMC). The computer separates the data into the genotypes of each contributor at every genetic locus, representing genotype uncertainty with probability (of allele pair possibilities). TrueAllele is able to determine error rates under different conditions for false positives and negatives, regardless of whether there were multiple contributors or whether the DNA template quantity was high or low.

14. TrueAllele has been validated and there are seven published peer reviewed validation papers on the TrueAllele Casework system. I began validating this system from its inception. In one validation paper, the results show that the computer is more sensitive in detecting lower quantities of DNA whereas human review essentially stops working at around one hundred picograms of DNA, which is just the beginning of a low template region. Unlike human review, TrueAllele computer interpretation extends all the way through low template range. In 2002, the New York State Police (NYSP) submitted the first NDIS validation for expert systems in the United States for TrueAllele Databank. The NYSP later published this seminal scientific validation in the Journal of Forensic Sciences, where it became the foundation for the Federal expert system validation standards. In September 2006, Cybergenetics received a contract to reanalyze and reinterpret the victim remains DNA data from the World Trade Center terrorist attack using its TrueAllele Casework system. TrueAllele enables fast, accurate and objective interpretation of complex DNA crime scene evidence by preserving all the information inherent in the DNA. A recent validation study with the NYSP has shown that TrueAllele Casework can provide information for almost all crime scene evidence items when DNA is present, compared with human interpretation that fails to get match information for two thirds of case items.

15. I received a grant from the National Institute of Justice to test the system on data generated from crime laboratories and my own laboratory, including samples of known composition and from casework. With a sample of known composition the genotypes and proportions are known, while casework samples come from crime scenes; both sample types are used in validations. TrueAllele is in use in Kern County,

California where the analysts report cases using the system. Virginia's Department of Forensic Science has trained TrueAllele analysts who conduct their own validations studies, and use the system to solve casework mixtures and prepare reports. Forensic scientists in Richland County, South Carolina use TrueAllele for their mixture analysis. There have been over twenty studies done on the system's reliability. TrueAllele has gone through twenty-five versions, refining the probability model over a ten year period. This testing has allowed the system to give appropriate answers on large sets of test data. These and other validation studies determine the system's specificity, sensitivity and reproducibility, which establish error rates.

16. In 2009, TrueAllele Casework results were presented in Indiana County court in the murder case Commonwealth of Pennsylvania v. Kevin Foley. This was the first time that automated computer interpretation was accepted into court as evidence, with the science supporting the resulting conviction of the defendant. Foley's conviction was affirmed by the state's appellate court in February of 2012.

17. At least six admissibility hearings have supported the reliability and admissibility of the TrueAllele analysis. In 2014, TrueAllele based evidence was ruled admissible in Louisiana and Ohio trial courts after substantial, but unsuccessful, Daubert challenges from the defense. In 2015, TrueAllele was admitted after Frye challenge in New York. Other unsuccessful TrueAllele Frye challenges were in California, Pennsylvania and Virginia.

18. California, Louisiana, Maryland, Massachusetts, New York, Pennsylvania, South Carolina and Virginia crime laboratories have purchased the TrueAllele Casework system. TrueAllele is also being used internationally in Oman, Australia, England and Northern Ireland.

19. Ten labs have purchased TrueAllele and four are currently using the system. TrueAllele began giving presentations in the State of New York in 2010, and gained approval from the DNA subcommittee

of the New York State Commission on Forensic Science in 2011 for its use in forensic casework. I have described the steps necessary for forensic casework approval in New York. When a lab purchases TrueAllele, it performs an internal validation on the equipment in accordance with the FBI quality assurance standards; the lab should measure how well the system works under a variety of different mixtures. Other people have conducted independent studies on the TrueAllele System. The forensic science community is moving toward using statistical computing in analyzing DNA mixtures to produce match statistics.

20. A study from the United States Department of Commerce National Institute for Standards and Technology ("NIST") on stochastic thresholds indicated that probabilistic genotyping is a more preferred approach to DNA interpretation. NIST has its own in-house TrueAllele computer. They have used it to characterize their standard reference mixtures materials for the forensic community. Through the use of TrueAllele, NIST knows the composition of the mixture standards that they send to other labs. NIST conducted its own independent study concerning TrueAllele. A forensic scientist at NIST indicated in a presentation that the community is moving forward with probabilistic genotyping. Threshold methods for DNA interpretation are an accepted practice amongst crime labs, but are viewed as antiquated by the community of scientists who develop statistical methods for interpreting forensic DNA evidence. The FBI's Scientific Working Group on DNA Analysis Methods (hereafter "SWGDAM") recognized probabilistic genotyping in 2010, and for the past year NIST has been promoting the approach.

21. Regarding general acceptance, TrueAllele was used in mass casualty identification of victims through DNA analysis. The TrueAllele System was used for the identification of victims in the 9/11 World Trade Center disaster. In addition, probabilistic genotyping and use of computers in interpreting DNA mixtures is a topic at conferences and a subject discussed amongst scientists. TrueAllele has also been involved in over ten defense cases, about half involving innocence project cases. Defense attorneys have written articles about TrueAllele. It is my understanding that the FBI and DNA testing laboratories throughout

the country will be moving toward some sort of probabilistic genotyping system. A scientist can duplicate my mixture analysis by properly using the TrueAllele system. Replicating TrueAllele results is advisable because of statistical computing considerations.

22. The TrueAllele System's computer source code is a trade secret. The source code contains about 170 thousand lines. The reliability of a software program is determined by testing and validation studies, and not by looking at the source code. The validity of software is assessed by how the program operates, not by reading the text. About half a dozen other systems exist that are similar to TrueAllele. Some are open source, while others are closed-source; commercial closed-source systems are validated by the developer. I have interpreted the STR data for this case, and I used the TrueAllele Casework System to determine the contributors' genetic profiles.

23. I have interpreted the existing STR data for this case using the TrueAllele Casework system. The existing data addressed in a Cellmark report (Case F901011) show DNA mixtures from a jacket (Items 2A, 3A, 4A) and a sweater (Item 12A). Reference genotypes for comparison were available from *Darryl Pinkins* (Item 7AP), *Roosevelt Glenn* (Case F06-0055), and questioned *hair* on a slide (Case FOR4863, Item 59D).

24. TrueAllele was able to statistically separate the quantitative STR mixture data from the evidence items to accurately determine genotypes.

25. TrueAllele inferred a 90% major contributor from the nonsperm fraction (NSF) of the Jacket Cutting #3-3 (Item 03AE1P) mixture to determine a possible *victim genotype*. This is a definite genotype, as if from a single DNA source. This definite genotype had been previously reported.

26. TrueAllele inferred a 90% major contributor from the sperm fraction (SF) of the Jacket Cutting #3-3 (Item 03AE2P) mixture to determine a *major jacket genotype*. This is a definite genotype, as if from a single DNA source. This definite genotype had been previously reported.

27. TrueAllele inferred an 80% major contributor from the sperm fraction (SF) of the Sweater Cutting (Item 12AE2P) mixture to determine a *major sweater genotype*. This is a definite genotype, as if from a single DNA source. This definite genotype had been previously reported.

28. TrueAllele inferred a 10% minor contributor from joint consideration of the sperm fraction (SF) of the Jacket Cutting #3-2 (Item 02AE2P) mixture, together with the sperm fraction (SF) of the Sweater Cutting (Item 12AE2P) mixture, to determine a *minor jacket-sweater genotype*. The major jacket and major sweater genotypes were used as assumed known genotypes in this calculation. This probabilistic minor genotype was not previously reported.

29. TrueAllele inferred a 5% minor contributor from joint consideration of the sperm fraction (SF) of the Jacket Cutting #3-3 (Item 03AE2P) mixture, together with the sperm fraction (SF) of the Jacket Cutting #3-4 (Item 04AE2P) mixture, to determine a *minor jacket-jacket genotype*. The major jacket genotype was used as an assumed known genotype in this calculation. This probabilistic minor genotype was not previously reported. It is more likely than not that this is a real genotype, and not a computer artifact.

30. There is no statistical support for a match between any of the five mixture genotypes in the preceding paragraphs and Darryl Pinkins' genotype. Thus it is my expert opinion that Darryl Pinkins' genetic DNA profile is excluded as having contributed to any of the mixture evidence in this case.

31. There is no statistical support for a match between any of these five mixture genotypes and Roosevelt Glenn's genotype. Thus it is my expert opinion that Roosevelt Glenn's genetic DNA profile is excluded as having contributed to any of the mixture evidence in this case.

32. The major jacket, major sweater, and hair genotypes are all different from each other. However, these three distinct genotypes share alleles at many genetic loci, and exhibit some statistical association. This genotype similarity and statistical association may be due to some relatedness of the three individuals.

33. Except as noted in the preceding paragraph, there is no statistical support for a match between any of the five mixture genotypes (victim, major jacket, major sweater, minor jacket-sweater, minor jacket-jacket), nor between any of these genotypes and those of Darryl Pinkins, Roosevelt Glenn, and the hair genotype.

34. It is my expert opinion that at least three (3) unknown individuals contributed their DNA to the mixture evidence, and that these distinct individuals are people other than Darryl Pinkins, Roosevelt Glenn, the victim, and the person who left the hair known as Exhibit 59D.

A F F I R M A T I O N

I, Mark Perlin, PhD, MD, PhD, Chief Executive Officer and Chief Scientific Officer for Cybergenetics, do hereby affirm, this 22nd day of April, 2015, under penalties for perjury, that the foregoing representations are true to the best of my knowledge and belief.

Dr. Mark W. Perlin
Chief Scientific Officer, Cybergenetics