

**IN THE CIRCUIT COURT OF COOK COUNTY
COUNTY DEPARTMENT - CRIMINAL DIVISION**

PEOPLE OF THE STATE OF ILLINOIS

v.

JUAN LUNA

:
:
:
:
:

No. 02 CR 15430

DECLARATION

I, BRUCE BUDOWLE, Ph.D, declare under penalty of perjury that the following statements are true and correct:

1. I am Dr. Bruce Budowle, Senior Scientist at the FBI Laboratory. My qualifications are summarized in my attached resume.

2. DNA databases, such as CODIS (Combined DNA Index System), are useful for developing investigative leads for a variety of crimes where there is no known suspect, for linking multiple cases committed by the same individual, and for identifying human remains attributed to a missing person. Such databases routinely assist law enforcement in solving crimes, and future crimes are prevented by more quickly identifying true perpetrators. In addition, searching a DNA profile derived from an evidence sample against a large repository of Convicted Felon DNA profiles in effect excludes an extremely large number of individuals as being the source of an evidence sample.

3. It has been conveyed to me that there is a desire to obtain the DNA profiles contained within the Convicted Felon database of the State of Illinois for the purposes of a search or, alternatively, for the State of Illinois to perform the search itself. The object of the proposed search, as it has been conveyed to me, is to locate profiles which match at nine or more loci. The likely purpose for obtaining the DNA profiles is to perform empirical studies to assess the validity of the current practices for generating random match probability estimates (as described in the NRC II Report and followed throughout the United States by DNA forensic science practitioners).

4. Because of their nature and design, United States Convicted Felon DNA databases, such as that in Illinois, are an inappropriate source of DNA data for assessing DNA profile rarity, and no valid analyses using such a repository can be carried out regarding the reliability of current statistical practices. Any results obtained from studies using such data would not be relevant and would be misleading for either supporting or refuting current forensic DNA statistical practices. Among the criteria that define the evaluation of the statistical legitimacy of using allele frequency estimates under current forensic practices are that the databases contain little or no duplicate profiles or profiles of close relatives. Therefore, before any legitimate

statistical frequency inferences could be drawn from such databases, it would be imperative to remove as many as possible of the duplicates or profiles contributed by close family members. It is widely known that the current felon DNA databases unavoidably contain duplicate profiles and profiles of relatives. Maintaining such profiles in a Convicted Felon database does not compromise its use for developing investigative leads. However, removal of such profiles would be a monumental task and such expenditure of resources would gain little insight into the intricacies of forensic statistics.

5. In the early 1990s there was much debate on the statistical practices used to assess the rarity of a DNA profile. One of the most notable criticisms raised by detractors was that the use of allele frequency data from heterogeneous databases is improper; it would violate the basic assumptions of allelic independence. In other words, major population groups, such as Caucasian and African American, are composed of subgroups, and the allele frequencies may be quite different among the subgroups within a major population group. These allele frequency differences would then cause the assumption that the alleles at a genetic marker are independent to be erroneous. Thus, multiplying allele frequencies to derive an estimate of the rarity of a DNA profile would be inappropriate. These criticisms have been shown to be of minimal concern when following current forensic practices of compiling broad population group data sets, such as African American and Caucasian. However, these concerns are directly applicable to the use of a Convicted Felon database to verify the assumptions of allelic independence, because the compiled profiles in the data set are not composed of profiles from only one major population group; instead the DNA profiles derive from individuals from different major population groups. A Convicted Felon database qualifies as one of the most heterogeneous DNA profile databases available. It is comprised of individuals from many different population groups (African American, Asian, Caucasian, Hispanic, Native American, and Oceanian). Anyone advocating the use of a Convicted Felon database for evaluating the validity of current forensic statistical practices would have to discard concerns about using a very heterogeneous database. Heterogeneity is an irresolvable problem because population affinity data are not maintained with the DNA profile data. This lack of information prevents the data being parsed into more homogeneous major population groups (i.e., the population partitions used by forensic scientists). Observed departures from expectations would have no relevance for questioning the reliability of statistical practices because the very heterogeneous data sets would be expected to violate the basic assumptions of independence.

6. Better defined population data, those partitioned into major population categories and which are sufficiently abundant, would provide a more meaningful assessment. Many studies on better defined population data sets already have addressed the validity of the basic assumptions of DNA forensic statistics and overwhelmingly support the current practices as reliable.

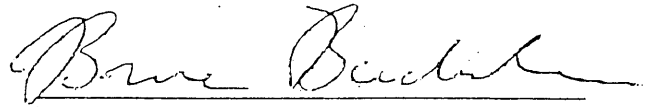
7. The strict product rule (i.e., the assumption that the alleles are statistically independent) is never applied by United States forensic practitioners. The recommendations of the NRC II Report are followed which require conservative correction factors because the assumption of independence may not be met. The current practices assume there are violations of the

assumption of independence. Thus, tests demonstrating a violation of independence add little to question the current practices.

8. Likely pairwise profile comparisons will be used to generate empirical data from the collection of Convicted Felon profiles. When conducting pairwise DNA profile comparisons using Convicted Felon database DNA profiles, it is important to recognize that the number of profile matches that might be found at 9 or more loci is entirely predictable, is directly related to the basic principles used to generate a random match probability estimate, and is grounded in well-established probability theory. The principle is the same as that for the well-known Birthday Scenario. If asked "assuming the probability of having a birthday is independent and all birthdays are equally distributed, how many people would have to be gathered in a room for odds to be better than even that two people share the same birthday?" most people answer this question with numbers of 180 or more. In fact, only 23 people need to be gathered for odds to be better than even that two will share the same birthday. This result may seem odd and counterintuitive; however, one must appreciate the question being asked. The question is not "what is the chance of a match between a specific person(s) and a specific birthday?" Instead, the question is "what is the chance that any two people might share any birthday, with the birthday not being specified." The latter is different in that it is not based solely on the number of people, but on the number of pairwise comparisons. When seeking a match of a specific birthday in a group of 23 people, there are only 23 comparisons made. For any two to share a birthday out of 23 people, there are 253 total pairwise comparisons. The same principle applies to "matches" and "partial matches" in the Convicted Felon database, even though it may seem counterintuitive with profile frequencies estimated to be less than one in a trillion. The issue again resides in the number of comparisons. The total number of pairwise comparisons, for example, for the CODIS database with approximately 3,000,000 profiles would be more than 4 trillion. Databases with only 100,000 samples would enable more than a few billion pairwise comparisons. Thus, with so many pairwise comparisons, matches at 9 or more loci are expected. Observing such partial matches would not call into question the current forensic practices. Indeed, it would support the current statistical practices.

9. Convicted Felon databases are excellent tools for providing investigative leads. However, they are extremely poor for inferences regarding the assumptions of current forensic statistical practices. The current statistical practices are well-established and grounded on many population studies readily available in the scientific literature. The DNA profiles in Convicted Felon databases are not good for quality population statistics studies, because they are not properly annotated for population affinity, are not parsed out into major population categories, and duplicates and relatives reside in the database. Any results obtained from population genetic and forensic statistics validity studies using the data in its current form would be meaningless.

I declare under penalty of perjury pursuant to 28 U.S.C. Section 1746 that the foregoing is true and correct and that those matters stated upon information and belief are true to the best of my knowledge.



Bruce Budowle

Subscribed and sworn to me on this 17th day of July, 2006.

County/City of Stafford
Commonwealth/State of Virginia
The foregoing instrument was acknowledged
before me this 17th day of July
2006 by
Bruce Budowle
(name of person seeking acknowledgment)
Sharon Gaye Cousin
Notary Public
My commission expires: May 31, 2008