Source Identification of Body Fluid Stains Using DNA Profiling

Jenifer A.L. Smith, Ph.D., and Bruce Budowle, Ph.D.
FBI Laboratory, 935 Pennsylvania Avenue, N.W., Washington D.C. 20535

In his 1995 commentary, published following the infamous OJ. Simpson trial, Dr. Bruce Weir stated "I look forward to the time when DNA profiles are of such extent and integrity that they are recognized as being as probative as fingerprints. Along with this recognition will come the presentation of matching profiles, in most situations, without numerical statements." (1). The time for source attribution has come.

In October of 1997, the DNA Analysis Unit I (DNAU I) of the FBI Laboratory implemented a policy that enabled examiners to report and offer an opinion regarding the source of an evidentiary body fluid stain. This policy was developed and implemented based on the recommendations of the report authored by National Research Council of the National Academy of Sciences (2). Such a conclusion can be reached if a match between a known and questioned sample is determined and if the frequency of occurrence of the evidence DNA profile is exceedingly rare. Analysis using either or both RFLP and PCR-based loci may provide sufficiently rare profiles to opine source attribution.

The report, entitled The Evaluation of Forensic DNA Evidence, (NRC II) (2) endorses the methods of DNA typing and statistical interpretation currently used by the forensic community within the United States as stated below:

"The technology for DNA profiling and the methods for estimating frequencies and related statistics have progressed to the point where the reliability and validity of properly collected and analyzed DNA data should not be in doubt."

The report also addressed the issue of uniqueness. The committee suggested that if sufficient information was obtained by DNA typing, the source of the DNA obtained from the stain may be uniquely identified. Attribution of the source of an evidentiary sample does not require uniqueness. Indeed, uniqueness (excluding identical twins) cannot be determined unless all members of a population(s) are typed. However, if a large number of loci are typed, the DNA profile from the evidence can be so rare that it is likely that a suspect with a matching profile is the source of the evidentiary stain.

To ensure that a high degree of confidence has been attained regarding source attribution, a threshold probability value (p) must be established. Following release of the NRC II Report, an approach was developed to determine a threshold value for examinations conducted in the DNAU I. The details of this approach are generally detailed below and are more fully described in Budowle *et al.* (3, in preparation).

An individual (excluding identical twins) can be identified as the source of an evidentiary profile to a reasonable degree of scientific certainty in the United States, if the adjusted probability (p) of the DNA profile(s) from the questioned specimen(s) satisfies the following condition:

$$p \leq 1 - (1-\alpha)^{1/N}$$

For the general case at the FBI, $\alpha = 0.01$, representing a confidence level of 99%, and N = 260 million, a conservative population size value equal to the population of the United States.

The adjusted probability (p) will be derived as follows:

1. According to NRC II Report recommendations (2), multiply the frequencies from the alleles of the profile loci to obtain an overall profile frequency from each of the four major databases (Caucasian, African-American, Southeastern Hispanic and Southwestern Hispanic) routinely used by DNAU I. Other population groups can be considered on a case-by-case basis.

2. Select the most common frequency.

3. Increase the frequency by a factor of ten. This incorporates the recommendation of the 1996 NRC II report that "It is safe to assume that within a race, the uncertainty of a value calculated from adequate databases by the product rule is within a factor of ten above and below the true value" (page 156), and represents an additional conservative feature of the approach. This probability is the adjusted probability p.

If p is less than 3.87×10^{-11}, ($1/p \approx 2.6 \times 10^{10}$) then the following statement can be included in the report:

"K1 is the source of the DNA obtained from specimen Q1 to a reasonable degree of scientific cer-

tainty", where K1 and Q1 represent the known and unknown specimens, respectively.

This statement is used in FBI Laboratory reports replacing the statement which provided the estimated frequencies associated with selecting an unrelated individual at random from the specific population(s), when the threshold is surpassed. The frequency calculations will be maintained in the case notes and provided if needed for discovery purposes.

To determine identity in ethnic populations or subgroups, such as Native Americans, the procedure is similar that described above except N will be based on either the size of the relevant subgroup and/or the potential contributors of the sample at the crime scene. The report will contain the statement:

"K1 is the source of the DNA obtained from specimen Q1 to a reasonable degree of scientific certainty", where K1 and Q1 represent the known and unknown specimens, respectively.

The above assessments involve determining the value for p for unrelated individuals. In some cases, it may be necessary to consider the probability that a relative of a suspect may have the same profile. If it is not possible to obtain known standards from pertinent siblings or other relatives, the conditional probability, p', can be calculated using formulae described in the NRC II report (2, page 113). These formulae are used in the Popstats program of the CODIS system and can be found in the publication "Popstats 5.1 Calculation Specification" on page 17 (4). The procedure for the DNAUI for determining identity among related individuals is as follows:

1. Determine the p' of the DNA profile in each database

2. Select the most common value of p'

3. If $p' \leq 1 - (1-\alpha)^{1/N'}$, where N' represents the appropriate family size of the matching individual (set to a conservative value of at least 10) and $\alpha = 0.01$, then the following statement can be included in the report:

"K1 is the source of the DNA obtained from specimen Q1 to a reasonable degree of scientific certainty", where K1 and Q1 represent the known and unknown specimens, respectively.

If the p or p' value(s) is greater than the established threshold, examiners in the DNAU I will report the estimated DNA profile frequency in the case report as they have done in the past.

In October, the first FBI Laboratory report stating source attribution was issued in the investigation of a serial rapist from Milwaukee. This evidence powerfully refuted several alibi witnesses who testified on behalf of the defendant Iran Shuttlesworth, a correctional officer. He was accused of kidnapping and first degree sexual assault of two Milwaukee women. FBI DNA examiner Thomas Callaghan testified and opined that Shuttlesworth was the source of the semen found on the victim's clothing. This was the first time that a DNA examiner directly identified an individual as the source of the stain based on DNA testing.

In conclusion, the time has arrived for society to enjoy the full potential of DNA profiling. In many cases sufficient information is obtained such that the DNA profile is so exceedingly rare to exceed the reciprocal of the world population by orders of magnitude. Source attribution can and should be provided. As anticipated, attorneys may question and judges may decide the weight of these determinations, but it is within the purview of the scientist to report and opine the identification (5).

REFERENCES

1. Weir B.S., (1995) DNA Statistics in the Simpson Matter. *Nature Genetics*, **11**: 365-368.

2. National Research Council, (1996) "The Evaluation of Forensic DNA Evidence", National Academy Press, Washington D.C.

3. Budowle, *et al.* (in prep)

4. Federal Bureau of Investigation, (1997) "Popstats 5.1 Calculation Specification", U.S. Department of Justice, Washington D.C.

5. State v. Curtis Buckner, (1997) 133 Wash.2d 63, 941 P.2d 667.