Reference Guide on Forensic DNA Evidence

Judith A. McKenna
Joe S. Cecil
Pamela Coukos

Judith A. McKenna, J.D., Ph.D., and Joe S. Cecil, J.D., Ph.D., are members of the staff in the Research Division of the Federal Judicial Center, Washington, D.C. Pamela Coukos, J.D., was a legal assistant in the Research Division of the Federal Judicial Center and is currently a law clerk in the United States District Court for the Eastern District of New York.

Contents

I. Introduction

This reference guide addresses technical issues that arise in considering the admissibility of and weight to be accorded analyses of *forensic samples* of *deoxyribonucleic acid* (DNA). We address only the best-established form of forensic DNA identification analysis, known as *restriction fragment length polymorphism* (RFLP) *analysis*.[1] Scientific knowledge in this area is evolving rapidly, and analytical techniques based on different principles are being developed. Technologies based on the *polymerase chain reaction* (PCR) technique of DNA amplification may become the dominant modes of analysis in the relatively near future. PCR-based analysis techniques are discussed briefly in section IV.A.

In this guide we set forth an analytical framework judges can use when DNA identification evidence is offered by expert witnesses. Therefore we describe only the basic principles of the technique; we do not address the general expert witness qualification matters raised by Federal Rule of Evidence 702 or procedural issues such as discovery.[2] Although we cite both state and federal cases for illustrative purposes, this reference guide is not intended as a guide to the status of DNA evidence in any particular jurisdiction.[3]

DNA analysis is based on well-established principles of the wide genetic variability among humans and the presumed uniqueness of an individual's genetic makeup (identical twins excepted). Laboratory techniques for isolating and observing the DNA of human *chromosomes* have long been used in nonforensic scientific settings. The forensic application of the technique involves comparing a known DNA sample obtained from a suspect with a DNA sample obtained from the crime scene, and often with one obtained from the victim. Such analyses typically are offered to support or refute the claim that a criminal suspect contributed a biological specimen (e.g., semen or blood) collected at a crime

1. Both RFLP and PCR may be supplanted by new techniques that rely on direct sequencing of genes, which may avoid some of the chief problems associated with current techniques. *See, e.g.,* Alec J. Jeffreys et al., *Minisatellite Repeat Coding as a Digital Approach to DNA Typing,* 354 Nature 204 (1991).

2. *See, e.g.,* Paul C. Giannelli, *Criminal Discovery, Scientific Evidence, and DNA*, 44 Vand. L. Rev. 793 (1991). *See also* Margaret A. Berger, Evidentiary Framework § II, in this manual.

3. Some states have adopted statutes that expressly provide that such analyses are admissible as evidence in criminal proceedings under certain circumstances. States that have adopted statutes addressing the admissibility of DNA identification evidence include: Indiana (Ind. Code § 35-37-4-13 (1993)); Louisiana (La. Rev. Stat. Ann. §15:441.1 (West 1992 & Supp. 1994)); Maryland (Md. Code Ann., Cts. & Jud. Proc. §10-915 (1989 & Supp. 1993)); Minnesota (Minn. Stat. Ann. §§ 634.25–634.26, 634.30 (West Supp. 1994)); Tennessee (Tenn. Code Ann. § 24-7-117 (Supp. 1993)); and Virginia (Va. Code Ann. §19.2-270.5 (Michie 1990)).

scene.[4] For example, an analyst may testify on the basis of a report that includes the following:

> Deoxyribonucleic acid (DNA) profiles for [the specific sites tested] were developed from specimens obtained from the crime scene, from the victim, and from the suspect. Based on these results, the DNA profiles from the crime scene match those of the suspect. The probability of selecting at random from the population an unrelated individual having a DNA profile matching the suspect's is approximately 1 in 200,000 in Blacks, 1 in 200,000 in Whites, and 1 in 100,000 in Hispanics.

An objection to this sort of testimony usually comes before the court when the defense moves to exclude the testimony and the report. Such a motion can be made before or during trial, depending on circumstances and the court's rules regarding in limine motions. Before the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[5] such disputes often were aired in hearings devoted to determining whether the theory and techniques of DNA identification were generally accepted by the relevant scientific community and so satisfied the *Frye* standard.[6] Since *Daubert*, general acceptance remains an issue but only one of several a court may consider when DNA evidence is offered. Frequently disputed issues include the validity of applying the standard RFLP technique to crime samples, the proper interpretation of test results as showing a match, and the appropriate statistical determination of the probability of a coincidental match.

To address concerns that had been raised regarding the use of DNA identification evidence in forensic contexts, the National Research Council (NRC) of the National Academy of Sciences convened a committee of scientists, forensic scientists working in law enforcement, and legal scholars. The committee's report[7] offers conclusions about the scientific validity and reliability of DNA analysis and makes recommendations concerning its use as evidence in court. The report, which recommends that courts take judicial notice of both the theory underlying DNA analysis and the ability of the most commonly used method of testing to distinguish reliably between different sources of DNA, confirmed the conclusion several courts had reached. However, the committee also recognized actual and potential problems with the use of current DNA identification methods and recommended steps by which the scientific and forensic communities could address them. Throughout this guide, we cite, where appropriate, the

4. Although the technique has other uses, this is the usual context in which such evidence comes before a federal court. Other considerations may arise where DNA analysis is used to narrow the field of suspects by comparing a crime sample with samples from a blood bank, to establish the commission of a crime where no body is found, or to establish parentage.

5. 113 S. Ct. 2786 (1993).

6. Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

7. Committee on DNA Technology in Forensic Science, National Research Council, DNA Technology in Forensic Science (1992) [hereinafter NRC Report].

NRC committee's position on issues still in dispute, but such citation should not be taken as endorsement of the committee's conclusions.

II. Overview of RFLP Analysis[8]

RFLP analysis is based on the observable variability of human genetic characteristics. Human genetic information (one's *genome*) is encoded primarily in chromosomal DNA, which is present in most body cells.[9] Except for sperm and egg cells, each DNA-carrying cell contains 46 chromosomes. Forty-four of these are arranged in homologous pairs of one *autosome* (nonsex chromosome) inherited from the mother and one autosome from the father. The cells also carry two sex chromosomes (an X from the mother and either an X or a Y from the father). Chromosomal DNA sequences vary in length and are made up of four organic bases (*adenine* (A), *cytosine* (C), *thymine* (T), and *guanine* (G)). A pairs only with T; C pairs only with G. These sequences of *base pairs* are arranged in long chains that form the twisted *double helix,* or ladder structure, of DNA. Thus, if the bases on one side of the helix or ladder are represented as CATAGAT, the complementary side would be GTATCTA.

Most DNA-carrying cells in a human contain the same information encoded in the approximately 3.3 billion base pairs per set of chromosomes in each cell.[10] More than 99% of the base pairs in human cells are the same for all individuals, which accounts for the many common traits that make humans an identifiable species. The remaining base pairs (about 3 million) are particular to an individual (identical twins excepted), which accounts for most of the wide variation that makes each person unique.

A *gene* (characteristic DNA sequence) is found at a particular site, or *locus*, on a particular chromosome. For instance, a gene for eye color is found at the same place or locus on the same chromosome in every individual.[11] Normal individuals have two copies of each gene at a given locus—one from the father and one from the mother. A locus on the DNA molecule where all humans have the

8. More detailed descriptions of these principles may be found in the NRC Report, *supra* note 7, at 27–40; United States v. Jakobetz, 955 F.2d 786, 791–93 (2d Cir.), *cert. denied*, 113 S. Ct. 104 (1992); and in United States v. Yee, 134 F.R.D. 161, 169 (N.D. Ohio 1991), *aff'd sub nom.* United States v. Bonds, 12 F.3d 540 (6th Cir. 1993).

9. Mature red blood cells do not carry chromosomal DNA. However, white blood cells do carry chromosomal DNA and can be used for DNA analysis.

10. The only significant universal exception to this fact is that a sex cell (i.e., an egg or sperm) contains only twenty-three chromosomes, which vary from cell to cell (e.g., from sperm to sperm). This accounts for the genetic individuality of offspring.

11. This consistency of genetic localization has made it possible for geneticists to begin mapping the human genome.

same genetic code is called *monomorphic*. However, genes vary. An individual may receive the genetic code for blue eyes from his or her mother and the genetic code for brown eyes from his or her father. An alternative form of a gene is known as an *allele*.[12] A locus where the allele differs among individuals is called *polymorphic*, and the difference is known as a *polymorphism*.

Although some polymorphisms have been found to govern what makes individuals observably distinct from one another (e.g., eye color), others serve no known function. Among these noncoding DNA regions are some in which certain base pair sequences repeat in tandem many times (e.g., CATCATCAT . . . paired with GTAGTAGTA . . .). This is known as a *Variable Number of Tandem Repeats*, or a VNTR. The number of base pairs and the sequence of pairs vary from locus to locus on one chromosome and from chromosome to chromosome. RFLP analysis allows scientists to determine the size of a repetitive sequence. Because the length of these sequences (sometimes called *band size*) of base pairs is highly polymorphic, although not necessarily unique to an individual, comparison of several corresponding sequences of DNA from known (suspect) and unknown (forensic) sources gives information about whether the two samples are from the same source.

Appendix A is a detailed schematic of the steps in RFLP analysis. First, DNA is extracted from an evidence sample collected at the crime scene. Second, it is digested by a *restriction enzyme* that recognizes a particular known sequence called a *restriction site* and cuts the DNA there. The result is many DNA fragments of varying sizes. Third, digested DNA from the crime sample is placed in a well at the end of a lane in an agarose gel, which is a gelatin-like material solidified in a slab about five inches thick. Digested DNA from the suspect is placed in another well on the same gel. Typically, control specimens of DNA fragments of known size, and, where appropriate, DNA specimens obtained from a victim, are run on the same gel. Mild electric current applied to the gel slowly separates the fragments in each lane by length, as shorter fragments travel farther than longer, heavier fragments. This procedure is known as *gel electrophoresis*.

Fourth, the resulting array of fragments is transferred for manageability to a sheet of nylon by a process known as *Southern blotting*. Either during or after this transfer, the DNA is *denatured* ("unzipped") by heating, separating the double helix into single strands. The weak bonds that connect the two strands are susceptible to heat and salinity. The double helix can be unzipped or denatured without disrupting the chain on either side. Fifth, a *probe*—usually with a radioactive tag—is applied to the membrane. The probe is a single strand of DNA that *hybridizes* with (binds to) its complementary sequence when it is applied to

12. When an individual's two alleles for a locus are the same (e.g., two blue-eye alleles), the individual is said to be *homozygous* for that locus. When the alleles are not the same, the individual is said to be *heterozygous* for that locus. *See* Office of Technology Assessment, U.S. Congress, Genetic Witness: Forensic Uses of DNA Tests 42 (1990) [hereinafter OTA Report].

the samples of denatured DNA. The DNA locus identified by a given probe is found by experimentation, and individual probes often are patented by their developers. Different laboratories may use different probes (i.e., they may test for alleles at different loci). Where different probes are used, test results are not comparable.

Finally, excess *hybridization* solution is washed off, and the nylon membrane is placed between at least two sheets of photographic film. Over time, the radioactive probe material exposes the film where the biological probe has hybridized with the DNA fragments. The result is an *autoradiograph*, or an *autorad*, a visual pattern of bands representing specific DNA fragments. An autorad that shows two bands in a single lane indicates that the source is heterozygous for that locus (i.e., he or she inherited a different allele from each parent). If the autorad shows only one band, the person may be homozygous for that allele (i.e., each parent contributed the same variant of the gene).[13] Together, the two alleles make up the person's *genotype* (genetic code) for the specific locus associated with the probe.

Once an appropriately exposed autorad is obtained, the probe is washed from the membrane, and the process is repeated with different probes that bind to different sequences of DNA. Three to five probes are typically used, the number depending in part on the amount of testable DNA recovered from the crime sample.[14] The result is a set of autorads, each of which shows the results of one probe.[15] In Appendix B there is a copy of an autorad with illustrative DNA patterns. Illustrations of the results of all probes depict an overlay of multiple films or the results from a multilocus probe.

If the two DNA samples are from the same source, and if the laboratory procedures are conducted properly, hybridized DNA fragments of approximately the same length should appear at the same point in the suspect and evidence specimen lanes. If, on visual inspection, the DNA band patterns for the suspect and the evidence sample appear to be aligned on the autorad, this impression is verified by a computerized measurement. If the two bands fall within a specified length, or *match window* (e.g., ± 2.5% of band length), a match is declared for that probe or allele.[16] For forensic purposes, a match means that the patterns are

13. However, the appearance of a single band may also indicate problems with the analysis, and the analyst should explore alternative explanations when homozygosity cannot be established (as by testing the DNA of the parents or performing other checks). *See* NRC Report, *supra* note 7, at 58.

14. Another type of DNA analysis uses a *multilocus probe*, which hybridizes with multiple locations simultaneously, yielding a more complex DNA pattern. In general, this technique is most appropriate in paternity cases. In the United States it is not preferred for analyses in criminal cases. *See* NRC Report, *supra* note 7, at 40.

15. As the membrane is compressed between two films for each probe, two autorads per probe actually result. However, one film per probe is checked during the process to see whether the process is complete. As a result, the image on the film removed first may be weaker than the final print and generally is not used.

16. The match window is intended to accommodate measurement error or differences that may result when the tests are run repeatedly on the same sample.

consistent with the conclusion that the two DNA samples came from the same source.[17] Taken together, the results of the probes form the *DNA profile*.[18]

17. The interpretation of autorads is discussed in more detail *infra* § VI.

18. Throughout this reference guide, we use the term DNA profile to refer to the autorad pattern yielded by the test probes. If it were feasible to map the entire genome, it is believed that a unique profile would be obtained for each individual.

III. Theory and Technique of RFLP Analysis

A. Is the Scientific Theory Underlying RFLP Valid?

The basic genetic theory underlying DNA profiling (e.g., the fundamental structure of DNA and the observable polymorphism of some areas of DNA) essentially is undisputed, as several courts have noted, and courts may want to consider whether the theory is now a proper subject for judicial notice.[19] The ability to discriminate between human genetic profiles, using enough test sites, is well accepted.[20]

B. Are the Laboratory Techniques Used in RFLP Analysis Valid and Reliable?

The basic laboratory procedures used in RFLP analysis have been used in nonforensic settings for many years and are generally accepted by molecular biologists.[21] Most courts addressing the issue have concluded that there also appears to be a broad consensus among molecular biologists that properly conducted RFLP analysis can produce reliable information relevant to the identification of a forensic sample of DNA. Courts reaching this conclusion have applied the *Frye* test,[22] the reliability test,[23] a combination of the two,[24] and the *Daubert*

19. *See, e.g.,* State v. Davis, 814 S.W.2d 593, 602 (Mo. 1991), *cert. denied sub nom.* Davis v. Missouri, 112 S. Ct. 911 (1992); People v. Castro, 545 N.Y.S.2d 985, 987 (1989).

20. The NRC committee has recommended that courts take judicial notice that "[t]he study of DNA polymorphisms can, in principle, provide a reliable method for comparing samples," and that "[e]ach person's DNA is unique (with the exception of identical twins), although the actual discriminatory power of any particular DNA test will depend on the sites of DNA variation examined." NRC Report, *supra* note 7, at 149.

21. The Office of Technology Assessment has concluded that "no scientific doubt exists that technologies available today accurately detect genetic differences." OTA Report, *supra* note 12, at 59.

22. Smith v. Deppish, 807 P.2d 144, 157–59 (Kan. 1991); State v. Davis, 814 S.W.2d 593, 600–03 (Mo. 1991), *cert. denied sub nom.* Davis v. Missouri, 112 S. Ct. 911 (1992); State v. Ford, 392 S.E.2d 781, 783–84 (S.C. 1990); State v. Schwartz, 447 N.W.2d 422, 424–25 (Minn. 1989); State v. Woodall, 385 S.E.2d 253, 259 (W. Va. 1989); Glover v. State, 787 S.W.2d 544, 547–48 (Tex. Ct. App. 1990), *aff'd,* 825 S.W.2d 127 (Tex. Crim. App. 1992); People v. Wesley, 633 N.E.2d 451 (N.Y. 1994).

23. State v. Brown, 470 N.W.2d 30, 32–33 (Iowa 1991); Caldwell v. State, 393 S.E.2d 436, 441 (Ga. 1990); Spencer v. Commonwealth, 384 S.E.2d 775, 783 & n.10 (Va. 1989), *cert. denied,* 493 U.S. 1036 (1990); Andrews v. State, 533 So. 2d 841, 846–47 (Fla. Dist. Ct. App. 1988) (applying a "relevancy/reliability" test).

24. United States v. Jakobetz, 995 F.2d 786, 792–97 (2d Cir.), *cert. denied*, 113 S. Ct. 104 (1992).

analysis.[25] The forensic applicability of RFLP analysis also has been recommended as a subject for judicial notice.[26]

C. With Respect to Disputed Issues, What Are the Relevant Scientific Communities?

Most courts have accepted molecular biologists as the relevant scientific community with respect to the laboratory techniques of isolating and probing the DNA. The statistical interpretation of the results has been more properly the province of population geneticists.[27] Individuals from other fields (e.g., genetic epidemiologists and biostatisticians) also may have the requisite background to testify about the analysis. Forensic scientists or laboratory technicians involved in the analysis often do not have a strong background in the relevant scientific discipline but may be knowledgeable about techniques of sample collection and preservation, forensic laboratory standards and procedures, and proficiency tests.

25. United States v. Bonds, 12 F.3d 540, 554–68 (6th Cir. 1993) (assessing DNA identification evidence in light of Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993)).

26. The NRC committee recommended that courts take judicial notice that

> [t]he current laboratory procedure for detecting DNA variation (specifically, single-locus probes analyzed on Southern blots without evidence of band shifting) is fundamentally sound, although the validity of any particular implementation of the basic procedure will depend on proper characterization of the reproducibility of the system (e.g., measurement variation) and the inclusion of all necessary scientific controls.

NRC Report, *supra* note 7, at 149.

27. United States v. Yee, 134 F.R.D. 161, 164 (N.D. Ohio 1991), *aff'd sub nom.* United States v. Bonds, 12 F.3d 540 (6th Cir. 1993); United States v. Jakobetz, 747 F. Supp. 250, 259–60 (D. Vt. 1990), *aff'd*, 955 F.2d 786 (2d Cir.), *cert. denied*, 113 S. Ct. 104 (1992).

IV. Sample Quantity and Quality

RFLP analysis requires a suitable sample of DNA. Several factors may affect a sample's suitability for analysis. For each factor claimed to affect a particular analysis, the court may want to have the experts address whether its influence is likely to cause a false positive result (incorrect identification of the suspect as a potential source of the forensic DNA) or merely an inconclusive or uninterpretable result.[28]

A. Did the Crime Sample Contain Enough DNA to Permit Accurate Analysis?

To be interpretable, the crime sample must contain enough DNA of sufficiently high molecular weight to allow isolation of longer DNA fragments, which are the most susceptible to degradation. Samples of blood, semen, or other DNA sources may be too small to permit analysis. We are aware of no evidence to suggest that small sample size increases the likelihood that an interpretable test will yield a false match for a given probe or allele. However, to the extent that small sample size precludes a full series of tests (three to five probes), it can significantly diminish the power of RFLP analysis to distinguish between DNA samples obtained from different individuals. In addition, the unavailability of additional DNA precludes repeated testing that might verify or refute the initial test.

The amount of testable DNA may be increased by a technique known as the polymerase chain reaction, or PCR amplification. PCR mimics DNA's self-replicating properties to make up to millions of copies of the original DNA sample in only a few hours.[29] Although the term PCR often is used loosely to refer to the

28. Some factors, including contamination, may cause a false negative result, but courts seldom encounter the problem of false negatives in criminal cases.

29. For descriptions of the PCR process, see OTA Report, *supra* note 12, at 47–50; George F. Sensabaugh & Cecilia Von Beroldingen, *The Polymerase Chain Reaction: Application to the Analysis of Biological Evidence, in* Forensic DNA Technology 63, 64–67 (Mark A. Farley & James J. Harrington eds., 1991); Lorne T. Kirby, DNA Fingerprinting: An Introduction 75–76 (1990); Russell Higuchi & Edward Blake, *Applications of the Polymerase Chain Reaction in Forensic Science, in* DNA Technology and Forensic Science 265, 266–68 (Jack Ballantyne et al. eds., 1989); Henry A. Erlich et al., *Recent Advances in the Polymerase Chain Reaction,* 252 Science 1643, 1649 (1991); NRC Report, *supra* note 7, at 40–42. PCR requires special attention to issues of quality control and contamination. Courts presented with evidence derived from PCR-based methods may want to consult the NRC Report, *supra* note 7, at 63–70, and the OTA Report, *supra* note 12, at 69–71.

entire process of replicating DNA and testing for the presence of matching alleles, the term properly refers only to the *replication* portion of that process. After amplifying a DNA sample with PCR, technicians must use other methods to determine whether a known and unknown sample match. Standard RFLP analysis can be used in many circumstances, but other techniques often are used, including a process using *sequence-specific oligonucleotide* (SSO) *probes*. Currently, one locus, called *HLA DQ* α,[30] is available for this process. The technique is faster than RFLP but less discriminating and therefore somewhat less powerful as an identification tool. A second technique, *amplified fragment length polymorphism* (AMP-FLP), is under development and may be used soon in criminal investigations. It relies on amplification of VNTR loci and usually uses gel electrophoresis. Other detection systems are either too novel or too unreliable to be addressed here.[31]

B. Was the Crime Sample of Sufficient Quality to Permit Accurate Analysis?

Exposure to heat, moisture, and ultraviolet radiation can degrade the DNA sample. Samples also may have been contaminated by exposure to chemical or bacterial agents that alter DNA, interfere with the enzymes used in the testing process, or otherwise make DNA difficult to analyze. Such exposure is known as *environmental insult*.[32]

Although old samples of DNA may be analyzed successfully, attention must be given to possible sample degradation due to age. Again, we are aware of no evidence to suggest that age-degraded samples are likely to produce false positive results. Courts may want to ask the experts whether the research has addressed the effects of specific types of environmental insult likely in a particular case.

C. How Many Sources of DNA Are Thought to Be Represented in the Crime Sample?

Often, the expected composition of a crime sample of DNA can be narrowed to a single perpetrator, a single victim, or both. However, a crime sample may be thought to include DNA from multiple sources, as where more than one person

30. *Human Leukocyte Antigens* (HLAs) are antigens (foreign bodies that stimulate an immune system response) located on the surface of most cells (excluding red blood cells and sperm cells). HLAs differ among individuals and are associated closely with transplant rejection. HLA DQ α is a particular class of HLA whose locus has been completely sequenced and thus can be used for forensic typing.

31. Some of these other techniques include direct sequencing of DNA samples, which promises theoretically absolute identification (and raises all the issues and concerns associated with capturing an individual's genetic code), but the process is not yet sufficiently developed for use in forensic applications. *See* NRC Report, *supra* note 7, at 43–44.

32. *See, e.g.,* Dwight E. Adams et al., *Deoxyribonucleic Acid (DNA) Analysis by Restriction Fragment Length Polymorphisms of Blood and Other Body Fluid Stains Subjected to Contamination and Environmental Insults*, 36 J. Forensic Sci. 1284 (1991).

is thought to have contributed to the crime sample of blood or semen. Male and female DNA extracted from such a sample may be distinguished, as can same-sex DNA where the alternative source is known and available for testing (e.g., a rape victim's husband). The presence of multiple, same-sex samples from unknown sources raises additional complications. Mixed samples can be difficult to interpret, although the intensity of different bands can offer clues. Courts are most likely to encounter the possibility of mixed samples when analysts report extra or anomalous bands; they should be alert to the possible explanations for such bands and inquire into the laboratory's procedures for testing those explanations.[33]

33. *See* NRC Report, *supra* note 7, at 58–59.

V. Laboratory Performance of the RFLP Analysis Offered as Evidence

Adherence to proper test procedures is critical to accurate comparison of DNA samples. Before the NRC issued its report, guidelines for laboratory procedures, quality control, and proficiency testing were promulgated jointly by the Technical Working Group on DNA Analysis Methods (TWGDAM) and the California Association of Criminalists Ad Hoc Committee on DNA Quality Assurance.[34] These guidelines (TWGDAM Guidelines) have been used by courts to assess the quality of individual test results and testing laboratories[35] and enjoy wide, albeit not universal, acceptance as appropriate laboratory pro cedures for RFLP analysis.[36] This section incorporates some of the TWGDAM Guidelines and can be used to identify disputed issues regarding the testing laboratory and the individual RFLP analysis. The report of the NRC also contains a detailed discussion of quality-assurance issues and strongly favors adoption of quality control and proficiency-testing standards.[37]

A. Has the Testing Laboratory Demonstrated a Record of Proficiency and Quality Control Sufficient to Permit Confidence That the Tests Were Conducted Properly?

1. Does the laboratory maintain appropriate documentation?

The TWGDAM Guidelines state that the DNA laboratory "must maintain documentation on all significant aspects of the DNA analysis procedure, as well as any related documents or laboratory records that are pertinent to the analysis or

34. TWGDAM is a practitioners' group, largely sponsored by the Federal Bureau of Investigation (FBI) and comprising representatives of the FBI, the National Institute of Standards and Technology, the Royal Canadian Mounted Police, and state forensic laboratories.

35. State v. Schwartz, 447 N.W.2d 422, 426 (Minn. 1989). In particular, the protocol used by the FBI has been held to be sufficiently reliable. United States v. Jakobetz, 955 F.2d 786, 799–800 (2d Cir.), *cert. denied,* 113 S. Ct. 104 (1992); Prater v. State, 820 S.W.2d 429, 436 (Ark. 1991).

36. United States v. Yee, 134 F.R.D. 161, 202–06 (N.D. Ohio 1991), *aff'd sub nom.* United States v. Bonds, 12 F.3d 540 (6th Cir. 1993). The NRC committee described the TWGDAM Guidelines as an "excellent starting point for a quality-assurance program," but recommended that certain technical recommendations be added. NRC Report, *supra* note 7, at 16, 99.

37. *See* NRC Report, *supra* note 7, at 97–109.

interpretation of results, so as to create a traceable audit trail."[38] The guidelines list seventeen areas that must be covered by the documentation.[39]

2. Has the laboratory's procedure been validated?

The TWGDAM Guidelines stress the importance of validation "to acquire the necessary information to assess the ability of a procedure to reliably obtain a desired result, determine the conditions under which such results can be obtained and determine the limitations of the procedure."[40] Each locus or probe to be used in the analysis should be validated separately.[41] The court may want to ask whether the laboratory's procedures have been validated in accordance with the minimum standards set forth by section 4 of the TWGDAM Guidelines.[42]

3. Has the laboratory been subjected to appropriate proficiency testing? With what results?

Studies of the quality and accuracy of laboratory analysis in many forensic areas have raised profound concerns about the accuracy of evidence presented in criminal cases.[43] Section 9 of the TWGDAM Guidelines recommends that DNA laboratories participate in appropriately designed proficiency-testing programs, preferably programs conducted by outside institutions.[44] The court may want to inquire about the relevant laboratory's participation in external or internal proficiency-testing programs, the results of such testing, and the significance of the results. Some commentators have suggested that, at least when an adequate history of repeated proficiency testing becomes available, expert testimony should report not only the likelihood of a coincidental match according to applicable *population genetics* but also the applicable laboratory error rates.[45] Some

38. Technical Working Group on DNA Analysis Methods (TWGDAM) & California Ass'n of Criminalists Ad Hoc Comm. on DNA Quality Assurance, *Guidelines for a Quality Assurance Program for DNA Analysis*, 18 Crime Laboratory Dig. 44, 52 (1991) [hereinafter *TWGDAM Guidelines*] (*see infra* Appendix C).

39. *Id.* §§ 3.1–3.17, at 52 (*see infra* Appendix C).

40. *Id.* § 4.1.1, at 53 (*see infra* Appendix C).

41. *Id.* §§ 4.1.1–4.1.3, at 53 (*see infra* Appendix C).

42. *Id.* § 4, at 53–56 (*see infra* Appendix C).

43. *See* Edward J. Imwinkelried, *The Debate in the DNA Cases Over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis,* 69 Wash. U.L.Q. 19, 25–27; and studies cited therein reporting error rates in forensic testing.

44. *TWGDAM Guidelines*, *supra* note 38, § 9.2, states:

> Blind Proficiency Testing. Ideally, blind proficiency test specimens should be presented to the testing laboratory through a second agency. These samples should appear to the examiner/analyst as routine evidence. The blind proficiency test serves to evaluate all aspects of the laboratory examination procedure, including evidence handling, examination/testing and reporting. It is highly desirable that the DNA laboratory participate in a blind proficiency test program, and every effort should be made to implement such a program.

45. Arguments over what information juries should be given concerning laboratory error focus in part on considerations of how to give jurors the most diagnostic information possible in a particular case and in part on policy considerations of how to increase the diagnosticity of DNA identification evidence overall by increasing incentives for laboratories to improve their procedures and submit to large-scale proficiency testing. *See, e.g.,* Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the Rest of*

also argue that the laboratory error rate should be combined with the probability of a coincidental match derived from the population analysis to yield an overall probability that the analyst would report a match when there was no true match.[46] This approach would incorporate the *Daubert* factor of laboratory error rate in such a way that it would go to weight, not admissibility. There is no reported federal decision in which a court has imposed this constraint on the expert's testimony.

In evaluating proficiency-testing results, a court may want to attach more weight to blind tests than to open tests and may inquire as to the type of samples that were actually tested (e.g., pristine samples or samples subjected to environmental insults typical of forensic casework). Anonymous testing may give some insights into whether the procedures are generally reliable, but nonanonymous testing provides stronger evidence of the quality of a laboratory's work. Several earlier studies of forensic laboratory work have found different error rates; differences were accounted for in part by whether false negatives (incorrect exclusions) and inconclusive results were counted as errors.[47] Because one purpose of proficiency testing is to help laboratories identify problem areas and correct them, it is important to focus on tests of a laboratory's current procedures where possible.

B. Was the Crime Sample in This Case Handled Properly?

When the normal chain-of-custody evidence is adduced, the court should allow inquiry into potential sources of error in sample handling by laboratory personnel.

1. What opportunities for laboratory mislabeling were present?

It is especially important that the laboratory follow procedures for ensuring that the known sample is not mislabeled as, or mixed with, the crime sample. These errors can lead to false positive results that incriminate the wrong suspect.

2. What sources of possible laboratory contamination were present?

Potential contamination of the sample by reagents and other substances found in the laboratory is also of concern, although such contamination is less likely than mislabeling to create a false positive result. Contamination is of particular concern when PCR amplification is used, because even a small fragment of DNA from a foreign source may be amplified many times.

Forensic Science, 13 Cardozo L. Rev. 361, 368–70 (1991); and Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 Judicature 222, 228–29 (1993). For a discussion of the competing considerations, see Richard Lempert, *Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 Cardozo L. Rev. 303, 323–28 (1991).

46. For examples of ways such a combination might be derived, see Saks & Koehler, *supra* note 45, at 368–70.

47. *See* OTA Report, *supra* note 12, at 79.

C. Was the RFLP Analysis in This Case Conducted Properly?

Minor departures from the TWGDAM Guidelines or other accepted standards are not necessarily fatal to the validity of the DNA evidence, and some courts have treated such departures as bearing on the weight of the evidence, not its admissibility.[48] Few courts have excluded DNA evidence because of a laboratory's failure to meet accepted standards, but a persuasive argument can be made that significant departures from accepted standards render the evidence unreliable and, hence, inadmissible. In any event, deviations from the laboratory's written procedures should be explored, and the court may want to have experts from both sides address the nature of errors that could be caused by deviations from the standards.[49]

48. United States v. Yee, 134 F.R.D. 161, 206–07 (N.D. Ohio 1991), *aff'd sub nom.* United States v. Bonds, 12 F.3d 540 (6th Cir. 1993).

49. *See* Thomas M. Fleming, Annotation, *Admissibility of DNA Identification Evidence*, 84 A.L.R.4th 313, 342–46 (1991 & Supp. 1993); State v. Schwartz, 447 N.W.2d 422, 426 (Minn. 1989). *See also* State v. Jobe, 486 N.W.2d 407, 420 (Minn. 1992) (*Frye* hearing is still required on whether the testing laboratory was in compliance with appropriate standards and controls, but not on the basic RFLP testing procedures themselves). The split of authority over whether compliance with proper test procedures is a foundational matter that should govern the admissibility of laboratory evidence, and over which party should bear the burden of showing compliance or noncompliance, is discussed in Imwinkelried, *supra* note 43.

VI. Comparison of DNA Profiles

The first step in interpreting the evidence is to determine if the DNA obtained from the crime sample is consistent with the DNA profile of the suspect or other known source. The RFLP procedure compares the length of the DNA fragments in the known sample with the length of the DNA fragments in the crime sample. Techniques used to declare profile matches vary in their details but typically use a two-step process of visual inspection and objective measurement of DNA bands that appear to present the same profile.

A. Does the DNA Profile from the Crime Sample Appear to Be Consistent with the DNA Profile of the Suspect?

A pair of lanes represents the results of one genetic probe on the suspect and crime samples, and each pair should be evaluated independently of every other pair. Distinct bands appearing, on visual inspection, to be at the same position for the two samples suggest that both sources of DNA have fragments of the same length (i.e., the same alleles). Only such a pattern will implicate the suspect as the source of the crime sample. If the bands do not appear to match, the analyst must determine if the test excludes the suspect as the source of the crime sample or if the test is inconclusive because of some shortcoming of the samples or the laboratory procedure. [50]

When bands of different specimens from the same source do not appear, on visual inspection, to be in the same position in the lane, an exclusion usually is declared. Occasionally, the two profiles obtained will appear to be highly similar, even though the bands do not appear in the same lane position. When the bands are consistently offset or shifted relative to one another, some analysts may declare a match, attributing the differences in position to a phenomenon known

50. *See, e.g.,* State v. Woodall, 385 S.E.2d 253, 260 (W. Va. 1989) (DNA test inconclusive due to poor quality of forensic sample). The distinction is most likely to become relevant to federal courts in the context of post-conviction efforts to reopen cases in which DNA evidence was not offered by the prosecution at the original trial. The FBI has reported that approximately 35% of the interpretable samples it has tested yielded conclusive "no match" decisions that excluded primary suspects. NRC Report, *supra* note 7, at 156. *See also* State v. Hammond, 604 A.2d 793, 800 (Conn. 1992) (FBI forensic analyst testified that DNA testing, which confirmed the results of blood type testing, showed defendant accused of rape could not have been the source of the semen stain; the court rejected the state's argument that these tests were unreliable due to "inherent flaws in the testing procedures themselves, or from contamination of the sample").

as *band-shift*.[51] Band shifting may occur with different samples from the same source because of differences in DNA concentration or other conditions within a specimen (e.g., salt concentration). Arguably, a match could be declared with offset but similar patterns when the DNA band sizes are consistent relative to internal lane markers of known size, which are in the lanes with the suspect and crime samples.[52] Declaring a match on the assumption that patterns would have matched in the absence of band-shift is controversial; the NRC committee recommends that samples that show apparent band shifting be declared "inconclusive" until the accuracy and reliability of proposed corrections (e.g., *monomorphic probes*) have been demonstrated through research.[53] Often, but not always, RFLP analyses can be repeated to avoid band shifting.

B. What Measurement Standard Was Used to Determine That the Bands Are Similar Enough to Declare a Match?

If the bands appear to be identical, the forensic analyst must measure the length of the bands by comparing their positions to standardized size markers included on the autorad. Repeated measurements of the same sample may differ slightly due to variation in laboratory materials and procedures, even if the suspect is the source of the crime sample. The standardized size markers are of known length (i.e., a known number of base pairs in a sequence), and the size of the sample fragment can be estimated in relation to these markers. A computer digitizes the images to be compared and determines whether they fall within a specified window of measurement variation. A valid analysis must apply an objective standard for declaring matching band patterns based on the variability observed when known crime samples are tested repeatedly.[54]

51. Band-shift is discussed in the NRC Report, *supra* note 7, at 60–61, 140–41.

52. *See, e.g.,* Caldwell v. State, 393 S.E.2d 436, 442–43 (Ga. 1990) (allowing testimony that DNA from a forensic sample matched suspect after band shifting was adjusted by measuring the position of the bands relative to internal lane markers).

53. NRC Report, *supra* note 7, at 61. *See also* People v. Keene, 591 N.Y.S.2d 733, 740 (N.Y. Sup. Ct. 1992) ("While the DNA principle and RFLP analysis are generally accepted in the scientific community, this court cannot find that the practice of using monomorphic probes to correct for band shift is a generally accepted test among molecular geneticists.").

54. For example, using a match window of ± 2.5% of the observed length of the fragment to estimate the true length of the fragment, if the measurement instrument reports that the length of a fragment is 2000 base pairs (bp), the true length is considered to fall within the window defined by 1950–2050 bp. *See also* NRC Report, *supra* note 7, at 72 (recommending a "precise and objective matching rule for declaring whether two samples match").

VII. Estimation of the Probability That the DNA Profiles Match by Coincidence

If a profile match is declared, it means only that the DNA profile of the suspect is consistent with that of the source of the crime sample. The crime sample may be from the suspect or from someone else whose profile, using the particular probes involved, happens to match that of the suspect. Expert testimony concerning the frequency with which the observed alleles are found in the appropriate comparison population is necessary for the finder of fact to make an informed assessment of the incriminating value of this match.[55]

The frequency with which an individual allele occurs in the comparison population is taken to be the probability of a coincidental match on that allele. These individual probabilities of a coincidental match are combined into an estimate of the probability of a coincidental match on the entire profile. This estimate is interpreted as the probability that a person selected at random from a comparison population would have a DNA profile that matches that of the crime sample. The probability estimate typically provided by a forensic expert cannot be interpreted strictly as the probability that an examiner will declare a match when the samples are actually from different sources. That probability is affected by other factors, the most important of which is the chance of laboratory error.

55. Statistical testimony concerning the likelihood of a DNA profile matching by coincidence is necessary to assess the probative value of the matching profile. NRC Report, *supra* note 7, at 74 ("To say that two patterns match, without providing any scientifically valid estimate (or, at least, an upper bound) of the frequency with which such matches might occur by chance is meaningless."). *See also* People v. Barney, 8 Cal. App. 4th 798, 817 (Cal. Ct. App. 1992) ("The statistical calculation step is the pivotal element of DNA analysis, for the evidence means nothing without a determination of the statistical significance of a match of DNA patterns."); D. H. Kaye, *The Forensic Debut of the National Research Council's DNA Report: Population Structure, Ceiling Frequencies and the Need for Numbers,* 34 Jurimetrics J. 369, 381 (1994) ("As a legal matter, a completely unexplained statement of a 'match' should be inadmissible because it is too cryptic to be weighed fairly by the jury, . . . [H]ow to present to a jury valid scientific evidence of a match is a legal rather than a scientific issue falling far outside the domain of the general acceptance test and the fields of statistics and population genetics. Thus, it would not be 'meaningless' to inform the jury that two samples match and that this match makes it more probable, in an amount that is not precisely known, that the DNA in the samples comes from the same person."). *But see* United States v. Martinez, 3 F.3d 1191, 1199 (8th Cir. 1993) (ruling that admitting evidence of a DNA profile match without evidence concerning the statistical probability of a coincidental match was not reversible error where defendant stipulated that statistical evidence was not required), *cert. denied*, 114 S. Ct. 734 (1994).

Differences in scientific opinion arise with respect to two main issues: (1) the appropriate method for computing the estimated probability of a coincidental match of a DNA profile; and (2) the selection of an appropriate comparison population. These issues are addressed below.

A. What Procedure Was Used to Estimate the Probability That the Individual Alleles Match by Coincidence?

Ascertaining an allele's frequency in a given population is essentially an empirical exercise. A sample of individuals is drawn from the designated population, their DNA is examined with genetic probes used in forensic analysis, and a table of frequencies is developed. For example, the FBI has constructed frequency tables using a *fixed-bin method*, in which standardized size markers are used to define boundaries of bins into which are sorted the fragment sizes observed in a sample population. Table 1, for example, shows a distribution of 858 alleles observed in a Caucasian sample, using the probe known as TBQ7, for the locus D10S28. The probability of a band of a specific length occurring by chance is computed by dividing the number of bands falling in the assigned bin by the total number of alleles observed in the sample of persons tested for that probe.[56] Thus, in Table 1, fragment lengths falling in Bin 7 are more common (75/858, or 0.087) than fragment lengths falling in Bin 20 (12/858, or 0.014). Similar tables have been and continue to be constructed using different probes in different populations.[57] The size of a matching DNA band is cross-referenced in the table of frequencies generated by the examination of the appropriate comparison sample.

56. If a bin contains fewer than five bands, the FBI merges that bin into an adjacent bin of higher frequency. This merging of bins is believed to yield a more conservative estimate of the probability of a random match. United States v. Yee, 134 F.R.D. 161, 172 (N.D. Ohio 1991), *aff'd sub nom.* United States v. Bonds, 12 F.3d 540 (6th Cir. 1993). Ranajit Chakraborty & Kenneth K. Kidd, *The Utility of DNA Typing in Forensic Work,* 254 Science 1735, 1738 (1991). An alternative approach used by some testing laboratories is the *floating bin method*, in which the frequency of an allele in a data set is calculated by counting the number of alleles falling into a bin centered on the allele of interest with a width specified by the matching rule.

57. Note, however, that courts have excluded DNA evidence when a laboratory's standards for declaring a match in individual samples have differed from the standards used in creating the allelic frequency tables. Under this analysis, if a ± 2.5% match window is used when the suspect's DNA is measured, the same standard should have been used when the alleles observed in the reference sample were sorted into bins. State v. Pennell, 584 A.2d 513, 521–22 (Del. Super. Ct. 1989) (court excluded evidence of population frequency because of different matching criteria for comparing prints with the database and for determining the population frequencies); People v. Castro, 545 N.Y.S.2d 985, 998 (1989) (DNA identification evidence declared inadmissible because, among other reasons, different criteria were used in comparing the sample prints with the population database).

Table 1
Rebinned Caucasian Population Data for TBQ7 for the Locus D10S28

Bin	Range (bp)	Count	Fraction
1	0–963	13	0.015
2	964–1077	44	0.051
3	1078–1196	38	0.044
4	1197–1352	15	0.017
5	1353–1507	34	0.040
6	1508–1637	67	0.078
7	1638–1788	75	0.087
8	1789–1924	71	0.083
9	1925–2088	40	0.047
10	2089–2351	51	0.059
11	2352–2522	16	0.019
12	2523–2692	14	0.016
13	2693–2862	36	0.042
14	2863–3033	42	0.049
15	3034–3329	41	0.048
16	3330–3674	56	0.065
17	3675–3979	39	0.045
18	3980–4323	62	0.072
19	4324–4821	58	0.068
20	4822–5219	12	0.014
21	5220–5685	6	0.007
22	5686–6368	23	0.027
23	6369–	5	0.006
Total		858	0.999

Source: Bruce Budowle et al., *A Preliminary Report on Binned General Population Data on Six VNTR Loci in Caucasians, Blacks and Hispanics from the United States,* 18 Crime Laboratory Dig. 9 tbl. 5 at 14 (1991).

Note: This table is presented for illustrative purposes; the FBI does not currently use the D10S28 locus in its analyses.

Assumptions of classical population genetics are used to estimate the probability that a person chosen at random from the specified population would ex-

hibit the same genotype as the suspect's.[58] The greater the probability of such a coincidental match, the lower the incriminating value of the evidence.

B. How Were the Probability Estimates for Coincidental Matches of Individual Alleles Combined for an Overall Estimate of a Coincidental Match of the Entire DNA Profile?

One of the most difficult and contentious issues in forensic use of DNA evidence is how to estimate the probability that two DNA profiles match by chance. This issue has become especially difficult in federal courts since the Supreme Court's decision in *Daubert*.[59] Because legal standards and scientific opinion are changing, judges are likely to benefit from information obtained from counsel on recent developments concerning estimation of the probability of a coincidental match.

What follows is a description of two techniques for estimating the probability of a coincidental match: the *product rule technique* and the *modified ceiling principle technique*, which was recommended by the NRC committee. The probability estimates can vary widely depending on the technique used.

1. Product rule technique

The product rule technique offers the most straightforward method of computing the probability of a matching DNA profile. To compute the probability of a random occurrence of a specific pattern of alleles in a DNA profile, the analyst multiplies the separate estimated probabilities of a random occurrence of each allele in the comparison population. When these individual probabilities are multiplied, the estimated probability of a distinctive pattern occurring at random may be less than one in several hundred thousand.[60]

The probability estimate resulting from the multiplication assumes that the individual alleles identified by genetic probes are *independent* of each other.[61] If the probabilities of the individual alleles are not independent (i.e., if certain alleles are likely to occur together in a person), multiplying the individual allele frequencies may underestimate or overestimate the true probability of matching alleles in the chosen population and thereby misstate the incriminating value of the evidence. Critics of the product rule technique contend that in some ethnic

58. The probability of a genotype for two distinct alleles at a single locus is $2p_1p_2$, where p_1 and p_2 are the relative frequencies of the alleles at a locus. This computation assumes that p_1 and p_2 are independent, an issue discussed *infra* note 61. Using the example above, the probability of a coincidental match of a genotype represented by alleles falling in Bin 20 and Bin 7 would be 2(0.014)(0.087) or 0.0024, or 24 out of 10,000, or 1 out of 417. When only a single band appears in a lane, most forensic laboratories will estimate the probability of the genotype as $2p_1$ (rather than p_1^2) to leave open the possibility of a second undetected allele. *See* NRC Report, *supra* note 7, at 78.

59. 113 S. Ct. 2786 (1993).

60. For an example of such an analysis, see OTA Report, *supra* note 12, at 67.

61. Two alleles will be independent if the occurrence of one allele is unrelated to the occurrence of another, much as tossing a pair of honest dice will yield two numbers that are independent.

subpopulations the alleles identified by commonly used genetic probes are not independent and that using a broad-based comparison population is therefore inappropriate. They prefer estimation techniques that do not require analysts to assume the *independence* of individual alleles in large comparison populations.[62]

2. Modified ceiling principle technique

The method recommended by the NRC committee involves conservative interpretations of existing population data.[63] As a preliminary test, the laboratory should examine its population database to determine if it contains a sample that matches the profile of the multiple alleles of the crime sample. If no match is found across multiple genetic probes, the expert reports that "the DNA pattern was compared to a database containing *N* individuals from the population and no match was observed."[64] This is an extremely conservative approach that yields probative information but does not give the fact finder any information concerning the probability of finding a matching profile by chance in the larger population.

The NRC committee has proposed a modified ceiling principle technique, which takes advantage of the computation techniques of population genetics but which includes adjustments that the committee believes make it an "appropriately conservative" approach. Although it noted that recent empirical studies have detected no evidence of a departure from independence within or across commonly used genetic probes,[65] the NRC committee chose "to assume for the sake of discussion that *population substructure* may exist and provide a

62. This issue is discussed in greater detail *infra* § VII.C.2.

63. The procedure described herein is the NRC committee's recommended interim solution for the computation of conservative estimates of a coincidental match of a specific DNA profile. The NRC report also calls for research that will assess the existence of *population substructures* by examining the frequency of alleles used for forensic identification in 100-person samples from fifteen to twenty genetically homogeneous ethnic populations in various regions of the world. This information would be combined in a reported frequency that represents the maximum possible likelihood of a coincidental match for any possible heritage, making it unnecessary to specify a comparison group based on an ethnic subpopulation. This technique is discussed in detail in the NRC Report, *supra* note 7, at 80–85. We present the modified ceiling principle interim technique because the more refined technique must await the findings of the proposed research on ethnic subpopulations.

64. NRC Report, *supra* note 7, at 91. The simplicity of this counting technique is offset by its strong conservative bias; the estimate of the likelihood of a match is bounded by the number of persons in the data set, which rarely exceeds a few hundred. Furthermore, the counting technique fails to take advantage of information from population genetics regarding the rarity of a match across multiple alleles. *See* Neil J. Risch & B. Devlin, *On the Probability of Matching DNA Fingerprints,* 255 Science 717, 720 (1992) (concluding that the counting rule is "unnecessarily conservative"). For these reasons, the NRC committee found the counting technique, when used to the exclusion of other methods, to be too conservative for use in estimating the frequency of particular DNA patterns. NRC Report, *supra* note 7, at 76.

65. *See also* NRC Report, *supra* note 7, at 80 ("Recent empirical studies concerning VNTR loci detected no deviation from independence within or across loci." (footnotes & citation omitted)); Richard Lempert, *DNA, Science and the Law: Two Cheers for the Ceiling Principle,* 34 Jurimetrics J. 41, 45–46 ("most empirical studies fail to find any forensically significant departures from Hardy-Weinberg or linkage equilibria within general populations" (footnote citing studies omitted)).

method for estimating population frequencies in a manner that adequately accounts for it."[66]

The NRC committee recommends that two adjustments be made to population frequencies derived from existing data to permit conservative estimates of the likelihood of a coincidental matching profile. First, the 95% upper confidence limit for the estimated allele frequency is computed for each of the existing population samples (Black, White, Native American, etc.). This upper bound of the *confidence interval* is intended to accommodate the uncertainties in current population sampling.[67] Second, the largest of these upper confidence limit estimates, or 10%, whichever is greater, is used to compute the joint probability of a coincidental match on the DNA profile of the crime sample. A lower bound of 10% is intended to address concerns that current population data sets may be substructured in unknown ways that would yield misleading estimates of a coincidental matching profile for members of subpopulations. The resulting probabilities then are multiplied as in the product rule computations.[68] The upper confidence limit from among the existing samples, or a 10% lower bound, is used to provide a probability estimate that errs, if at all, in a conservative direction (i.e., that is more favorable to a suspect).[69]

66. NRC Report, *supra* note 7, at 80. The NRC committee noted four considerations supporting this policy choice:

> (1) It is possible to provide conservative estimates of population frequency, without giving up the inherent power of DNA typing.
>
> (2) It is appropriate to prefer somewhat conservative numbers for forensic DNA typing, especially because the statistical power lost in this way can often be recovered through typing of additional loci, where required.
>
> (3) It is important to have a general approach that is applicable to any loci used for forensic typing. Recent empirical studies pertain only to the population genetics of the VNTR loci in current use. However, we expect forensic DNA typing to undergo much change over the next decade—including the introduction of different types of DNA polymorphisms, some of which might have different properties from the standpoint of population genetics.
>
> (4) It is desirable to provide a method for calculating population frequencies that is independent of the ethnic group of the subject. *Id*.

The propriety of such policy considerations in shaping the recommendations of the NRC committee has been questioned, as have been some of their scientific underpinnings. *See* Lempert, *supra* note 65 (acknowledging validity of some criticisms of the NRC recommendations but concluding that using the ceiling principle is preferable to the way random match probabilities were presented before the NRC report); William C. Thompson, *Evaluating the Admissibility of New Genetic Identification Tests: Lessons from the "DNA War,"* 84 J. Crim. L. & Criminology 22, 80 (1993) ("A major theme of critics from all perspectives is that the 'ceiling principle' is not a principle of science. It is an arbitrary policy statement, and can be accepted or rejected only as such." (footnote omitted)); David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts*, 7 Harv. J.L. & Tech. 101, 147 n.189 (1993) ("Plainly, the NRC panel's desire for a single method of calculating an upper bound on genotype frequencies in any likely population or subpopulation is not a pronouncement about science, but a mere preference for one jurisprudential policy over another.").

67. NRC Report, *supra* note 7, at 92.

68. This approach is discussed in greater detail in the NRC Report, *supra* note 7, at 91–92.

69. NRC Report, *supra* note 7, at 82–85.

C. What Is the Relevant Comparison Population for Estimating the Probability of a Coincidental Allelic Match?

Disputes over the appropriate comparison population have focused on cases in which the product rule technique has been used.[70] The extent to which the product rule technique may underestimate the probability of a coincidental match has been hotly disputed by population geneticists and other scholars.[71] The dispute centers on disagreements over the adequacy of commonly used comparison populations and the role of racial and ethnic subpopulations in probability estimation. Specifying the appropriate comparison population may be of considerable importance, as the estimates of a coincidental match can vary greatly depending on the population selected. For example, the prevalence of certain alleles may vary greatly across races—some alleles are common in Black populations and infrequent in White populations, and vice versa. If a DNA profile for a Black suspect is compared with frequency estimates based on a White population, the estimated likelihood of a chance match may be in error by some unknown amount. Similarly, if a DNA profile of a member of a subpopulation with a distinct frequency distribution of alleles is compared with frequency estimates based on an inappropriate larger population, an error of unknown magnitude may result. Concern over accuracy of estimates of a coincidental match has focused attention on the assumptions used in selecting a comparison population and the scientific validity of the methods used to estimate the probability of a coincidental match.

1. Is the comparison population consistent with the population of possible sources of the DNA?

If the modified ceiling principle technique is not used, an appropriate comparison population must be designated. Such a designation is guided by the characteristics of the population of individuals who might have been the source of the sample.[72] For example, if a rape victim saw her assailant and described him as White, and there is no more information to implicate a member of a specific subpopulation or group, the comparison population should be those who appear

70. Although the modified ceiling principle technique was intended to eliminate the necessity of specifying a comparison group, a dispute has emerged over which databases are to be considered when estimating the ceiling frequency for alleles (i.e., all databases or only those for major population groups). Thompson, *supra* note 66, at 80–81. This issue involves differing interpretations of the NRC committee's intended definition of the modified ceiling principle rather than a dispute over issues of science.

71. *See, e.g.,* R. C. Lewontin & Daniel L. Hartl, *Population Genetics in Forensic DNA Typing*, 254 Science 1745 (1991); Chakraborty & Kidd, *supra* note 56; Eric S. Lander, *Invited Editorial: Research on DNA Typing Catching Up with Courtroom Application*, 48 Am. J. Hum. Genet. 819, 821 (1991); Risch & Devlin, *supra* note 64. This debate is summarized in the NRC Report, *supra* note 7, at 79–80; Thompson, *supra* note 66; and Kaye, *supra* note 66.

72. NRC Report, *supra* note 7, at 85 ("[F]requencies should properly be based on the population of possible perpetrators, rather than on the population to which a particular suspect belongs." (footnotes omitted)). *See also* Lempert, *supra* note 45, at 310; and Kaye, *supra* note 66, at 137–38 & n.155.

to be White and who were in a position to commit the assault.[73] Comparisons based on members of populations or subpopulations who appear to be non-White therefore would be inappropriate. Similarly, where there is no information indicating the race or ethnicity of the perpetrator, the comparison population should be designated by the characteristics of those in a position to commit the assault. The race or ethnicity of the suspect is irrelevant.[74]

In other cases, however, the pool of alternative suspects may be limited to members of a distinct isolated community or a specific ethnic subgroup. This circumstance has arisen in federal courts where the defendant was a member of a Native American tribe and the crime occurred on the tribal reservation.[75] The typical databases of allele frequencies used for probability estimation address broader population groups, such as Blacks, Whites, Native Americans, Hispanics from the southeastern United States, and Hispanics from the southwestern United States. The extent to which broad racial and cultural comparison populations must correspond to the characteristics of the suspect population turns on the extent to which the distribution of alleles tested for in the forensic analysis differs for the suspect population and the comparison database. This remains a disputed issue among scientists. Difficulty in resolving this issue was responsible, in part, for the NRC committee's recommending a technique that does not require the designation of a suspect population.

2. Does the comparison population conform to characteristics that allow the estimation of the joint occurrence of matching alleles by multiplication of the probabilities of the individual alleles?

The comparison population must conform to assumptions that underlie the technique used to compute the estimate of a coincidental match, or at least conform sufficiently that minor deviations are of little consequence in computing the probability estimates. The computation of the probability of a random match by the product rule technique is based on the assumption that the individual alleles of the DNA profile are independent of one another. According to the principles of population genetics, the independence of alleles may be assumed only where the comparison population mixes freely and mates randomly (with respect to the alleles) such that the distribution of alleles within the comparison population is homogeneous.[76] If the comparison population does not conform to

73. *See, e.g.,* United States v. Jakobetz, 955 F.2d 786 (2d Cir.) (rape committed at a rest stop along an interstate highway by person identified only as White), *cert. denied*, 113 S. Ct. 104 (1992).

74. *See* Kaye, *supra* note 66, at 137–38 & n.155.

75. *E.g.*, United States v. Two Bulls, 918 F.2d 56 (8th Cir. 1990), *and vacated*, 925 F.2d 1127 (8th Cir. 1991); United States v. Martinez, 3 F.3d 1191 (8th Cir. 1993), *cert. denied*, 114 S. Ct. 734 (1994).

76. A comparison population whose members freely mix and randomly mate with respect to the relevant genes, resulting in a homogeneous distribution of alleles within the population, is said to be in *Hardy-Weinberg equilibrium*. Fleming, *supra* note 49, at 322. If the distribution of alleles within the population is not homogeneous, such a state may not exist, and the probe-detected alleles may not be independent, thereby violating an assumption of the product rule technique. *Id.* The assumptions of the product rule technique may also be violated if the detected alleles are in such proximity that they tend to be inherited together, resulting in

these assumptions, the alleles may not be independent, and the computation of probability estimates may be incorrect.

Opponents of the product rule technique argue that it is inappropriate to use broad racial and cultural characteristics to specify the comparison population, because in reality such groups do not mix freely with respect to relevant genes. Broad racial groups, they argue, disguise subpopulations that would be more appropriate for comparison. More specifically, opponents charge that use of broad racial groups violates the assumption of independence justifying the multiplication of the separate probabilities assigned to each probe. For example, a White comparison group that includes diverse ethnic groups (e.g., persons of Polish, Italian, or Irish descent) may mask differences in the distribution of alleles among subgroups. Opponents charge that no meaningful estimate of the probability of a particular DNA profile can be developed without specifying a suitable subpopulation that meets the demanding assumptions of the product rule technique.[77] When alleles are not independent, as when a comparison group contains a substructure, the product rule technique may underestimate the probability that the forensic and suspect DNA patterns match by coincidence.

Proponents of the product rule technique acknowledge that some substructuring may exist in the comparison populations typically used; but they argue that it is inappropriate to apply the assumption strictly and that the probability estimates generally are accurate in spite of violations of the strict assumption regarding the absence of substructuring.[78] Furthermore, proponents claim that the conservative features of the fixed-bin method more than compensate for any underestimation.[79] The NRC committee notes that what little empirical evidence existed at the time of its report appeared to support this contention.[80]

At the time of this writing, two federal courts of appeals have approved the admission of probability estimates based on the product rule technique, and one court of appeals has approved a variation of the product rule technique with a conservative adjustment for substructuring. Before the *Daubert* decision, the

linkage disequilibrium. *Id.* The latter problem often is avoided by probing for alleles of genes known to be on different chromosomes. *Id.*

77. Eric S. Lander, *DNA Fingerprinting on Trial*, 339 Nature 501 (1989); Lewontin & Hartl, *supra* note 71, at 1746.

78. Chakraborty & Kidd, *supra* note 56, at 1736; OTA Report, *supra* note 12, at 68 ("Consensus exists that genetic departures as extreme as those for rare disease alleles do not exist for alleles detected by forensic DNA probes.").

79. Chakraborty & Kidd, *supra* note 56, at 1738. For a discussion of the fixed-bin method, see *supra* § VII.A.

80. NRC Report, *supra* note 7, at 80:

> Recent empirical studies concerning VNTR loci detected no deviation from independence within or across loci. Moreover, pairwise comparisons of all five-locus DNA profiles in the FBI database showed no exact matches; the closest match was a single three-locus match among 7.6 million pairwise comparisons. These studies are interpreted as indicating that multiplication of gene frequencies across loci does not lead to major inaccuracies in the calculation of genotype frequency—at least not for the specific polymorphic loci examined. (citations omitted)

See also Lempert, *supra* note 65, at 45–46.

U.S. Court of Appeals for the Second Circuit upheld the admission of probability estimates under Rule 702 of the Federal Rules of Evidence.[81] Soon after the *Daubert* decision, the U.S. Court of Appeals for the Sixth Circuit held that disputes over probability estimation techniques go to weight, not admissibility.[82] The district court had found probability estimates based on the product rule technique admissible under *Frye*'s test of general acceptance within the scientific community. The Sixth Circuit reinterpreted these findings under the *Daubert* standards and concluded, "[I]t is clear from this record that the DNA evidence and testimony would have met the more liberal Rule 702 test adopted by the Supreme Court."[83] The court noted that the theory and methods that support the FBI estimates can and have been tested and that they have received at least some degree of peer review and evaluation.[84] The court also found that the theory and methods were generally accepted by the relevant scientific community, or at least not disfavored by a substantial part of that community.[85] The court noted that the dispute over ethnic substructure goes to the accuracy of the probability estimates and thus to the weight of the evidence, not its admissibility.[86] Because the NRC report was not part of the trial record, the Sixth Circuit did not consider it or the merits of the modified ceiling principle technique. However, the court noted that the substance of the criticisms presented in the NRC report, including the possibility of ethnic substructure, was before the court in the form of expert testimony.[87]

The U.S. Court of Appeals for the Ninth Circuit upheld the admission of probability estimates computed in a way that takes into account some of the

81. United States v. Jakobetz, 955 F.2d 786, 798–800 (2d Cir.) (no abuse of discretion where evidence was admitted under relevancy standard prior to *Daubert* and the trial judge concluded that conservative estimate arising from FBI's fixed-bin method compensates for any departure from assumptions of the analysis arising from a possible substructure in the comparison population), *cert. denied*, 113 S. Ct. 104 (1992).

82. United States v. Bonds, 12 F.3d 540, 564 (6th Cir. 1993), *aff'g* United States v. Yee, 134 F.R.D. 161 (N.D. Ohio 1991) ("This substructure argument involves a dispute over the accuracy of the probability results, and thus this criticism goes to the weight of the evidence, not its admissibility."). One additional federal court of appeals has considered but has not ruled on the admissibility of probability estimates under the product rule. An earlier decision by the U.S. Court of Appeals for the Eighth Circuit was vacated and set for *en banc* review, but the appeal was dismissed following the death of the appellant. United States v. Two Bulls, 918 F.2d 56 (8th Cir. 1990), *and vacated*, 925 F.2d 1127 (8th Cir. 1991). In a more recent decision, the Eighth Circuit was not required to decide whether it was error to admit evidence of a profile match without an accompanying probability estimate because the defendant had asked the court to exclude all statistical evidence of the probability of a match. United States v. Martinez, 3 F.3d 1191, 1198–99 (8th Cir. 1993), *cert. denied*, 114 S. Ct. 734 (1994).

83. *Bonds*, 12 F.3d at 557.

84. *Id.* at 558–60. The court acknowledged concern over the failure to conduct blind proficiency tests and the lack of specific information concerning error rates, but it noted that error rate is but one of the factors that bear on admissibility of scientific evidence. *Id.* at 560.

85. *Id.* at 562 ("[E]ven substantial criticism as to one theory or procedure will not be enough to find that the theory/procedure is not generally accepted. Only when a theory or procedure does not have the acceptance of most of the pertinent scientific community, and in fact a substantial part of the scientific community disfavors the principle or procedure, will it not be generally accepted.").

86. *Id.* at 564 (The potential of ethnic substructure does not mean that the theory and procedures used by the FBI are not generally accepted; it means only that there is a dispute over whether the results are as accurate as they might be and what, if any, weight the jury should give those results.).

87. *Id.* at 552–53.

concerns expressed in the NRC report.[88] The defendant, a member of the Navajo tribe, challenged his conviction for rape and murder, contending, among other things, that the existing comparison databases underrepresent the members of his tribe.[89] Because members of his tribe form an isolated subgroup that may have a distinct genetic substructure, he contended that the probability estimates derived from existing databases would likely understate the probability of a random match between the forensic sample and members of his tribe. In the absence of a database with appropriate representation of members of the Navajo tribe, the prosecution used a technique similar to that used in the modified ceiling principle technique—allele frequencies were examined for several Native American tribal populations and the estimates were based on the highest frequency.[90]

The court found the probability estimates generated by a variation of the modified ceiling principle technique satisfied the standard of Rule 702 as interpreted in *Daubert,* and it went on to assess whether the probative value of such evidence was outweighed by the potential for prejudice under Rule 403.[91] Of particular concern was the possibility "that the jury will accept the DNA evidence as a statement of source probability (i.e., the likelihood that the defendant is the source of the evidentiary sample)"[92] rather than as an estimate of the rarity

88. United States v. Chischilly, 30 F.3d 1144 (9th Cir. 1994).

89. *Id.* at 1155. The crime occurred in an isolated area of the Navajo tribal reservation, and people with an opportunity to commit the crime were members of the tribe.

90. The details of this technique are not reported in the opinion. A footnote quotes one of the government's witnesses as follows: "[T]hey looked at the allele frequencies in several American Indian segment tribal populations and picked up the one containing largest frequency (sic)." The court noted:

> While not calculated pursuant to the NRC Report's controversial recommendation to adopt the ceiling principle, the one in 2563 probability that was introduced at trial was nonetheless arguably calculated on the basis of somewhat conservative statistical assumptions, was premised on the favorable assumption that the source of the sperm was a Native American and was emphasized at the expense of a much smaller probability of a random match that Government witnesses testified would be statistically defensible. (footnotes omitted).

Id. at 1158.

91. The court noted:

> Notwithstanding *Daubert's* express preference for exposing novel scientific theories and methodologies to the glare of the adversarial process, *Daubert* enjoins watchful assessment of the risk that a jury would assign undue weight to DNA profiling statistics even after hearing appellant's opposing evidence, the testimony of Government witnesses under vigorous cross-examination and the careful instructions of the district court on burdens of proof. Of particular concern is where the Government seeks to present probability testimony derived from statistical analysis, the third main phase of DNA profiling. Numerous hazards attend the courtroom presentation of statistical evidence of any sort. Accordingly, Rule 403 requires judicial vigilance against the risk that such evidence will inordinately distract the jury from or skew its perception of other, potentially exculpatory evidence lacking not so much probative force as scientific gloss. (footnotes omitted).

Id. at 1156.

92. *Id.* at 1156. In support of its concern, the court cited Lempert, *supra* note 45, at 306 ("Unfortunately, the careless presentation of evidence . . . may make it look as if the question of the rareness of the evidence DNA profile and the probability that the defendant's matching DNA is the source of the evidence profile are identical."); and Janet C. Hoeffel, *Note, The Dark Side of DNA Profiling: Unreliable Scientific Evidence Meets the Criminal Defendant,* 42 Stan. L. Rev. 465, 515 (1990) (asserting that "juries often erroneously equate the

of the DNA profile. Estimating the probability that the defendant was the source of the crime sample requires not only an estimate of a coincidental match based on a comparison population but consideration of other evidence bearing on guilt or innocence. In addition, the court indicated concern that even when the source probability is correctly assessed, jurors might equate such probability estimates with guilt, thereby ignoring the possibility that there may be a noncriminal explanation of how the suspect could have been the source of the evidentiary sample. The solution to such potential problems, the court concluded, is careful oversight by the district court to limit the opportunity for misrepresentation of statistical evidence. The court of appeals approved the admission of the estimates, holding that the district court exercised such control by presenting statistical estimates as the probability of a random match, not the probability of the defendant's innocence, and by admitting probability estimates based on conservative statistical assumptions.[93]

As a result, federal courts are at odds with a number of state courts that have considered explicitly the NRC report, cited the modified ceiling principle technique with favor, and expressed the hope that it will gain general acceptance in the scientific community.[94] Unlike the federal courts that have addressed the issue, many of the state courts continue to follow variations of the *Frye* standard

frequency of the accused's blood type in the population with the probability of innocence, discounting the other evidence in the case pointing to guilt or innocence, such as the fact that a close relative is also suspected"). *See also* William C. Thompson & Edward L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials, The Prosecutor's Fallacy and the Defense Attorney's Fallacy,* 11 Law & Hum. Behav. 167, 170–71 (1987); Jonathan J. Koehler, *Error and Exaggeration in the Presentation of DNA Evidence at Trial,* 34 Jurimetrics J. 21 (1993).

93. United States v. Chischilly, 30 F.3d at 1158.

94. Commonwealth v. Lanigan, 596 N.E.2d 311, 316 (Mass. 1992) ("This [ceiling] principle is entirely in keeping with the hope that we expressed in Commonwealth v. Curnin, that the scientific community would 'generally agree on a means of arriving at a conservative estimate of the probability of another person having the same alleles and thus resolve all uncertainties and variables in favor of the defense.'" (citation omitted)); People v. Barney, 8 Cal. App. 4th 798, 821 (Cal. Ct. App. 1992) ("There must be some common ground, some sufficiently conservative method of determining statistical significance, as to which there is general scientific agreement. . . . The NRC report on DNA analysis appears to point the way to such common ground." (citations omitted)); State v. Vandebogart, 616 A.2d 483, 494 (N.H. 1992) ("The NRC asserts that the ceiling principle can account for any error caused by possible population substructure. Therefore, the admissibility of population frequency estimates do not necessarily await resolution of the population substructure issue, as long as the relevant scientific community generally accepts a method for calculating statistical probabilities."). Vargas v. State, Nos. 92-556, 92-557, 92-558 (consolidated), 1994 WL 231360, *13 (Fla. App. 1 Dist. June 1, 1994) ("The discussion of the modified ceiling principle in the cases appears to confirm that a more conservative calculation may be possible, which would be generally accepted in the relevant scientific community, . . ."); People v. Watson, 629 N.E.2d 634, 647–48 (Ill. App. 1994) ("Because the match of DNA patterns is a matter of substantial significance and because this case has the potential for becoming a significant precedent in this jurisdiction, we believe the trial court should be given the opportunity to determine whether the recently promulgated ceiling principle is appropriate under *Frye* for calculating the probability estimate to be applied to a match declaration in the present case. Accordingly, we remand this cause to the trial court for such a determination. At least in our view, the NRC Report, which was not previously available to the trial court, suggests that the DNA evidence should be admitted on the basis of this more conservative probability calculation for which the requisite consensus may now exist."); State v. Bloom, 516 N.W.2d 159, 160 (Minn. 1994) (testimony regarding statistical probability of a matching DNA profile derived by using the "interim ceiling method" admissible as an exception to Minnesota's rule against admitting statistical probability evidence on the issue of identity in criminal prosecutions).

and look to the NRC report as expressing a consensus of scientific opinion. Yet, while some state courts have embraced the modified ceiling principle technique, a portion of the scientific community has questioned its scientific validity.[95] The possibility of a convergence of scientific opinion around the ceiling principle technique seems remote, and the NRC plans to impanel another committee to consider scientific criticisms of the technique and more recent research.[96]

95. B. Devlin et al., *Statistical Evaluation of DNA Fingerprinting: A Critique of the NRC's Report*, 259 Science 748, 749 (1993); Peter Aldhous, *Geneticists Attack NRC Report As Scientifically Flawed*, 259 Science 755 (1993); Joel E. Cohen, *The Ceiling Principle Is Not Always Conservative in Assigning Genotype Frequencies for Forensic DNA Testing*, 51 Am. J. Hum. Genet. 1165, 1166 (1992) (letter to the editor); B. Devlin et al., *Comments on the Statistical Aspects of the NRC's Report on DNA Typing*, 39 J. Forensic Sci. 28 (1994). *See also* People v. Wallace, 14 Cal. App. 4th 651, 660 (Cal. Ct. App. 1993):

> [R]ecent developments have shown that general acceptance may not be easily achieved. It appears that some proponents of DNA analysis, rather than attempting to come to terms with the NRC report or some other compromise on statistical calculation, have taken the offensive and attacked the report's proposed new methods of statistical calculation as unsound.

96. Rorie Sherman, *New Scrutiny for DNA Testing,* Nat'l L.J., Oct. 18, 1993, at 3.

Appendix A: Schematic of Single-Locus Probe RFLP Analysis

Appendix B: Example of Illustrated DNA Patterns on Autorad

Autorad courtesy of Cellmark Diagnostics.

The autorad on the preceding page depicts DNA evidence from a criminal case in which blood stains on the pants and shirt of the defendant yielded DNA that was compared with the DNA samples known to be from the victim and the defendant. The autorad depicts DNA patterns in ten parallel lanes. The images in the lanes represent the following sources:

The lanes labeled "λ," "1kb," and "TS" show control samples of DNA. These serve as quality-control checks.

The lane labeled "D" shows the pattern obtained from a known DNA sample obtained from the defendant.

The lane labeled "jeans" shows the print of DNA from blood stains on the defendant's pants.

The two lanes labeled "shirt" show the prints of DNA from blood stains on the defendant's shirt.

The lane labeled "V" is a print of the DNA from a blood sample known to be from the victim.

The DNA from the blood stains on the defendant's clothing (lanes labeled "jeans" and "shirt") do not match his own blood sample ("D") but do match that of the victim ("V").

Appendix C: TWGDAM Guidelines

The following excerpts were taken from Technical Working Group on DNA Analysis Methods & California Ass'n of Criminalists Ad Hoc Comm. on DNA Quality Assurance, *Guidelines for a Quality Assurance Program for DNA Analysis*, 18 Crime Laboratory Dig. 44, 52–56, 60–62 (1991).

3. Documentation

The DNA laboratory must maintain documentation on all significant aspects of the DNA analysis procedure, as well as any related documents or laboratory records that are pertinent to the analysis or interpretation of results, so as to create a traceable audit trail. This documentation will serve as an archive for retrospective scientific inspection, reevaluation of the data, and reconstruction of the DNA procedure. Documentation must exist for the following topic areas:

3.1 Test Methods and Procedures for DNA Typing

This document must describe in detail the protocol currently used for the analytical testing of DNA. This protocol must identify the standards and controls required, the date the procedure was adopted and the authorization for its use. Revisions must be clearly documented and appropriately authorized.

3.2 Population Data Base to include number, source and ethnic and/or racial classification of samples.

3.3 Quality control of critical reagents (such as commercial supplies and kits which have expiration dates) to include lot and batch numbers, manufacturer's specifications and internal evaluations.

3.4 Case files/case notes - Must provide foundation for results and conclusions contained in formal report.

3.5 Data analysis and reporting

3.6 Evidence handling protocols

3.7 Equipment calibration and maintenance logs

3.8 Proficiency testing

3.9 Personnel training and qualification records

3.10 Method validation records

3.11 Quality assurance and audit records

3.12 Quality assurance manual

3.13 Equipment inventory
3.14 Safety manuals
3.15 Material safety data sheets
3.16 Historical or archival records
3.17 Licenses and certificates

4. Validation

4.1 General Considerations for Developmental Validation of the DNA Analysis Procedure

4.1.1 Validation is the process used by the scientific community to acquire the necessary information to assess the ability of a procedure to reliably obtain a desired result, determine the conditions under which such results can be obtained and determine the limitations of the procedure. The validation process identifies the critical aspects of a procedure which must be carefully controlled and monitored.

4.1.2 Validation studies must have been conducted by the DNA laboratory or scientific community prior to the adoption of a procedure by the DNA laboratory.

4.1.3 Each locus to be used must go through the necessary validation.

4.1.4 The DNA primers, probe(s) or oligonucleotides selected for use in the forensic DNA analysis must be readily available to the scientific community.

4.1.5 The validation process should include the following studies (Report of a Symposium on the Practice of Forensic Serology 1987, and Budowle *et al*. 1988):

4.1.5.1 Standard Specimens - The typing procedure should have been evaluated using fresh body tissues, and fluids obtained and stored in a controlled manner. DNA isolated from different tissues from the same individual should yield the same type.

4.1.5.2 Consistency - Using specimens obtained from donors of known type, evaluate the reproducibility of the technique both within the laboratory and among different laboratories.

4.1.5.3 Population Studies - Establish population distribution data in different racial and/or ethnic groups.

4.1.5.4 Reproducibility - Prepare dried stains using body fluids from donors of known types and analyze to ensure that the strain specimens exhibit accurate, interpretable and reproducible DNA types of profiles that match those obtained on liquid specimens.

4.1.5.5 Mixed Specimen Studies - Investigate the ability of the system to detect the components of mixed specimens and define the limitations of the system.

4.1.5.6 Environmental Studies - Evaluate the method using known or previously characterized samples exposed to a variety of environmental conditions. The samples should be selected to represent the types of specimens to be routinely analyzed by the method. They should resemble actual evidence materials as closely as possible so that the effects of factors such as matrix, age and degradative environment (temperature, humidity, UV) of a sample are considered.

4.1.5.7 Matrix Studies - Examine prepared body fluids mixed with a variety of commonly encountered substances (e.g. dyes, soil) and deposited on commonly encountered substrates (e.g. leather, denim).

4.1.5.8 Nonprobative Evidence - Examine DNA profiles in nonprobative evidentiary stain materials. Compare the DNA profiles obtained for the known liquid blood versus questioned blood deposited on typical crime scene evidence.

4.1.5.9 Nonhuman Studies - Determine if DNA typing methods designed for use with human specimens detect DNA profiles in nonhuman source stains.

4.1.5.10 Minimum Sample - Establish quantity of DNA needed to obtain a reliable typing result.

4.1.5.11 On-site Evaluation - Set up newly developed typing methods in the case working laboratory for on-site evaluation of the procedure.

4.1.5.12 It is essential that the results of the developmental validation studies be shared as soon as possible with the scientific community through presentations at scientific/professional meetings. It is imperative that details of these studies be available for peer review through timely publications in scientific journals.

4.2 Characterization of Loci

During the development of a DNA analysis system, basic characteristics of the loci must be determined and documented. (Baird 1989; AABB Standards Committee 1990.)

4.2.1 Inheritance - DNA loci used in forensic testing shall have been validated by family studies to demonstrate the mode of inheritance. Those DNA loci used in parentage testing should have a low frequency of mutation and/or recombination.

4.2.2 Gene Mapping - The chromosomal location of the polymorphic loci used for forensic testing shall be submitted to or recorded in the Yale Gene Library or the International Human Gene Mapping Workshop.

4.2.3 Detection - The molecular basis for detecting the polymorphic loci shall be documented in the scientific or technical literature.

4.2.3.1 For RFLP this includes the restriction enzyme and the probes used.

4.2.3.2 For PCR this includes the primers and probes if used.

4.2.4 Polymorphism - The type of polymorphism detected shall be known.

4.3 Specific Developmental Validation of RFLP Procedures

4.3.1 Restriction - The conditions and control(s) needed to ensure complete and specific restriction must be demonstrated.

4.3.2 Separation - Parameters for the reproducible separation of DNA fragments must be established.

4.3.3 Transfer - Parameters for the reproducible transfer of DNA fragments must be established.

4.3.4 Detection - The hybridization and stringency wash conditions necessary to provide the desired degree of specificity must be determined.

4.3.5 Sizing - The precision of the sizing procedure must be established.

4.4 Specific Developmental Validation of PCR Based DNA Procedures

4.4.1 Amplification

4.4.1.1 The PCR primers must be of known sequence.

4.4.1.2 Conditions and measures necessary to protect pre-amplification samples from contamination by post PCR materials should be determined (see Section 7.5).

4.4.1.3 The reaction conditions such as thermocycling parameters and critical reagent concentrations (primers, polymerase and salts) needed to provide the required degree of specificity must be determined.

4.4.1.4 The number(s) of cycles necessary to produce reliable results must be determined.

4.4.1.5 Potential for differential amplification must be assessed and addressed.

4.4.1.6 Where more than one locus is amplified in one sample mixture, the effects of such amplification on each system (alleles) must be addressed and documented.

4.4.2 Detection of PCR Product

The validation process will identify the panel of positive and negative controls needed for each assay described below.

4.4.2.1 Characterization without hybridization

When a PCR product is characterized directly, appropriate standards for assessing the alleles shall be established (e.g., size markers).

4.4.2.2 Characterization with hybridization

(a) Hybridization and stringency wash conditions necessary to provide the desired degree of specificity must be determined.

(b) For assays in which the amplified target DNA is to be bound directly to a membrane, some mechanism should be employed to ensure that the DNA has been applied to the membrane.

(c) For assays in which the probe is bound to the membrane, some mechanism should be employed to show that adequate amplified DNA is present in the sample (e.g., a probe which reacts with any amplified allele or a product yield gel).

4.5 Internal Validation of Established Procedures (ASCLD 1986)

Prior to implementing a new DNA analysis procedure, or an existing DNA procedure developed by another laboratory that meets the developmental criteria described under Section 4.1, the forensic laboratory must first demonstrate the reliability of the procedure in-house. This internal validation must include the following:

4.5.1 The method must be tested using known samples.

4.5.2 If a modification which materially effects the results of an analysis has been made to an analytical procedure, the modified procedure must be compared to the original using identical samples.

4.5.3 Precision (e.g., measurement of fragment lengths) must be determined by repetitive analyses to establish criteria for matching.

4.5.4 The laboratory must demonstrate that its procedures do not introduce contamination which would lead to errors in typing.

4.5.5 The method must be tested using proficiency test samples. The proficiency test may be administered internally, externally or collaboratively.

7. Analytical Procedures

7.1 Sample Evaluation and Preparation

7.1.1 General characterization of the biological material should be performed prior to DNA analysis. Evidence samples submitted should be evaluated to determine the appropriateness for DNA analysis.

7.1.2 When semen is identified, a method of differential extraction should be employed, and when appropriate, each of the DNA fractions typed (see Section 4.1.5.10).

7.1.3 Testing of evidence and evidence samples should be conducted to provide the maximum information with the least consumption of the sample. Whenever possible, a portion of the original sample should be retained or returned to the submitting agency as established by laboratory policy.

7.2 DNA Isolation

7.2.1 The DNA isolation procedure should protect against sample contamination.

7.2.2 The effectiveness of the DNA isolation procedure should be evaluated by regular use of an appropriate cellular source of human DNA.

7.3 Procedures for Estimating DNA Recovery:

A procedure should be used for estimating the quality (extent of DNA degradation) and quantity of DNA recovered from the specimens. One or more of the following procedures may be employed to evaluate the effectiveness of the DNA recovery.

7.3.1 Yield Gel - Yield gels must include a set of high molecular weight DNA calibration standards for quantitative estimate of yield.

7.3.2 UV Absorbance - Absorbance and wavelength standards or a high molecular weight DNA calibration standard may be used.

7.3.3 Fluorescence - Approximate quantification of extracted DNA can be accomplished by comparison with known concentrations of high molecular weight DNA.

7.3.4 Hybridization - Quantitation with human/primate specific probes requires an appropriate set of human DNA standards.

7.4 Analytical Procedures for RFLP Analysis

7.4.1 Restriction Enzymes

7.4.1.1 Prior to its initial use, each lot of restriction enzyme should be tested against an appropriate viral, human or other DNA standard which produces an expected DNA fragment pattern under standard digestion conditions. The restriction enzyme should also be tested under conditions that will reveal contaminating nuclease activity.

7.4.1.2 Demonstration of Restriction Enzyme Digestion - Digestion of extracted DNA by the restriction enzyme should be demonstrated using a test gel which includes:

(a) Size Marker - Determines approximate size range of digested DNA.

(b) Human DNA Control - Measures the effectiveness of restriction enzyme digestion of genomic human DNA.

7.4.2 Analytical Gel - The analytical gel used to separate restriction fragments must include the following:

7.4.2.1 Visual Marker - Visual or fluorescent markers which are used to determine the end point of electrophoresis.

7.4.2.2 Molecular Weight Size Markers - Markers which span the RFLP size range and are used to determine the size of unknown restriction fragments. Case samples must be bracketed by molecular weight size marker lanes.

7.4.2.3 Human DNA Control - A documented positive human DNA control of known type which produces a known fragment pattern with each probe and serves as a systems check for the following functions:

(a) electrophoresis quality and resolution
(b) sizing process
(c) probe identity
(d) hybridization efficiency
(e) stripping efficiency

7.4.2.4 A procedure should be available to interpret altered migration of DNA fragments.

7.4.3 Southern Blots/Hybridization - The efficiency of blotting, hybridizations and stringency washes are monitored by the human DNA control and size markers.

7.4.4 Autoradiography - The exposure intensity is monitored by the use of multiple X-ray films or by successive exposures in order to obtain films of the proper intensity for image analysis.

7.4.5 Image and Data Processing - The functioning of image and data processing is monitored by the human DNA control allelic values.

7.5 Analytical Procedures for PCR Based Techniques

7.5.1 Internal Controls and Standards

The laboratory's QC guidelines should contain specific protocols to assess critical parameters in normal operations which include the following:

7.5.1.1 Negative controls to be included with each sample set are: (a) a reagent blank and (b) an amplification blank.

7.5.1.2 A human DNA known type must be introduced at the amplification step as a positive control and carried through the remainder of the typing.

7.5.1.3 Where appropriate, substrate controls should be collected from the evidence (e.g., unstained areas adjacent to stained areas, hair shafts adjacent to hair roots) and should be processed at the same time as the evidence samples.

7.5.1.4 Where feasible, the sample should be split for duplicate analysis as early as possible prior to amplification.

7.5.1.5 To characterize amplified fragment length polymorphisms, markers which span the allele size range must be used. Case samples must be bracketed by marker lanes.

Glossary of Terms

Many of the following terms are defined specifically within the context of forensic typing of human DNA. Some of these terms have broader or slightly different meanings in other molecular biology applications. These terms and definitions were adapted in part from the following: Committee on DNA Technology in Forensic Science, National Research Council, DNA Technology in Forensic Science (1992); Office of Technology Assessment, U.S. Congress, Genetic Witness: Forensic Uses of DNA Tests (1990); Lorne T. Kirby, DNA Fingerprinting: An Introduction (1990).

Adenine (A). One of the four bases, or nucleotides, that make up the DNA double helix. Adenine only binds to thymine. See Nucleotide.

Allele. An alternative form of a gene or VNTR at a specific locus. Some genes have two variants (e.g., an allele for eye color may be blue or brown); others have more. Alleles are inherited separately from each parent. At the same loci on any two homologous chromosomes, an individual may have two different alleles (heterozygous) or the same allele (homozygous).

Amplified Fragment Length Polymorphism (AMP-FLP). A DNA identification technique that uses PCR-amplified DNA fragments of varying lengths based on VNTRs.

Autoradiograph (Autorad). In RFLP analysis, the x-ray film (or print thereof) showing the positions of radioactively marked lengths (bands) of DNA.

Autosome. Any chromosome other than the sex chromosomes X or Y. See Chromosome.

Band-Shift. Movement of DNA fragments in one lane of a gel at a different rate than fragments of identical length in another lane, resulting in the same pattern "shifted" up or down relative to the comparison lane. Band-shift does not necessarily occur at the same rate in all portions of the gel.

Band Size. Length of DNA fragment measured in base pairs. See Base Pair.

Base Pair (bp). Two complementary nucleotides bonded together at the matching bases (A and T or C and G) to form one segment of the DNA double helix. The length of a DNA fragment often is measured in numbers of base

pairs (1 kilobase (kb) =1000 bp); base pair numbers also are used to describe the location of an allele on the DNA strand.

Chromosome. A rod-like structure composed of DNA which carries part of the genome. Most normal human cells contain forty-six chromosomes, twenty-two autosomes and one sex chromosome (X) inherited from the mother, and twenty-two autosomes and one sex chromosome (either X or Y) inherited from the father.

Confidence Interval. Interval set around an allele frequency that accounts for uncertainty in measurement of the allele. The upper bound of this interval (the upper confidence limit) may be used as a conservative estimate of the frequency of the allele.

Cytosine (C). One of the four bases, or nucleotides, that make up the DNA double helix. Cytosine only binds to guanine. See Nucleotide.

Denature, Denaturation. The process of splitting, as by heating, two complementary strands of the DNA double helix into single strands in preparation for hybridization with biological probes.

Deoxyribonucleic Acid (DNA). Basic molecule of heredity. DNA is composed of nucleotide building blocks, each containing a base (A, C, G, or T), a phosphate, and a sugar. These nucleotides are linked together in a double helix—two strands of DNA molecules paired up at complementary bases (A with T, C with G). See Adenine, Cytosine, Guanine, Thymine.

DNA Profile. The pattern of band lengths on an autorad representing the combined results of multiple probes.

Double Helix. Ladder structure of DNA.

Environmental Insult. Exposure of DNA to external agents such as heat, moisture, and ultraviolet radiation, or chemical or bacterial agents. Such exposure can interfere with the enzymes used in the testing process, or otherwise make DNA difficult to analyze.

Fixed-Bin Method. In a fixed-bin method, preestablished bins are designated by certain absolute base pair ranges, and band sizes are sorted into these existing bins.

Floating Bin Method. In a floating bin method, the bin is centered on the base pair length of the allele in question, and the width of the bin is defined by the laboratory's matching rule (e.g., ± 2.5% of band size).

Forensic Sample. A sample of DNA associated with the commission of a crime, such as a DNA sample obtained from a blood stain at a crime scene. See Suspect Sample.

Gel Electrophoresis. In RFLP analysis, the process of sorting DNA fragments by size by applying an electric current to an agarose gel. The different-sized fragments move at different rates through the gel.

Gene. A distinctive, ordered sequence of nucleotide base pairs on a chromosome. The gene is the fundamental unit of heredity; each gene provides a "code" for a specific biological characteristic.

Genome. The complete genetic makeup of an organism, comprising 50,000–100,000 genes in humans. See Gene.

Genotype. The genetic code of an organism (as distinguished from phenotype, which refers to how the genetic code expresses itself, as in physical appearance).

Guanine (G). One of the four bases, or nucleotides, that make up the DNA double helix. Guanine only binds to cytosine. See Nucleotide.

Hardy-Weinberg Equilibrium. A condition in which the allele frequencies within a large, random, intrabreeding population are unrelated to patterns of mating. In this condition, the occurrence of alleles from each parent will be independent and have a joint frequency estimated by the product rule. See Independence, Linkage Disequilibrium.

Heterozygous. Having a different allele at a given locus on each of a pair of homologous chromosomes, one inherited from each parent. See Allele.

HLA DQ α. A particular class of Human Leukocyte Antigen (HLA) whose locus has been sequenced completely and thus can be used for forensic typing. See Human Leukocyte Antigen.

Homologous Chromosomes. The forty-four autosomes in the normal human genome are in homologous pairs that share an identical sequence of genes, but may have different alleles at the same loci.

Homozygous. Having the same allele at a given locus on each of a pair of homologous chromosomes, one inherited from each parent. See Allele.

Human Leukocyte Antigen (HLA). Antigen (foreign body that stimulates an immune system response) located on the surface of most cells (excluding red blood cells and sperm cells). HLAs differ among individuals and are associated closely with transplant rejection. See HLA DQ α.

Hybridize, Hybridization. Pairing up of complementary strands of DNA from different sources at the matching base pair sites. For example, a primer with the sequence AGGTCT would bond with the complementary sequence TCCAGA on a DNA fragment.

Independent, Independence. A condition in which the occurrence of alleles is uncorrelated, permitting accurate estimation of frequency of the joint occur-

rence by the product rule. In large populations, the presence of distinct subpopulations may violate this condition. See Hardy-Weinberg Equilibrium.

Linkage Disequilibrium. The nonrandom association of one allele at one locus and another allele at a different locus (i.e., they appear together with greater frequency than expected by chance). See Hardy-Weinberg Equilibrium.

Locus, Loci. A specific location or locations on a chromosome.

Match Window. Size of a bin for DNA fragment length measurement and comparison purposes.

Modified Ceiling Principle Technique. An adjustment to account for possible population substructure in statistical estimates of a coincidental match on a DNA profile. In multiplying allele frequencies, an expert would use the upper bound of the confidence interval for the subpopulation in which the allele frequency is most common. This ceiling frequency is chosen so that the allele frequencies used in the calculation exceed the allele frequency in any of the subgroups, providing theoretically more conservative (i.e., favoring the suspect) estimates of the likelihood of a matching DNA profile.

Monomorphic, Monomorphism. A single form of a genetic trait.

Monomorphic Probe. A probe that detects the same allele and hence the same pattern in everyone.

Multilocus Probe. A probe that marks multiple sites (loci). RFLP analysis using a multilocus probe will yield an autorad showing a striped pattern of thirty or more bands. Rarely used now in forensic applications in the United States.

Nucleotide. A unit of DNA consisting of a base (A, C, G, or T) and attached to a phosphate and a sugar group. See Deoxyribonucleic Acid.

Oligonucleotide. A synthetic polymer made up of fewer than 100 nucleotides; used as a primer or a probe in PCR. See Primer.

Polymerase Chain Reaction (PCR). A process that mimics DNA's own replication processes to make up to millions of copies of the original genetic material in a few hours.

Polymorphic, Polymorphism. Multiple forms of a genetic trait.

Population Genetics. The study of genetic composition of groups of individuals.

Population Substructure. In population genetics, the theory that allele frequencies are not spread randomly within large heterogeneous racial groups, but vary greatly between smaller ethnic subpopulations which do not mix freely.

Primer. An oligonucleotide that attaches to one end of a DNA fragment and provides a point for more complementary nucleotides to attach and replicate the DNA strand. See Oligonucleotide.

Probe. In forensics, a short segment of DNA used to detect certain alleles. The probe hybridizes, or matches up, to a specific complementary sequence. Probes allow visualization of the hybridized DNA, either by radioactive tag (usually used for RFLP analysis) or biochemical tag (usually used for HLA DQ α).

Product Rule Technique. Technique for calculating genotype frequencies by multiplying allele frequencies observed in a population database. Assumes that alleles are inherited independently.

Replication. The synthesis of new DNA from existing DNA. See Polymerase Chain Reaction.

Restriction Enzyme. Protein that cuts double-stranded DNA at specific base pair sequences (different enzymes recognize different sequences). See Restriction Site.

Restriction Fragment Length Polymorphism (RFLP) Analysis. Analysis of individual variations in the lengths of DNA fragments produced by digesting sample DNA with a restriction enzyme.

Restriction Site. A sequence marking the location at which a restriction enzyme cuts DNA into fragments. See Restriction Enzyme.

Sequence-Specific Oligonucleotide (SSO) Probes. Also, Allele-Specific Oligonucleotide (ASO) Probes. Oligonucleotide probes used in a PCR-associated detection technique to identify the presence or absence of certain base pair sequences identifying different alleles. The probes are visualized by an array of dots rather than by the "bar codes" associated with RFLP analysis.

Single-Locus Probe. A probe that only marks a specific site (locus). RFLP analysis using a single-locus probe will yield an autorad showing one band if the individual is homozygous, two bands if heterozygous.

Southern Blotting. Named for its inventor, a technique by which processed DNA fragments, separated by gel electrophoresis, are transferred onto a nylon membrane in preparation for the application of biological probes.

Thymine (T). One of the four bases, or nucleotides, that make up the DNA double helix. Thymine only binds to adenine. See Nucleotide.

Variable Number of Tandem Repeats (VNTR). Multiple copies of virtually identical base pair sequences, arranged in succession at a specific locus on a chromosome. The number of repeats varies from individual to individual, thus providing a basis for individual recognition.

References on Forensic DNA Evidence

Committee on DNA Technology in Forensic Science, National Research Council, DNA Technology in Forensic Science (1992).

Forensic DNA Technology (Mark A. Farley & James J. Harrington eds., 1991).

Lorne T. Kirby, DNA Fingerprinting: An Introduction (1990).

Office of Technology Assessment, U.S. Congress, Genetic Witness: Forensic Uses of DNA Tests (1990).

David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts*, 7 Harv. J.L. & Tech. 101 (1993) .

David H. Kaye, *The Forensic Debut of the National Research Council's DNA Report: Population Structure, Ceiling Frequencies and the Need for Numbers*, 34 Jurimetrics J. 369 (1994).