Victoria State Coroner's Inquest into the Death of Jaidyn Leskie
Report prepared by Dan E. Krane
December 4, 2003

This report has been prepared at the request of Ms. Rowena Orr and Mr. James Kennan.

I am an Associate Professor in the Department of Biological Sciences at Wright State University in Dayton, Ohio, USA. I have a B.S. degree with a double major in Biology and Chemistry from John Carroll University (Cleveland, Ohio), and a Ph. D. from the Biochemistry program of the Cell and Molecular Biology Department of the Pennsylvania State University (State College, Pennsylvania). I have also done postdoctoral research using the tools of molecular biology to answer questions in the fields of population genetics and molecular evolution in the Genetics Department of the Washington University Medical School (St. Louis, Missouri) and in the Department of Organismic and Evolutionary Biology of Harvard University (Cambridge, Massachusetts). I have published more than 30 scholarly papers in a variety of topics including population genetic studies of the genetic diversity of human populations at DNA typing loci, of organisms exposed to environmental stressors, and the use of DNA typing in forensic science. I am also the lead author of a widely used textbook, *Fundamental Concepts of Bioinformatics.* And, I am the founder and president of Forensic Bioinformatic Services, Inc., a consulting company that since its establishment in April of 2002 has reviewed the DNA evidence in hundreds of criminal cases where STR-DNA profiling has been performed. Since 1991 I have testified in over 45 criminal cases (from at least 12 of the 50 United States of America and in US Federal court) that have involved forensic DNA typing.

I have reviewed materials supplied to me Ms. Orr and Mr. Kennan including reports from: John Scheffer (dated 30 July, 2003); Robert Goetz (dated 20 September, 2003); SallyAnn Harbison (dated 26 November, 2003); Barry Boettcher (dated 20 October, 2003 and 13 November, 2003); and Bruce S. Weir (dated 10 November, 2003 and 24 November, 2003) as well as the working notes and other relevant materials related to the Victoria State Police Forensic Service Centre (VPFSC) cases 6603/967 and 2831/978. I have also communicated directly with Professor William C. Thompson prior to the completion of a report he intends to submit regarding the findings in these cases.

Like the other experts who have reviewed the relevant case materials and have prepared the reports mentioned above, I agree that the DNA profiles associated with the

condom (samples “1 i” and “1 ii” from case 2831/978) and from the bib and track-pants (samples “70 iii a” and “70 vi a” from case 6603/967) match each other across all tested STR loci for which results are available (no results were obtained for two loci during the testing on the bib and no results were available for four loci during the testing on the pants). The alleles observed in the tested loci for these four samples also correspond to alleles observed in the reference standard from the complainant, “Ms. P.,” in the rape investigation associated with case 2831/978.

As with all forensic DNA analyses, there are only three possible explanations for a match between a reference and an evidentiary sample:

1) the biological material associated with the evidentiary sample in fact originated from the person who provided the reference sample;

2) the true source of evidentiary sample is not the person who provided the reference sample but the two individuals do coincidentally match at all loci that have been tested;

3) an error has occurred (either accidental or deliberate) in the handling/collection, testing or interpretation of the evidentiary and/or reference sample.

The first alternative: the same person contributed both samples

From Mr. Sheffer’s 30 July, 2003 briefing paper I understand that Detective Senior Sergeant Roland Legg has concluded that Ms. P. has no connection to the death of Jaidyn Leskie (case 6603/967). I have not been provided a copy of Detective Senior Sergeant Legg’s report. However, if the first of the three alternative explanations listed above is effectively eliminated from consideration, only two logically possible explanations remain as to how Ms. P’s DNA profile has been found to be associated with items believed to have been worn by Jaidyn Leskie at the time of his disappearance.

The second alternative: Coincidental matching

Random match probabilities for related and unrelated individuals:

The likelihood of the second possible explanation, a coincidental match between the perpetrator of a crime and another individual chosen randomly from a population, can be rigorously addressed with the tools and approaches employed in the discipline of population genetics. In the Leskie investigation, DNA typing information was obtained from two different sets of loci – those examined by a commercially available test kit

known as Profiler Plus and those examined by a different kit known as Green I. Only seven of the nine Profiler Plus loci yielded reliable DNA profile information for the bib (and only five of the nine Profiler Plus loci yielded reliable DNA profile information for the track pants) and Ms. P was only typed with the Profiler Plus kit. I have used an approach recommended by the second United States National Research Council report on forensic DNA profiling (and the allele frequencies for Caucasian Australians) to determine the chance of a coincidental match between Ms. P and an unrelated Caucasian Australian at the seven Profiler Plus loci where she shares alleles common to those associated with the Leskie bib. Those calculations suggest that fewer than one in 1.3 billion unrelated Caucasian Australians would similarly match at those seven loci.

Individuals related to Ms. P are more likely to share alleles with her than unrelated Caucasian Australians. Specifically, the chance of a coincidental match across the same seven polymorphic loci with any given: full sibling is one in 1,300; parent is one in 1.2 million; half-sibling, aunt or uncle is one in 150 million; and cousin is one in more than 19 billion. If Ms. P has 16 half-siblings (as Ms. Orr has represented to me), the chance that one of those 16 half-siblings would have the same DNA profile across those seven polymorphic loci is 16/150 million or, approximately one in 9.3 million.

Barring the existence of an unknown full-sibling of Ms. P, and accepting that the assumptions and databases used by VPFSC are correct and reliable, the chance of a coincidental match between her and a randomly selected related or unrelated Caucasian Australian can be reasonably removed from consideration.

Cold hit probabilities:

It should be pointed out that the concordance between the DNA profiles associated with the bib and track pants from the Jaidyn Leskie investigation and Ms. P was not detected as the result of choosing Ms. P strictly at random from the Caucasian Australian population. Instead, the match is more properly described as having been found as a "cold hit" that resulted from a scan (or trawl) of a database of DNA profiles generated during the course of other investigations in Australia. The relevant question for such a cold hit is more appropriately stated as: "What is the chance that one of the more than 19,000 individuals in the database at the time of the trawl would perfectly match the DNA profile associated with Jaidyn Leskie's bib across these seven polymorphic loci?" rather than "What is the chance that an unrelated Caucasian Australian might perfectly match the DNA profile associated with Jaidyn Leskie's track pants across these seven polymorphic loci?" Since the primary difference between these kinds of matches is the manner in which a suspect is first identified, it is generally accepted that it is not possible to convert one type of case into the other (for instance, by

simply retesting a reference sample once a “cold hit” has been identified). It is also generally accepted in the scientific community that the statistical significance of those two kinds of DNA profile matches should be determined differently. However, there are at least three different commonly held opinions on how the statistics associated with “cold hits” should be generated and presented.

The first group to formally address this issue was a body of experts appointed to the Committee on DNA Science by the United State’s National Research Council in 1992. The position of this group is that database searches should be used to identify potential suspects but not to calculate frequency estimates. When successful, suspects identified by these searches would then be tested at a completely different group of independent genetic markers that would also be compared to the evidence. If these additional genetic loci also match between the suspect and evidence sample, they alone would be used to compute probabilities that reflect the significance of a match. With this methodology the genetic markers used in the original database search are specifically and deliberately excluded from any statistical calculation.

A second committee of prominent experts advocated a significantly different approach in 1996. They specifically recommended that, “When the suspect is found by a search of DNA databases, the random match probability should be multiplied by N, the number of persons in the database.” (The Evaluation of Forensic DNA Evidence, 1996, National Research Council Press, p. 40, 161). Proponents of this approach feel that the first method is too conservative. Their alternative method differs in three ways: 1) no testing is performed at additional loci; 2) genetic markers used in the original database search are included in the statistical calculations; and 3) the size of the database being searched (N) is taken into consideration.

A third group is comprised of individual scientists who have published peer-reviewed manuscripts in which they argue that a “cold hit” should actually be given more weight than a match found in a “probable cause” case. Their position is based on the thinking that not only has the defendant been found to match the evidence, but many more individuals have been found to not match. In “probable cause” cases where only a single match is found during the course of DNA testing, there is at least still a formal possibility that one or more untested people may also match the evidence –that possibility becomes increasingly less likely as the database used for a cold hit becomes larger. Proponents of this approach also feel that the first method is too conservative. Their method differs from it in three ways: 1) no testing is performed at additional loci; 2) genetic markers used in the original database search are included in the statistical calculations; and 3) the size of the database being searched (N) is taken into consideration. It also differs from the second in one very important way: the effect of the

database size on the significance of a match is precisely opposite – large databases generate the most damning statistics for a defendant while, in the second approach, the larger the database the less damning the statistics become to a defendant. The second and third approaches are diametrically opposed with respect to implications of the size of the database that is searched.

The proponents of each of these three approaches include many eminent scholars in the field of genetics and statistics. For instance, the blue ribbon panel of experts that generated the first National Research Council on DNA typing report (which supports the first approach as described in paragraph 4 above) includes Drs. Mary-Claire King, Richard Lempert, Eric Lander, Ruth Macklin, Thomass Marr, Victor McKusick, Philip Reilly and Sandy Zabel. Members of the second National Research Council on DNA Typing (which recommends the second approach as described above) include prominent population geneticists and statisticians such as Drs. James Crow, Arno Motulsky, Thomas Nagylaki, Mashotoshi Nei, David Siegmond and Stephen Stigler. The third approach (described in paragraph 6 above) is one that has been principally advocated by very influential and often cited geneticists and statisticians such as Drs. David Balding, Peter Donnelly and Bruce Weir (as in publications such as: Errors and Misunderstandings in the Second NRC Report, D. J. Balding, Jurimetrics, Summer 1997, 37:469-476; Evaluating DNA Profile Evidence When the Suspect s Identified through a Database Search, D. J. Balding and P. Donnelly, Journal of Forensic Science, 1996, 41:603-607; and Interpreting DNA Evidence, I. W. Evett and B. S. Weir, Sinauer Press, 1998, pp. 219- 222). This appears to represent a genuine split between three fundamentally different approaches by experts who are significant both in number and in eminence within their fields.

While I acknowledge that rigorous and compelling arguments have been presented by all three groups, I am personally in agreement with the position of the first group. Given that Ms. P was identified as matching the Leskie bib and track pants, it is possible to use the information from the three different loci that the Green I test kit examines to generate a statistic as this first group suggested. In fact, DNA profile associated with the Leskie bib and track pants as well as the inside of the condom from Ms. P's rape investigation do also match at those three Green I loci. The chance that two unrelated Caucasians (using allele frequencies from the US Caucasian population) would coincidentally match at those three loci is less than one in 1,200. (Both the second and the third proposed solutions to this question would generate much more unlikely probabilities that would effectively eliminate from consideration the chance of a coincidental match between Ms. P and a randomly selected related or unrelated Caucasian Australian.) In my opinion, this is the only approach that can consistently

generate conservative statistics that still reflect the power of DNA typing methodologies. My own analyses of the population databases generated by crime laboratories within the United States suggests that it may not actually be possible to resolve the conflict between the other two groups due to subtle and pervasive population substructuring within a general population (let alone within the subpopulation of those that are DNA profiled as part of a criminal investigation). An abundant number of genetic markers that are currently available allow the creation of a useful database consisting of one set of genetic markers *and* the subsequent generation of probative forensic evidence relying upon information from a different set of loci. In short, this approach allows an extremely conservative estimate of the significance of a cold hit match to be generated that still reflects the resolving power of DNA typing techniques. In those instances where greater resolving power might be needed (perhaps as in this case where an admittedly very conservative approach estimates that the chance of a coincidental match could be as likely as one in 1,200) an easily implemented solution would be to simply examine additional, previously untested loci.

The third alternative: the match is due to error

There are essentially three opportunities for errors arise in DNA profiling analyses: interpretation of testing results for the evidence and/or reference sample(s); handling/collection of evidence sample(s); and contamination with DNA amplified during the testing process itself.

Interpretation error:

As described at the beginning of this report, I, like the other experts who have reviewed the relevant case materials and have also prepared reports, agree that the DNA profiles associated with the condom (samples "1 i" and "1 ii" from case 2831/978) and from the bib and track-pants (samples "70 iii a" and "70 vi a" from case 6603/967) match each other across all tested loci for which results are available (no results were obtained for two loci, during the testing on the bib). Using the typing information as it appears in Table 1 of Mr. Sheffer's 30 July, 2003 report, the alleles observed at the tested loci for these four samples also correspond to the alleles observed in the reference standard from the complainant, Ms. P., in the rape investigation associated with case 2831/978. I have requested but have not received for review the electronic data that was generated during the course of the DNA testing process for these two cases. It is possible that a review of that data would allow me to detect, among other things, indications of low levels of signal associated with alleles that Mr. Sheffer does not list in Table 1 of his 30 July, 2003 report. A review of the control samples (a positive, negative and reagent blank) as well as the evidence samples for low level signals that would correspond to Ms. P's alleles in

particular is something that I would personally like to examine before definitively concluding that no interpretation errors (particularly regarding the possibility of contamination) occurred in either of the cases being considered here. However, I found no reason to believe that any interpretation errors have been made with regard to any of these samples in my review of the testing laboratory's notes that were generated during the test that they performed and I am prepared to assume that Mr. Sheffer's summations are also correct, in this regard.

Contamination during handling/collection of samples:

It is my understanding that the investigations into the rape of Ms. P and the abduction of Jaidyn Leskie were both physically and temporally separated such that samples (or those individuals handling them) could not have come into contact with each other prior to their arrival at the VPFSC.

However, the samples from the condom associated with the rape investigation (samples "1 i" and "1 ii" from case 2831/978) and the bib and track-pants (samples "70 iii a" and "70 vi a" from case 6603/967) were examined at the VPFSC within what can be considered a reasonable window of opportunity for contamination. (DNA was extracted either by or under supervision of the same senior forensic scientist and case manager/reporting officer, Mr. Max Jones, from the condom on Monday, 2 February, 1998 and from Jaidyn Leskie's clothing just two days later on Wednesday, 4 February, 1998. Contamination of Jaidyn Leskie's clothes by material from Ms. P's reference sample could not have occurred since that reference sample did not arrive in the lab until October of 1999.) Mr. Sheffer's 30 July, 2003 briefing paper describes in considerable detail the stringent safeguards that are in place to minimize the possibility of physical contact between evidence samples (let alone between evidence samples in different cases). I did not observe a record of any obvious errors or omissions relative to the VPFSC's documented procedures in my review of the laboratory notes generated during the course of the two relevant investigations (2831/978 and 6603/967).

Still, even if evidence samples from the two cases did not come into direct physical contact, secondary transfer (by way of common contact with a surface or implement) of biological material from the condom or articles of Ms. P's clothing in the rape investigation to the bib and track pants could have occurred. I agree with Dr. Harbison's assessment that no amount of verification of good intentions or examination of laboratory notes, methods and procedures can eliminate from consideration the possibility of such secondary transfer.

I understand but I am not persuaded by Mr. Sheffer's position that primary and/or secondary transfer is unlikely to have occurred for at least two reasons: 1) the Ms. P's DNA profile associated with the inside and outside of the condom is part of a mixture yet only Ms. P's alleles appear to be associated with the bib and track pants; and 2) no other samples including negative controls associated with both investigations display indications of Ms. P's allele's being present.

First, the DNA profile reported for the "inside condom" as described on the first table of Mr. Scheffer's report only barely qualifies as a "mixture" in that only Ms. P's alleles are observed at all but one of the seven Profiler Plus loci for which information is also available from the bib in the Leskie case. In that one exception (the D21 locus) the very much weaker signal associated with the single allele (a 28) that is not also observed in Ms. P (32.2 and 33.2 alleles) may have been lost due to "allelic dropout." (In some instances, an STR test will detect only one of the two alleles from a particular contributor at a particular locus. Generally this occurs when the quantity of DNA is relatively low, either because the sample is limited or because the DNA it contains is degraded, and hence the test is near its threshold of sensitivity. The potential for allelic dropout complicates the process of interpretation because analysts must decide whether a mismatch between two profiles reflects a true genetic difference or simply the failure of the test to detect all of the alleles in one of the samples. The occurrence of "allelic dropout" usually cannot be independently verified – the only evidence that this phenomenon has occurred is the "inconsistency" that it purports to explain.) As Dr. Harbison points out in her report, "it is entirely possible that the very low level minor component of the apparent mixture in condom sample (i), would not be detected in the DNA profile from the bib and track-pants that has about $1/10^{th}$ the intensity of the DNA profile from the condom."

Second, there is no good reason to expect that contamination would uniformly affect all samples if and when it does occur. In much the same way, *Salmonella*-tainted meat placed on one part of a countertop at the beginning of food preparation does not mean that all food subsequently prepared for the same (or even a later) meal will be similarly tainted – just those items that also come in contact with the same portion of the countertop can be reasonably expected to have a chance of being tainted.

Contamination with DNA amplified during testing:

It is also possible for the polymerase chain reaction (PCR) amplification process involved with contemporary DNA profiling to generate large amounts of material that can act as a source of contamination of other samples. The Profiler Plus amplification of the condom sample was performed on 23 November, 1999 and that the Profiler Plus

amplifications of the bib and track pants samples were first performed on 20 December, 2002. Since the testing on the condom precedes the testing of the bib and track pants this type of contamination remains a formal possibility. However, almost thirteen months seems to be an unreasonably large window of opportunity for contamination – especially given that no other samples tested in the laboratory during that time period seem to have been similarly affected.

Indications that error may have occurred:

The key factor that suggests that contamination of the bib and track pants samples occurred within the laboratory is the fact that the same personnel processed samples from the two cases within a plausible window of opportunity for contamination to have occurred. In an inquest such as this, it is particularly important to bear in mind that the presence of a DNA profile on an article is *not* itself evidence that sinister circumstances were associated with its transfer. In fact, it is quite uncommon for DNA tests themselves to say *anything* about the circumstances (or even the time frame) associated with the transfer. The DNA testing procedures that were employed in these investigations are exceptionally sensitive (literally being able to generate interpretable DNA profiles from as little material as is associated with a finger print). While that sensitivity represents a great strength of the methodology, it also constitutes a significant weakness in that it allows DNA profiles to be obtained from vanishingly small amounts of contaminating material.

Further, as pointed out by Dr. Harbison, it is somewhat surprising that it was even possible to obtain DNA testing results from the bib and track pants given that they had been submerged for what may have been months. These conditions are almost the antithesis of the circumstances that are generally held to be most conducive to preserving the integrity of a DNA sample. (The tested stains are described as being very weak and dilute in the contemporaneous VPFSC notes.) Also, if testing results could be obtained, it is even more surprising that none of Jaidyn Leskie's own alleles were also observed upon his own clothing.

Both of these additional surprising results could have been effectively addressed (and the test itself made more credible) if the testing laboratory had performed substrate controls at the same time that they had tested the putative stains on the bib and track pants. If similar results had been obtained from portions of the bib and track pants that did not appear to be stained it would have been a further indication that contamination may have occurred. Alternatively, if no typing results were obtained or if only alleles consistent with Jaidyn Leskie were observed, it would have supported the assertion that the profile that matches Ms. P was in fact associated with the stains.

Conclusions

A ten locus STR-DNA profile match such as the one observed between the condom of the rape investigation (VPFSC case 2831/978) and the bib in the Jaidyn Leskie abduction investigation (VPFSC case 6603/967) is exceedingly unlikely to be the result of coincidence. Population substructure (including the existence of related individuals) as well as the artifices of a database trawl do increase the likelihood of such a match. However, I firmly agree with Dr. Harbison that on the balance of probabilities, accidental contamination by DNA from what is described as the inside of the condom sample is the most likely cause of the DNA profiling results produced from the bib and track pants.

I have carefully reviewed all of the materials provided to me and I have not withheld any issues of significance to this inquest.

Sincerely,

Dan E. Krane
Associate Professor of Biological Sciences,
Wright State University,
Dayton, OH USA 45435-0001