Reference Guide on DNA Evidence

DAVID H. KAYE AND GEORGE F. SENSABAUGH, JR.

David H. Kaye, M.A., J.D., is Regents' Professor of Law, Arizona State University College of Law, Tempe, Arizona

George F. Sensabaugh, Jr., D. Crim., is Professor, School of Public Health, University of California, Berkeley, California.

CONTENTS

I. Introduction

Deoxyribonucleic acid, or DNA, is a molecule that encodes the genetic information in all living organisms. Its chemical structure was elucidated in 1954. More than thirty years later, samples of human DNA began to be used in the criminal justice system, primarily in cases of rape or murder. The evidence has been the subject of extensive scrutiny by lawyers, judges, and the scientific community.[1] It is now admissible in virtually all jurisdictions, but debate lingers over the safeguards that should be required in testing samples and in presenting the evidence in court.[2] Moreover, there are many types of DNA analysis, and still more are being developed.[3] New problems of admissibility arise as advancing methods of analysis and novel applications of established methods are introduced.

This reference guide addresses technical issues that arise in considering the admissibility of and weight to be accorded analyses of DNA, and it identifies legal issues whose resolution requires scientific information.[4] The goal is to present the essential background information and to provide a framework for resolving the possible disagreements among scientists or technicians who testify as to the results and import of forensic DNA comparisons.

A. Summary of Contents

Section I lists the major objections that can be raised to the admission of DNA evidence. It also outlines the types of scientific expertise that go into the analysis of DNA samples.

---

[1] At the request of various government agencies, the National Research Council empaneled two committees for the National Academy of Sciences that produced book-length reports on forensic DNA technology, with recommendations for enhancing the rigor of laboratory work and improving the presentation of the evidence in court. Committee on DNA Technology in Forensic Science, National Research Council, DNA Technology in Forensic Science (1992) [hereinafter NRC I]; Committee on DNA Forensic Science: An Update, National Research Council, The Evaluation of Forensic DNA Evidence (1996) [hereinafter NRC II]. One author of this guide served on both committees, the other served on the second committee (NRC II), and we have drawn on those reports. We also have relied extensively on the version of this reference guide on DNA evidence by Judith A. McKenna, Joe S. Cecil, and Pamela Coukos that appeared in the 1994 edition of the *Reference Manual on Scientific Evidence*.

[2] *See* D.H. Kaye, *DNA, NAS, NRC, DAB, RFLP, PCR, and More: An Introduction to the Symposium on the 1996 NRC Report on Forensic DNA Evidence*, 37 Jurimetrics J. 395 (1997); William C. Thompson, *Guide to Forensic DNA Evidence, in* Expert Evidence: A Practitioner's Guide to Law, Science, and the *FJC Manual* 185 (Bert Black & Patrick W. Lee eds., 1997).

[3] Emerging systems of DNA analysis are described and contrasted to the established methods and markers in National Comm'n on the Future of DNA Evidence Research & Dev. Working Group, Report to the Commission (forthcoming 2000).

[4] Leading cases are collected in tables in NRC II, *supra* note 1, at 205–11. For subsequent developments, see D.H. Kaye, *DNA Identification in Criminal Cases: Lingering and Emerging Evidentiary Issues*, *in* Proceedings of the Seventh International Symposium on Human Identification 12 (1997).

Section II gives an overview of the scientific principles behind DNA typing. It describes the structure of DNA and how this molecule differs from person to person. These are basic facts of molecular biology. The section also defines the more important scientific terms. It explains at a general level how DNA differences are detected. These are matters of analytical chemistry and laboratory procedure. Finally, the section indicates how it is shown that these differences permit individuals to be identified. This is accomplished with the methods of probability and statistics.

Sections III and IV outline basic methods used in DNA testing. Section III describes methods that begin by using the polymerase chain reaction (PCR) to make many copies of short segments of DNA. Section IV examines the theory and technique of the older procedure of variable number tandem repeat (VNTR) profiling.

Section V considers issues of sample quantity and quality common to all methods of DNA profiling. Section VI deals with laboratory performance. It outlines the types of information that a laboratory should produce to establish that it can analyze DNA reliably and that it has adhered to established laboratory protocols.

Section VII examines issues in the interpretation of laboratory results. To assist the courts in understanding the extent to which the results incriminate the defendant, it enumerates the hypotheses that need to be considered before concluding that the defendant is the source of the crime-scene samples, and it explores the issues that arise in judging the strength of the evidence. It focuses on questions of statistics, probability, and population genetics.

Section VIII takes up novel applications of DNA technology, such as the forensic analysis of non-human DNA. It identifies questions that can be useful in judging whether a new method or application has the scientific merit and power claimed by the proponent of the evidence.

An appendix provides detail on technical material, and a glossary defines selected terms and acronyms encountered in genetics, molecular biology, and forensic DNA work.[5]

B. Objections to DNA Evidence

The usual objective of forensic DNA analysis is to detect variations in the genetic material that differentiate individuals one from another.[6] Laboratory techniques for isolating and analyzing DNA have long been used in scientific research and medicine. Applications of these techniques to forensic work usually

---

[5] The glossary also defines a number of other terms that may be used by experts in these fields.

[6] Biologists accept as a truism the proposition that, except for identical twins, human beings are genetically unique.

involve comparing a DNA sample obtained from a suspect with a DNA sample obtained from the crime scene. Often, a perpetrator's DNA in hair, blood, saliva, or semen can be found at a crime scene,[7] or a victim's DNA can be found on or around the perpetrator.[8]

In many cases, defendants have objected to the admission of testimony of a match or its implications.[9] Under *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[10] the district court, in its role as "gatekeeper" for scientific evidence, then must ensure that the expert's methods are scientifically valid and reliable. Because the basic theory and most of the laboratory techniques of DNA profiling are so widely accepted in the scientific world, disputed issues involve features unique to their forensic applications or matters of laboratory technique. These include the extent to which standard techniques have been shown to work with crime-scene samples exposed to sunlight, heat, bacteria, and chemicals in the environment; the extent to which the specific laboratory has demonstrated its ability to follow protocols that have been validated to work for crime-scene samples; possible ambiguities that might interfere with the interpretation of test results; and the validity and possible prejudicial impact of estimates of the probability of a match between the crime-scene samples and innocent suspects.

C. Relevant Expertise

DNA identification can involve testimony about laboratory findings, about the statistical interpretation of these findings, and about the underlying principles of molecular biology. Consequently, expertise in several fields might be required to establish the admissibility of the evidence or to explain it adequately to the jury. The expert who is qualified to testify about laboratory techniques might

---

[7] *E.g.*, United States v. Beasley, 102 F.3d 1440 (8th Cir. 1996) (two hairs were found in a mask used in a bank robbery and left in the abandoned get-away car); United States v. Two Bulls, 918 F.2d 56 (8th Cir. 1990), *vacated for reh'g en banc, app. dismissed due to death of defendant*, 925 F.2d 1127 (1991) (semen stain on victim's underwear).

[8] *E.g.*, United States v. Cuff, 37 F. Supp. 2d 279 (S.D.N.Y. 1999) (scrapings from defendant's fingernails); State v. Bible, 858 P.2d 1152 (Ariz. 1993) (bloodstains on defendant's shirt); People v. Castro, 545 N.Y.S.2d 985 (Bronx Co. Sup. Ct. 1989) (bloodstains on defendant's watch). For brevity, we refer only to the typical case of a perpetrator's DNA at a crime scene. The scientific and legal issues in both situations are the same.

[9] Exclusion of the testimony can be sought before or during trial, depending on circumstances and the court's rules regarding pretrial motions. Pretrial requests for discovery and the appointment of experts to assist the defense also can require judicial involvement. *See, e.g.*, Dubose v. State, 662 So. 2d 1189 (Ala. 1995) (holding that due process was violated by the failure to provide an indigent defendant with funds for an expert); Paul C. Giannelli, *The DNA Story: An Alternative View*, 88 J. Crim. L. & Criminology 380, 414–17 (1997) (book review) (criticizing the reluctance of state courts to appoint defense experts and to grant discovery requests); Paul C. Giannelli, *Criminal Discovery, Scientific Evidence, and DNA*, 44 Vand. L. Rev. 791 (1991); NRC II, *supra* note 1, at 167–69.

[10] 509 U.S. 579 (1993).

not be qualified to testify about molecular biology, to make estimates of population frequencies, or to establish that an estimation procedure is valid.[11]

Trial judges ordinarily are accorded great discretion in evaluating the qualifications of a proposed expert witness, and the decisions depend on the background of each witness. Courts have noted the lack of familiarity of academic experts—who have done respected work in other fields—with the scientific literature on forensic DNA typing,[12] and on the extent to which their research or teaching lies in other areas.[13] Although such concerns may give trial judges pause, they rarely result in exclusion of the testimony on the ground that the witness simply is not qualified as an expert.[14]

The scientific and legal literature on the objections to DNA evidence is extensive.[15] By studying the scientific publications, or perhaps by appointing a special master or expert adviser to assimilate this material, a court can ascertain where a party's expert falls in the spectrum of scientific opinion. Furthermore, an expert appointed by the court under Rule 706 could testify about the scientific literature generally or even about the strengths or weaknesses of the particular arguments advanced by the parties.[16]

---

11. *See* 1 McCormick on Evidence § 203, at 875 n.40 (John W. Strong ed., 1992). Nevertheless, if previous cases establish that the testing and estimation procedures are legally acceptable, and if the computations are essentially mechanical, then highly specialized statistical expertise might not be essential. Reasonable estimates of DNA characteristics in major population groups can be obtained from standard references, and many quantitatively literate experts could use the appropriate formulae to compute the relevant profile frequencies or probabilities. NRC II, *supra* note 1, at 170. Limitations in the knowledge of a technician who applies a generally accepted statistical procedure can be explored on cross-examination. *E.g.*, State v. Colbert, 896 P.2d 1089 (Kan. 1995) (in view of general acceptance of databases, estimate of probability was admissible despite an expert's concessions that he was not a population geneticist and was not qualified to explain how the databases applied to the town of Coffeyville); State v. Harvey, 699 A.2d 596, 637 (N.J. 1997) (statistician not required).

12. *E.g.*, State v. Copeland, 922 P.2d 1304, 1318 n.5 (Wash. 1996) (noting that defendant's statistical expert "was also unfamiliar with publications in the area," including studies by "a leading expert in the field" whom he thought was "a guy in a lab somewhere").

13. *E.g., id.* (noting that defendant's population genetics expert "had published little in the field of human genetics, only one non-peer reviewed chapter in a general text, had two papers in the area rejected, was uninformed of the latest articles in the field, had misused a statistical model . . . , had no graduate students working under him, had not received any awards in his field in over ten years, had not received a research grant in about eight years, and made about $100,000 testifying as an expert in 1990–91").

14. *E.g.*, Commonwealth v. Blasioli, 685 A.2d 151 (Pa. Super. Ct. 1996) (professor of ecology and evolutionary biology was said to be qualified, but "barely").

15. *See, e.g.*, Bruce S. Weir, *A Bibliography for the Use of DNA in Human Identification, in* Human Identification: The Use of DNA Markers 179–213 (Bruce S. Weir ed., 1995); NRC II, *supra* note 1, at 226–39 (list of references).

16. Some courts have appointed experts to address general questions relating to DNA profiling. *E.g.*, United States v. Bonds, 12 F.3d 540 (6th Cir. 1993); United States v. Porter, Crim. No. F06277-89, 1994 WL 742297 (D.C. Super. Ct. Nov. 17, 1994) (mem.). Whether a court should appoint its own expert instead of an expert for the defense when there are more specific disputes is more controversial.

II. Overview of Variation in DNA and Its Detection

A. DNA, Chromosomes, Sex, and Genes

DNA is a complex molecule that contains the "genetic code" of organisms as diverse as bacteria and humans.[17] The molecule is made of subunits that include four nucleotide bases, whose names are abbreviated to A, T, G, and C.[18] The physical structure of DNA is described more fully in the appendix, but for general purposes it suffices to say that a DNA molecule is like a long sequence of these four letters, where the chemical structure that corresponds to each letter is known as a base pair.

Most human DNA is tightly packed into structures known as chromosomes, which are located in the nuclei of most cells.[19] If the bases are like letters, then each chromosome is like a book written in this four-letter alphabet, and the nucleus is like a bookshelf in the interior of the cell. All the cells in one individual contain copies of the same set of books. This library, so to speak, is the individual's genome.[20]

In human beings, the process that produces billions of cells with the same genome starts with sex. Every sex cell (a sperm or ovum) contains 23 chromosomes. When a sperm and ovum combine, the resulting fertilized cell contains 23 pairs of chromosomes, or 46 in all. It is as if the father donates half of his collection of 46 books, and the mother donates a corresponding half of her collection. During pregnancy, the fertilized cell divides to form two cells, each of which has an identical copy of the 46 chromosomes. The two then divide to form four, the four form eight, and so on. As gestation proceeds, various cells specialize to form different tissues and organs. In this way, each human being has immensely many copies[21] of the original 23 pairs of chromosomes from the fertilized egg, one member of each pair having come from the mother and one from the father.

All told, the DNA in the 23 chromosomes contains over three billion letters (base pairs) of genetic "text."[22] About 99.9% is identical between any two individuals. This similarity is not really surprising—it accounts for the common features that make humans an identifiable species. The remaining 0.1% is particular to an individual (identical twins excepted). This variation makes each

---

[17] Some viruses use a related nucleic acid, RNA, instead of DNA to encode genetic information.

[18] The full names are adenine, thymine, guanine, and cytosine.

[19] A few types of cells, such as red blood cells, do not contain nuclei.

[20] Originally, "genome" referred to the set of base pairs in an egg or sperm, but the term also is used to designate the ordered set in the fertilized cell.

[21] The number of cells in the human body has been estimated at more than 10^{15} (a million billion).

[22] If the base pairs were listed as letters in a series of books, one piled on top of the other, the pile would be as high as the Washington Monument.

person genetically unique.

A gene is a particular DNA sequence, usually from 1,000 to 10,000 base pairs long, that "codes" for an observable characteristic.[23] For example, a tiny part of the sequence that directs the production of the human group-specific complement protein (GC)[24] is

G C A A A A T T G C C T G A T G C C A C A C C C A A G G A A C T G G C A[25]

This gene always is located at the same position, or locus, on chromosome number 4. As we have seen, most individuals have two copies of each gene at a given locus—one from the father and one from the mother.

A locus where almost all humans have the same DNA sequence is called monomorphic ("of one form"). A locus at which the DNA sequence varies among individuals is called polymorphic ("of many forms"). The alternative forms are called alleles. For example, the GC protein gene sequence has three common alleles that result from single nucleotide polymorphisms (SNPs, pronounced "snips")—substitutions in the base that occur at a given point.[26] In the scientific literature, the three alleles are designated Gc★1F, Gc★1S, and Gc★2, and the sequences at the variable sites are shown in Figure 1.

Figure 1. The variable sequence region of the group-specific component gene. The base substitutions that define the alleles are shown in bold.

Allele ★2:	G C A A A A T T G C C T G A T G C C A C A C C C A A G G A A C T G G C A
Allele ★1F:	G C A A A A T T G C C T G A T G C C A C A C C C A **C** G G A A C T G G C A
Allele ★1S:	G C A A A A T T G C C T G A **G** G C C A C A C C C A **C** G G A A C T G G C A

In terms of the metaphor of DNA as text, the gene is like an important paragraph in the book; a SNP is a change in a letter somewhere within that paragraph, and the two versions of the paragraph that result from this slight change are the alleles. An individual who inherits the same allele from both parents is

---

[23] The genetic code consists of "words" that are three nucleotides long and that determine the structure of the proteins that are manufactured in cells. *See, e.g.*, Elaine Johnson Mange & Arthur P. Mange, Basic Human Genetics 107 (2d ed. 1999).

[24] This "GC" stands for "group-specific component," and not for the bases guanine and cytosine.

[25] The full GC gene is nearly 42,400 base pairs in length. The product of this gene is also known as vitamin D–binding protein. GC is one of the five loci included in the polymarker (PM) typing kit, which is widely used in forensic testing.

[26] *See* R.L. Reynolds & G.F. Sensabaugh, *Use of the Polymerase Chain Reaction for Typing Gc Variants, in* 3 Advances in Forensic Haemogenetics 158 (H.F. Polesky & W.R. Mayr eds. 1990); Andreas Braun et al., *Molecular Analysis of the Gene for the Human Vitamin-D-binding Protein (Group-specific Component): Allelic Differences of the Common Genetic GC Types*, 89 Hum. Genetics 401 (1992). These are examples of *point mutations*.

called a homozygote.[27] An individual with distinct alleles is termed a heterozygote.[28]

Regions of DNA used for forensic analysis usually are not genes, but parts of the chromosome without a known function. The "non-coding" regions of DNA have been found to contain considerable sequence variation, which makes them particularly useful in distinguishing individuals. Although the terms "locus," "allele," "homozygous," and "heterozygous" were developed to describe genes, the nomenclature has been carried over to describe all DNA variation—coding and non-coding alike—for both types are inherited from mother and father in the same fashion.

B. Types of Polymorphisms and Methods of Detection

By determining which alleles are present at strategically chosen loci, the forensic scientist ascertains the genetic profile, or genotype, of an individual. Genotyping does not require "reading" the full DNA sequence; indeed, direct sequencing is technically demanding and time-consuming.[29] Rather, most genetic typing focuses on identifying only those variations that define the alleles and does not attempt to "read out" each and every base as it appears.[30]

For instance, simple sequence variation, such as that for the GC locus, is conveniently detected using a sequence-specific oligonucleotide (SSO) probe. With GC typing, probes for the three common alleles (which we shall call A_1, A_2, and A_3) are attached to designated locations on a membrane. When DNA with a given allele (say, A_1) comes in contact with the probe for that allele, it sticks.[31] To get a detectable quantity of DNA to stick, many copies of the variable sequence region of the GC gene in the DNA sample have to be made.[32] All this DNA then is added to the membrane. The DNA fragments with the allele A_1 in them stick to the spot with the A_1 probe. To permit these fragments to be seen, a chemical "label" that catalyses a color change at the spot where the DNA

---

[27] For example, someone with the Gc*2 allele on both number 4 chromosomes is homozygous at the GC locus. This homozygous GC genotype is designated as 2,2 (or simply 2).

[28] For example, someone with the Gc*2 allele on one chromosome and the Gc*1F allele on the other is heterozygous at the GC locus. This heterozygous genotype is designated as 2,1F.

[29] However, automated machinery for direct sequencing has been developed and is used at major research centers engaged in the international endeavor to sequence the human genome (and the genomes of other organisms). *See* R. Waterston & J.E. Sulston, *The Human Genome Project: Reaching the Finish Line*, 282 Science 53 (1998).

[30] For example, genetic typing at the GC locus focuses on the sequence region shown in Figure 1; the remainder of the 42,300 base pairs of the GC gene sequence is the same for almost all individuals and is ignored for genetic typing purposes.

[31] This process of hybridization is described in Part B of the Appendix.

[32] The polymerase chain reaction (PCR) is used to make many copies of the DNA that is to be typed. PCR is roughly analogous to copying and pasting a section of text with a word processor. *See infra* the Appendix, Part D.

binds to its probe can be attached when the copies are made. A colored spot showing that the A_1 allele is present thus should appear on the membrane.[33]

Another category of polymorphism is characterized by the insertion of a variable number of tandem repeats (VNTR) at a locus.[34] The core unit of a VNTR is a particular short DNA sequence that is repeated many times end-to-end. This repetition gives rise to alleles with length differences; regions of DNA containing more repeats are larger than those containing fewer repeats. Genetic typing of polymorphic VNTR loci employs electrophoresis, a technique that separates DNA fragments based on size.[35]

The first polymorphic VNTRs to be used in genetic and forensic testing had core repeat sequences of 15–35 base pairs. Alleles at VNTR loci of this sort generally are too long to be measured precisely by electrophoretic methods—alleles differing in size by only a few repeat units may not be distinguished. Although this makes for complications in deciding whether two length measurements that are close together result from the same allele, these loci are quite powerful for the genetic differentiation of individuals, for they tend to have many alleles that occur relatively rarely in the population. At a locus with only twenty such alleles (and most loci typically have many more), there are 210 possible genotypes.[36] With five such loci, the number of possible genotypes is 210^5, which is more than 400 billion. Thus, VNTRs are an extremely discriminating class of DNA markers.

More recently, the attention of the genetic typing community has shifted to repetitive DNA characterized by short core repeats, two to seven base pairs in length. These non-coding DNA sequences are known as short tandem repeats (STRs).[37] Because STR alleles are much smaller than VNTR alleles, electrophoretic detection permits the exact number of base pairs in an STR to be determined, permitting alleles to be defined as discrete entities. Figure 2 illustrates the nature of allelic variation at a polymorphic STR locus. The first allele has nine tandem repeats, the second has ten, and the third has eleven.[38]

Figure 2. Three Alleles of an STR with the Core Sequence ATTT

ATTTATTTATTTATTTATTTATTTATTTATTTATTT

ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT

ATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT

---

[33] This approach can be miniaturized and automated with hybridization chip technology. *See infra* Glossary of Terms ("chip").

[34] VNTR polymorphisms also are referred to as minisatellites.

[35] We describe one form of electrophoresis often used with VNTR loci *infra* § IV.

[36] There are 20 homozygous genotypes and another $(20 \times 19)/2 = 190$ heterozygous ones.

[37] They also are known as microsatellites.

[38] To conserve space, the figure uses alleles that are unrealistically short. A typical STR is in the range of 50–350 base pairs in length. In contrast, a typical VNTR is thousands of base pairs long.

Although there are fewer alleles per locus for STRs than for VNTRs, there are many STRs, and they can be analyzed simultaneously.[39] As more STR loci are included, STR testing becomes more revealing than VNTR profiling at four or five loci.[40]

Full DNA sequencing is employed at present only for mitochondrial DNA (mtDNA).[41] Mitochondria are small structures found inside the cell. In these organelles, certain molecules are broken down to supply energy. Mitochondria have a small genome that bears no relation to the chromosomal genome in the cell nucleus.[42] Mitochondrial DNA has three features that make it useful for forensic DNA testing. First, the typical cell, which has but one nucleus, contains hundreds of identical mitochondria.[43] Hence, for every copy of chromosomal DNA, there are hundreds of copies of mitochondrial DNA. This means that it is possible to detect mtDNA in samples containing too little nuclear DNA for conventional typing.[44] Second, the mtDNA contains a sequence region of about a thousand base pairs that varies greatly among individuals. Finally, mitochondria are inherited mother to child,[45] so that siblings, maternal half-siblings, and others related through maternal lineage possess the same mtDNA sequence.[46] This last feature makes mtDNA particularly useful for associating persons related through their maternal lineage—associating skeletal remains to a family, for example.[47]

---

[39] The procedures for simultaneous detection are known as multiplex methods. *See infra* Glossary of Terms ("capillary electrophoresis," "chip"). Mass spectrometry also can be applied to detect STR fragments. *Id.*

[40] Usually, there are between seven and fifteen STR alleles per locus. Thirteen loci that have ten STR alleles each can give rise to 55^{13}, or 42 billion trillion, possible genotypes.

[41] The first use of this mtDNA analysis as evidence in a criminal case occurred in Tennessee in *State v. Ware,* No. 03C01-9705CR00164, 1999 WL 233592 (Tenn. Crim. App. Apr. 20, 1999). *See* Mark Curriden, *A New Evidence Tool: First Use of Mitochondrial DNA Test in a U.S. Criminal Trial*, A.B.A.J., Nov. 1996, at 18.

[42] In contrast to the haploid nuclear genome of over three billion base pairs, the mitochondrial genome is a circular molecule 16,569 base pairs long.

[43] There are from 75 to 1,000 or so mitochondria per cell.

[44] Even so, because the mitochondrial genome is so much shorter than the nuclear genome, it is a tiny fraction of the total mass of DNA in a cell.

[45] Although sperm have mitochondria, these are not passed to the ovum at fertilization. Thus the only mitochondria present in the newly fertilized cell originate from the mother.

[46] Evolutionary studies suggest an average mutation rate for the mtDNA control region of one nucleotide difference every 300 generations, or one difference every 6,000 years. Consequently, one would not expect to see many examples of nucleotide differences between maternal relatives. On the other hand, differences in the bases at a specific sequence position among the copies of the mtDNA within an individual have been seen. This heteroplasmy, which is more common in hair than other tissues, counsels against declaring an exclusion on the basis of a single base pair difference between two samples.

[47] *See, e.g.*, Peter Gill et al., *Identification of the Remains of the Romanov Family by DNA Analysis*, 6 Nature Genetics 130 (1994).

Just as genetic variation in mtDNA can be used to track maternal lineages, genetic variations on the Y chromosome can be used to trace paternal lineages. Y chromosomes, which contain genes that result in development as a male rather than a female, are found only in males and are inherited father to son.[48] Markers on this chromosome include STRs and SNPs,[49] and they have been used in cases involving semen evidence.[50]

In sum, DNA contains the genetic information of an organism. In humans, most of the DNA is found in the cell nucleus, where it is organized into separate chromosomes. Each chromosome is like a book, and each cell has the same library of books of various sizes and shapes. There are two copies of each book of a particular size and shape, one that came from the father, the other from the mother. Thus, there are two copies of the book entitled "Chromosome One," two copies of "Chromosome Two," and so on. Genes are the most meaningful paragraphs in the books, and there are differences (polymorphisms) in the spelling of certain words in the paragraphs of different copies of each book. The different versions of the same paragraph are the alleles. Some alleles result from the substitution of one letter for another. These are SNPs. Others come about from the insertion or deletion of single letters, and still others represent a kind of stuttering repetition of a string of extra letters. These are the VNTRs and STRs. In addition to the 23 pairs of books in the cell nucleus, another page or so of text resides in each of the mitochondria, the power plants of the cell.

The methods of molecular biology permit scientists to determine which alleles are present. The next two sections describe how this is done. Section III discusses the procedures that can distinguish among all the known alleles at certain loci. Section IV deals with the "RFLP" procedures that measure the lengths of DNA fragments at a scale that is not fine enough to resolve all the possible alleles.

---

[48] *See infra* note 110.

[49] *See, e.g.*, M.F. Hammer et al., *The Geographic Distribution of Human Y Chromosome Variation*, 145 Genetics 787 (1997). The Y chromosome is used in evolutionary studies along with mtDNA to learn about human migration patterns. *Id.*; Michael F. Hammer & Stephen L. Zegura, *The Role of the Y Chromosome in Human Evolutionary Studies*, 5 Evolutionary Anthropology 116 (1996). The various markers are inherited as a single package (known as a haplotype).

[50] They also were used in a family study to ascertain whether President Thomas Jefferson fathered a child of his slave, Sally Hemings. *See* Eugene A. Foster et al., *Jefferson Fathered Slave's Last Child*, 396 Nature 27 (1998); Eliot Marshall, *Which Jefferson Was the Father?*, 283 Science 153 (1999).

III. DNA Profiling with Loci Having Discrete Alleles

Simple sequence variations and STRs occur within relatively short fragments of DNA. These polymorphisms can be analyzed with so-called PCR-based tests (PCR = polymerase chain reaction). The three steps of PCR-based typing are (1) DNA extraction, (2) amplification, and (3) detection of genetic type using a method appropriate to the polymorphism. This section discusses the scientific and technological foundations of these three steps and the basis for believing that the DNA characteristics identified in the laboratory can help establish who contributed the potentially incriminating DNA.[51]

A. DNA Extraction and Amplification

DNA usually can be found in biological materials such as blood, bone, saliva, hair, semen, and urine.[52] A combination of routine chemical and physical methods permit DNA to be extracted from cell nuclei and isolated from the other chemicals in a sample.[53] Thus, the premise that DNA is present in many biological samples and can be removed for further analysis is firmly established.[54]

Just as the scientific foundations of DNA extraction are clear, the procedures for amplifying DNA sequences within the extracted DNA are well established. The first National Academy of Sciences committee on forensic DNA typing described the amplification step as "simple . . . analogous to the process by which cells replicate their DNA."[55] Details of this process, which can make millions of copies of a single DNA fragment, are given in the Appendix.

---

[51] The problem of drawing an inference about the source of the evidence DNA, which is common to all forms of DNA profiling, is taken up in section VII.

[52] *See, e.g.*, NRC I, *supra* note 1, at 28, tbl.1.1.

[53] *See, e.g.*, Michael L. Baird, *DNA Profiling: Laboratory Methods*, *in* 1 Modern Scientific Evidence: The Law and Science of Expert Testimony § 16-2.2, at 667 (David L. Faigman et al. eds., 1997) [hereinafter Modern Scientific Evidence]; Catherine T. Comey et al., *DNA Extraction Strategies for Amplified Fragment Length Polymorphism Analysis*, 39 J. Forensic Sci. 1254 (1994); Atsushi Akane et al., *Purification of Forensic Specimens for the Polymerase Chain Reaction (PCR) Analysis*, 38 J. Forensic Sci. 691 (1993).

[54] *See, e.g.*, NRC I, *supra* note 1, at 149 (recommending judicial notice of the proposition that "DNA polymorphisms can, in principle, provide a reliable method for comparing samples," "although the actual discriminatory power of any particular DNA test will depend on the sites of DNA variation examined"); NRC II, *supra* note 1, at 9 ("DNA typing, with its extremely high power to differentiate one human being from another, is based on a large body of scientific principles and techniques that are universally accepted.").

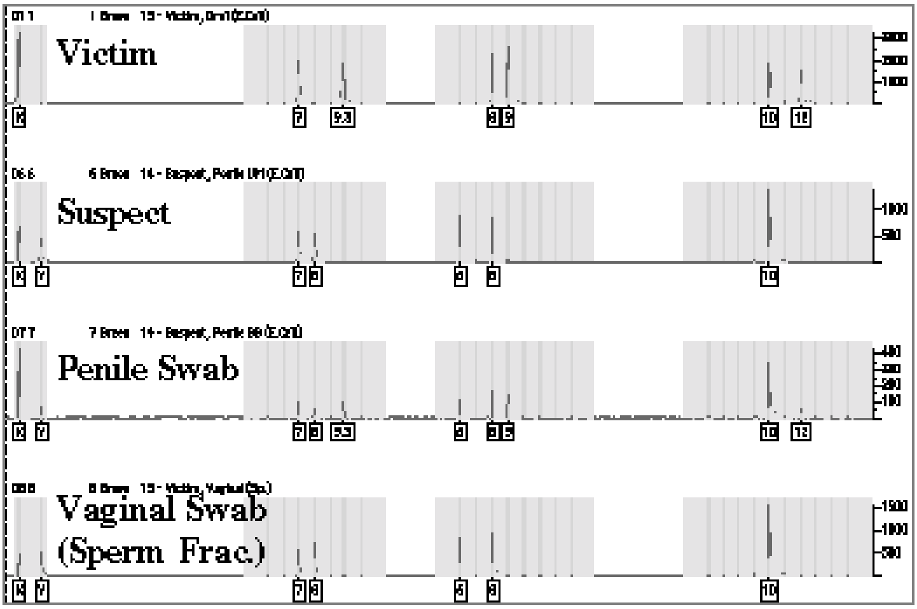
[55] NRC I, *supra* note 1, at 40. The second committee used similar language, reporting that "[t]he PCR process is relatively simple and easily carried out in the laboratory." NRC II, *supra* note 1, at 70. *But see* NRC I, *supra*, at 63 ("Although the basic exponential amplification procedure is well understood, many technical details are not, including why some primer pairs amplify much better than others, why some loci cause systematically unfaithful amplification, and why some assays are much more sensitive to variations in conditions."). For these reasons, PCR-based procedures are validated by experiment.

For amplification to work properly and yield copies of only the desired sequence, however, care must be taken to achieve the appropriate biochemical conditions and to avoid excessive contamination of the sample.[56] A laboratory should be able to demonstrate that it can faithfully amplify targeted sequences with the equipment and reagents that it uses[57] and that it has taken suitable precautions to avoid or detect handling or carryover contamination.[58]

B. DNA Analysis

To determine whether the DNA sample associated with a crime could have come from a suspect, the genetic types as determined by analysis of the DNA amplified from the crime-scene sample are compared to the genetic types as determined for the suspect. For example, Figure 3 shows the results of STR typing at four loci in a sexual assault case.[59]

Figure 3. Sexual Assault Case (CTTA)



---

[56] *See* NRC I, *supra* note 1, at 63–67; NRC II, *supra* note 1, at 71.

[57] *See* NRC I, *supra* note 1, at 63–64.

[58] Carryover occurs when the DNA product of a previous amplification contaminates samples or reaction solutions. *See id.* at 66.

[59] The initials CTTA refer to these loci, which are known as CPO, TPO, THO, and amelogenin.

The peaks result from DNA fragments of different sizes.[60] The bottom row shows the profile of sperm DNA isolated from a vaginal swab. These sperm have two alleles at the first locus (indicating that both X and Y chromosomes are present),[61] two alleles at the second locus (consisting of 7 and 8 repeat units), two at the third locus (a 6 and an 8), and one (a 10 on each chromosome) at the fourth.[62] The same profile also appears in the DNA taken from the suspect. DNA from a penile swab from the suspect is consistent with a mixture of DNA from the victim and the suspect.

Regardless of the kind of genetic system used for typing—STRs, Amp-FLPs,[63] SNPs, or still other polymorphisms[64]—some general principles and questions can be applied to each system that is offered for courtroom use. As a beginning, the nature of the polymorphism should be well characterized. Is it a simple sequence polymorphism or a fragment length polymorphism? This information should be in the published literature or in archival genome databanks.[65]

Second, the published scientific literature also can be consulted to verify claims that a particular method of analysis can produce accurate profiles under various conditions.[66] Although such validation studies have been conducted for all the discrete-allele systems ordinarily used in forensic work, determining the point at which the empirical validation of a particular system is sufficiently convincing to pass scientific muster may well require expert assistance.

Finally, the population genetics of the marker should be characterized. As new marker systems are discovered, researchers typically analyze convenient collections of DNA samples from various human populations[67] and publish studies

---

[60] 60. The height of (more, precisely, the area under) each peak is related to the amount of DNA in the gel.

[61] 61. The X-Y typing at the first locus is simply used to verify the sex of the source of the DNA. XY is male, and XX is female. *See infra* note 110. That these markers show that the victim is female and the suspect male helps demonstrate that a valid result has been obtained.

[62] 62. Although each sperm cell contains only one set of chromosomes, a collection of many sperm cells from the same individual contains both sets of chromosomes. *See infra* note 90.

[63] 63. "Amp-FLP" is short for "Amplified Fragment Length Polymorphism." The DNA fragment is produced by amplifying a longish sequence with a PCR primer. The longer Amp-FLPs, such as DS180, overlap the shorter VNTRs. In time, PCR methods will be capable of generating longer Amp-FLPs.

[64] 64. *See supra* § II; *infra* Appendix, Part C (Table A-1).

[65] 65. Primary data regarding gene sequence variation is increasingly being archived in publicly accessible computer databanks, such as GenBank, rather than in the print literature. *See* Victor A. McKusick, *The Human Genome Project: Plans, Status, and Applications in Biology and Medicine, in* Gene Mapping: Using Law and Ethics as Guides 18, 35 (George J. Annas & Sherman Elias eds., 1992). This trend is driven by an explosion of new data coupled with the fact that most of the detected variation has no known biological significance and hence is not particularly noteworthy.

[66] 66. *Cf.* NRC I, *supra* note 1, at 72 ("Empirical validation of a DNA typing procedure must be published in appropriate scientific journals.").

[67] 67. The samples come from diverse sources, such as blood banks, law enforcement personnel, paternity cases, and criminal cases. Reliable inferences probably can be drawn from these samples. *See infra* note 178.

of the relative frequencies of each allele in these population samples. These database studies give a measure of the extent of genetic variability at the polymorphic locus in the various populations, and thus of the potential probative power of the marker for distinguishing between individuals.

At this point, the existence of PCR-based procedures that can ascertain genotypes accurately cannot be doubted.[68] Of course, the fact that scientists have shown that it is possible to extract DNA, to amplify it, and to analyze it in ways that bear on the issue of identity does not mean that a particular laboratory has adopted a suitable protocol and is proficient in following it. These laboratory-specific issues are considered in section VI.[69]

IV. VNTR Profiling

VNTR profiling, described in section II, was the first widely used method of forensic DNA testing. Consequently, its underlying principles, its acceptance within the scientific community, and its scientific soundness have been discussed in a great many opinions.[70] Because so much has been written on VNTR profiling, only the basic steps of the procedure will be outlined here.

---

[68] *See, e.g.*, United States v. Shea, 159 F.3d 37 (1st Cir. 1998) (DQA, Polymarker, D1S80), *cert. denied*, 119 S. Ct. 1480 (1999); United States v. Lowe, 145 F.3d 45 (1st Cir. 1998) (DQA, Polymarker, D1S80); United States v. Beasley, 102 F.3d 1440, 1448 (8th Cir. 1996) (DQA, Polymarker); United States v. Hicks, 103 F.3d 837 (9th Cir. 1996) (DQA); United States v. Gaines, 979 F. Supp. 1429 (S.D. Fla. 1997) (DQA, Polymarker, D1S80); State v. Hill, 895 P.2d 1238 (Kan. 1995) (DQA); Commonwealth v. Rosier, 685 N.E.2d 739 (Mass. 1997) (STRs); Commonwealth v. Vao Sok, 683 N.E.2d 671 (Mass. 1997) (DQA, Polymarker, D1S80); State v. Moore, 885 P.2d 457 (Mont. 1994) (DQA), *overruled on other grounds in* State v. Gollehon, 906 P.2d 697 (Mont. 1995); State v. Harvey, 699 A.2d 596 (N.J. 1997) (DQA, Polymarker); State v. Lyons, 924 P.2d 802 (Or. 1996) (DQA); State v. Moeller, 548 N.W.2d 465 (S.D. 1996) (DQA); State v. Begley, 956 S.W.2d 471 (Tenn. 1997) (DQA); State v. Russell, 882 P.2d 747, 768 (Wash. 1994) (DQA).

[69] Some commentators have assumed or argued that some or all of these issues are aspects of admissibility under Federal Rule of Evidence 702. *E.g.*, 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 (1991); Barry C. Scheck, *DNA and* Daubert, 15 Cardozo L. Rev.. 1959, 1979–87 (1994); William C. Thompson, *Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence*, 37 Jurimetrics J. 405, 417 (1997). This reading of *Daubert* is rejected in *United States v. Shea*, 957 F. Supp. 331, 340–41 (D.N.H. 1997), but the protocols of a specific laboratory and the proficiency of its analysts are factors that affect probative value under Federal Rule of Evidence 403. *See* Margaret A. Berger, *Laboratory Error Seen Through the Lens of Science and Policy*, 30 U.C. Davis L. Rev. 1081 (1997); Edward J. Imwinkelried, *The Case Against Evidentiary Admissibility Standards that Attempt to "Freeze" the State of a Scientific Technique*, 67 U. Colo. L. Rev. 887 (1996).

[70] *See* NRC II, *supra* note 1, at 205–11 (listing leading cases and status as of 1995, by jurisdiction). The first reported appellate opinion is *Andrews v. State*, 533 So. 2d 841 (Fla. Dist. Ct. App. 1988).

1. Like profiling by means of discrete allele systems,[71] VNTR profiling begins with the extraction of DNA from a crime-scene sample. (Because this DNA is not amplified, however, larger quantities of higher quality DNA[72] are required.)

2. The extracted DNA is "digested" by a restriction enzyme that recognizes a particular, very short sequence; the enzyme cuts the DNA at these restriction sites. When a VNTR falls between two restriction sites, the resulting DNA fragments will vary in size depending on the number of core repeat units in the VNTR region.[73] (These VNTRs are thus referred to as a restriction fragment length polymorphism, or RFLP.)

3. The digested DNA fragments are then separated according to size by gel electrophoresis. The digest 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. 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 in a fixed time than longer, heavier fragments.

4. The resulting array of fragments is transferred for manageability to a sheet of nylon by a process known as Southern blotting.[74]

5. The restriction fragments representing a particular polymorphic locus are "tagged" on the membrane using a sequence-specific probe labeled with a radioactive or chemical tag.[75]

6. The position of the specifically bound probe tag is made visible, either by autoradiography (for radioactive labels) or by a chemical reaction (for chemical labels). For autoradiography, the washed nylon membrane is placed between

---

[71] *See supra* § III.

[72] "Quality" refers to the extent to which the original, very long strands of DNA are intact. When DNA degrades, it forms shorter fragments. RFLP testing requires fragments that are on the order of at least 20,000–30,000 base pairs long.

[73] *See supra* § II.

[74] This procedure is named after its inventor, Edwin Southern. Either before or during this transfer, the DNA is denatured ("unzipped") by alkali treatment, separating each double helix (*see infra* Appendix, Figure A-1) into two single strands. The weak bonds that connect the two members of a base pair are easily broken by heat or chemical treatment. The bonds that hold a base to the backbone and keep the backbone intact are much stronger. Thus, the double-stranded helix separates neatly into two single strands, with one base at each position.

[75] This locus-specific probe is a single strand of DNA that binds to its complementary sequence of denatured DNA in the sample. *See supra* § II.B. 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 (or different restriction enzymes) are used, test results are not comparable.

two sheets of photographic film. Over time, the radioactive probe material exposes the film where the biological probe has hybridized with the DNA fragments.[76] 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 individual who is the source of the DNA is a heterozygote at that locus. If the autorad shows only one band, the person may be homozygous for that allele (that is, each parent contributed the same allele), or the second band may be present but invisible for technical reasons. The band pattern defines the person's genotype at the locus associated with the probe.

Once an appropriately exposed autorad is obtained, the probe is stripped from the membrane, and the process is repeated with a separate probe for each locus tested. Three to five probes are typically used, the number depending in part on the amount of testable DNA recovered from the crime-scene sample. The result is a set of autorads, each of which shows the results of one probe.[77] If the crime-scene and suspect samples yield bands that are closely aligned on each autorad, the VNTR profiles[78] from the two samples are considered to match.[79]

A. Validity of the Underlying Scientific Theory

The basic theory underlying VNTR profiling is textbook knowledge. The molecular structure of DNA,[80] the presence of highly polymorphic VNTR loci,[81] and the existence of methods to produce VNTR fragments and measure their lengths are not in doubt.[82] Indeed, some courts have taken judicial notice of

---

[76] One film per probe is checked during the process to see whether the process is complete. Because this can weaken the image, the other film is left undisturbed, and it is used in comparing the positions of the bands.

[77] For a photograph of an autorad, see, e.g., NRC II, *supra* note 1, at 68 fig. 2.4.

[78] Each autorad reveals a single-locus genotype. The collection of single-locus profiles, one for each single-locus probe, sometimes is called a multi-locus VNTR profile. A "multi-locus probe," however, is a single probe that produces bands on a single autorad by hybridizing with VNTRs from many loci at the same time. It is, in other words, like a cocktail of single-locus probes. Because it is more difficult to interpret autoradiographs from multi-locus probes, these probes are no longer used in criminal cases in the United States.

[79] Issues that arise in interpreting autoradiographs and declaring matches are considered *infra* § IV.

[80] *See supra* § II.

[81] Studies of the population genetics of VNTR loci are reviewed in NRC II, *supra* note 1. *See also infra* § VII.

[82] *See, e.g.*, NRC I, *supra* note 1, at 149 (recommending judicial notice of the proposition that "DNA polymorphisms can, in principle, provide a reliable method for comparing samples," but cautioning that "the actual discriminatory power of any particular DNA test will depend on the sites of DNA variation examined"); NRC II, *supra* note 1, at 9 ("DNA typing, with its extremely high power to differentiate one human being from another, is based on a large body of scientific principles and techniques that are universally accepted."); *id.* at 36 ("Methods of DNA profiling are firmly grounded in molecular technology. When profiling is done with appropriate care, the results are highly reproducible.").

these scientific facts.[83] In short, the ability to discriminate between human DNA samples using a relatively small number of VNTR loci is widely accepted.

B. Validity and Reliability of the Laboratory Techniques

The basic laboratory procedures for VNTR analysis have been used in other settings for many years: "The complete process—DNA digestion, electrophoresis, membrane transfer, and hybridization—was developed by Edwin Southern in 1975 . . . . These procedures are routinely used in molecular biology, biochemistry, genetics, and clinical DNA diagnosis . . . ."[84] Thus, "no scientific doubt exists that [these technologies] accurately detect genetic differences."[85]

Before concluding that a particular enzyme-probe combination produces accurate profiles as applied to crime-scene samples at a particular laboratory, however, courts may wish to consider studies concerning the effects of environmental conditions and contaminants on VNTR profiling as well as the laboratory's general experience and proficiency with these probes.[86] And the nature of the sample and other considerations in a particular case can affect the certainty of the profiling. The next two sections outline the type of inquiry that can help assess the accuracy of a profile in a specific case.

V. Sample Quantity and Quality

The primary determinants of whether DNA typing can be done on any particular sample are (1) the quantity of DNA present in the sample and (2) the extent to which it is degraded. Generally speaking, if a sufficient quantity of reasonable quality DNA can be extracted from a crime-scene sample, no matter what the

---

[83] *See, e.g.*, State v. Fleming, 698 A.2d 503, 507 (Me. 1997) (taking judicial notice that "the overall theory and techniques of DNA profiling [are] scientifically reliable if conducted in accordance with appropriate laboratory standards and controls"); State v. Davis, 814 S.W.2d 593, 602 (Mo. 1991); People v. Castro, 545 N.Y.S.2d 985, 987 (N.Y. Sup. Ct. 1989); cases cited, NRC II, *supra* note 1, at 172 n.15.

[84] NRC I, *supra* note 1, at 38.

[85] Office of Tech. Assessment, Genetic Witness: Forensic Uses of DNA Tests 59 (1990). The 1992 NRC report therefore recommends 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 I, *supra* note 1, at 149. The 1996 report reiterates the conclusion that "[t]he techniques of DNA typing [including RFLP analysis] are fully recognized by the scientific community." NRC II, *supra* note 1, at 50. It insists that "[t]he state of the profiling technology and the methods for estimating frequencies and related statistics have progressed to the point where the admissibility of properly collected and analyzed DNA data should not be in doubt." *Id.* at 36.

[86] *See supra* note 69.

nature of the sample, DNA typing can be done without problem. Thus, DNA typing has been performed successfully on old blood stains, semen stains, vaginal swabs, hair, bone, bite marks, cigarette butts, urine, and fecal material. This section discusses what constitutes sufficient quantity and reasonable quality in the contexts of PCR-based genetic typing[87] and VNTR analysis by Southern blotting.[88] Complications due to contaminants and inhibitors also are discussed. Finally, the question of whether the sample contains DNA from two or more contributors is considered.

A. Did the Sample Contain Enough DNA?

The amount of DNA in a cell varies from organism to organism. The DNA in the chromosomes of a human cell, for example, is about two thousand times greater than that in a typical bacterium.[89] Within an organism, however, DNA content is constant from cell to cell. Thus, a human hair root cell contains the same amount of DNA as a white cell in blood or a buccal cell in saliva.[90] Amounts of DNA present in some typical kinds of samples are indicated in Table A-2 of the Appendix. These vary from a trillionth or so of a gram for a hair shaft to several millionths of a gram for a post-coital vaginal swab. RFLP typing requires a much larger sample of DNA than PCR-based typing. As a practical matter, RFLP analysis requires a minimum of about 50 billionths of a gram of relatively non-degraded DNA,[91] while most PCR test protocols recommend samples on the order of one to five billionths of a gram for optimum yields.[92] Thus, PCR tests can be applied to samples containing ten to five hundred-fold less nuclear

---

[87] *See supra* § III.

[88] *See supra* § IV.

[89] A human egg or sperm cell contains half as much DNA; hence, the haploid human genome is about one thousand times larger than the typical bacterial genome.

[90] A human cell contains about six picograms of DNA. (A picogram (pg) is one trillionth (1/1,000,000,000,000) of a gram.) Sperm cells constitute a special case, for they contain half a genetic complement (that which the father passes along to an offspring) and so contain half as much DNA (about 3 pg). The 3 pg of DNA varies from sperm cell to sperm cell because each such cell has a randomly drawn half of the man's chromosomes. The DNA in a semen sample contains many of these cells; being a mixture of the many combinations, it contains all the man's alleles.

[91] RFLP analysis has been performed successfully on smaller amounts of DNA but at a cost of longer autoradiograph exposure times. From the standpoint of the reliability of the typing, what is important is the strength of the banding pattern on the autoradiograph or lumigraph. Threshold amounts of DNA may result in weak bands, and some bands could be missed because they are too weak to be observed.

[92] Although the polymerase chain reaction can amplify DNA from the nucleus of a single cell, chance effects may result in one allele being amplified much more than another. To avoid preferential amplification, a lower limit of about ten to fifteen cells' worth of DNA has been determined to give balanced amplification. PCR tests for nuclear genes are designed to yield no detectable product for samples containing less than about 20 cell equivalents (100–200 pg) of DNA. This result is achieved by limiting the number of amplification cycles.

DNA than that required for RFLP tests.[93] Moreover, mitochondrial DNA analysis works reliably with DNA from even fewer cells. As noted in section II, cells contain only one nucleus, but hundreds of mitochondria. Consequently, even though there rarely is sufficient DNA in a hair shaft to allow testing with nuclear DNA markers, the mitochondrial DNA often can be analyzed.[94]

These sample-size requirements help determine the approach to be taken for a DNA typing analysis. Samples which, from experience, are expected to contain at least fifty to one hundred billionths of a gram of DNA typically are subjected to a formal DNA extraction followed by characterization of the DNA for quantity and quality. This characterization typically involves gel electrophoresis of a small portion of the extracted DNA. This test, however, does not distinguish human from non-human DNA. Since the success of DNA typing tests depends on the amount of human DNA present, it may be desirable to test for the amount of human DNA in the extract.[95] For samples that typically contain small amounts of DNA, the risk of DNA loss during extraction may dictate the use of a different extraction procedure.[96]

Whether a particular sample contains enough human DNA to allow typing cannot always be predicted in advance. The best strategy is to try; if a result is obtained, and if the controls (samples of known DNA and blank samples) have behaved properly, then the sample had enough DNA.

B. Was the Sample of Sufficient Quality?

The primary determinant of DNA quality for forensic analysis is the extent to which the long DNA molecules are intact. Within the cell nucleus, each molecule of DNA extends for millions of base pairs. Outside the cell, DNA spontaneously degrades into smaller fragments at a rate that depends on temperature,

---

[93] The great sensitivity of PCR for the detection of DNA, even under these "safe" conditions, is illustrated by the successful genetic typing of DNA extracted from fingerprints. Roland A.H. van Oorschot & Maxwell K. Jones, *DNA Fingerprints from Fingerprints*, 387 Nature 767 (1997).

[94] *E.g.*, M.R. Wilson et al., *Extraction, PCR Amplification, and Sequencing of Mitochondrial DNA from Human Hair Shafts*, 18 Biotechniques 662 (1995). Of course, mitochondrial DNA analysis can be done with other sources of mtDNA.

[95] This test entails measuring the amount of a human-specific DNA probe that binds to the DNA in the extract. This test is particularly important in cases where the sample extract contains a mixture of human and microbial DNA. Vaginal swabs, for example, are expected to contain microbial DNA from the vaginal flora as well as human DNA from the female and sperm donor. Similarly, samples that have been damp for extended periods of time often contain significant microbial contamination; indeed, in some cases, little or no human DNA can be detected even though the extract contains significant amounts of DNA.

[96] Boiling a sample for a few minutes releases DNA, and this DNA is used directly for PCR without first characterizing the DNA. The boiling step usually is conducted in the presence of a resin that adsorbs inhibitors of PCR.

exposure to oxygen, and, most importantly, the presence of water.[97] In dry biological samples, protected from air, and not exposed to temperature extremes, DNA degrades very slowly. In fact, the relative stability of DNA has made it possible to extract usable DNA from samples hundreds to thousands of years old.[98]

RFLP analysis requires relatively non-degraded DNA, and testing DNA for degradation is a routine part of the protocol for VNTR analysis. In RFLP testing, a restriction enzyme cuts long sequences of DNA into smaller fragments. If the DNA is randomly fragmented into very short pieces to begin with, electrophoresis and Southern blotting will produce a smear of fragments rather than a set of well-separated bands.[99]

In contrast, PCR-based tests are relatively insensitive to degradation. Testing has proved effective with old and badly degraded material such as the remains of the Tsar Nicholas family (buried in 1918, recovered in 1991)[100] and the Tyrolean Ice Man (frozen for some 5,000 years).[101] The extent to which degradation affects a PCR-based test depends on the size of the DNA segment to be amplified. For example, in a sample in which the bulk of the DNA has been degraded to fragments well under 1,000 base pairs in length, it may be possible to amplify a 100 base-pair sequence, but not a 1,000 base-pair target. Consequently, the shorter alleles may be detected in a highly degraded sample, but the larger ones may be missed.[102] As with RFLP analysis, this possibility would have to be considered in the statistical interpretation of the result.

---

[97] Other forms of chemical alteration to DNA are well studied, both for their intrinsic interest and because chemical changes in DNA are a contributing factor in the development of cancers in living cells. Most chemical modification has little effect on RFLP analysis. Some forms of DNA modification, such as that produced by exposure to ultraviolet radiation, inhibit the amplification step in PCR-based tests, while other chemical modifications appear to have no effect. George F. Sensabaugh & Cecilia von Beroldingen, *The Polymerase Chain Reaction: Application to the Analysis of Biological Evidence, in* Forensic DNA Technology 63 (Mark A. Farley & James J. Harrington eds., 1991).

[98] This has resulted in a specialized field of inquiry dubbed "ancient DNA." Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical, and Forensic Specimens (Bernd Herrmann & Susanne Hummel eds., 1993); Svante Paabo, *Ancient DNA: Extraction, Characterization, Molecular Cloning, and Enzymatic Amplification*, 86 Proc. Nat'l Acad. Sci. USA 1939 (1989).

[99] Practically speaking, RFLP analysis can yield interpretable results if the bulk of the DNA in a sample exceeds 20,000–30,000 base pairs in length. Partial degradation of the DNA can result in the weakening or loss of the signal from large restriction fragments. This effect is usually evident from the appearance of the restriction fragment banding pattern. Another indication of degradation is smearing in the background of the banding pattern. If there is evidence that degradation has affected the banding pattern, the statistical interpretation of a match should account for the possibility that some allelic bands might not have been detected.

[100] Gill et al., *supra* note 47.

[101] Oliva Handt et al., *Molecular Genetic Analyses of the Tyrolean Ice Man*, 264 Science 1775 (1994).

[102] For example, typing at a genetic locus such as D1S80, for which the target allelic sequences range in size from 300 to 850 base pairs, may be affected by the non-amplification of the largest alleles ("allelic dropout").

Allelic dropout of this sort does not seem to be a problem for STR loci, presumably because the size differences between alleles at a locus are so small (typically no more than 50 base pairs). If there is a degradation effect on STR typing, it is "locus dropout": in cases involving severe degradation, loci yielding smaller PCR products (less than 180 base pairs) tend to amplify more efficiently than loci yielding larger products (greater than 200 base pairs).[103]

Surprising as it may seem, DNA can be exposed to a great variety of environmental insults without any effect on its capacity to be typed correctly. Exposure studies have shown that contact with a variety of surfaces, both clean and dirty, and with gasoline, motor oil, acids, and alkalis either have no effect on DNA typing or, at worst, render the DNA untypable.[104]

Although contamination with microbes generally does little more than degrade the human DNA,[105] other problems sometimes can occur with both RFLP[106] and PCR-based analyses.[107] Nevertheless, there are procedures that identify or avoid these anomalies.[108] Therefore, the validation of DNA typing

---

[103] J.P. Whitaker et al., *Short Tandem Repeat Typing of Bodies from a Mass Disaster: High Success Rate and Characteristic Amplification Patterns in Highly Degraded Samples*, 18 Biotechniques 670 (1995).

[104] 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); Roland A.H. van Oorschot et al., *HUMTH01 Validation Studies: Effect of Substrate, Environment, and Mixtures*, 41 J. Forensic Sci. 142 (1996). Most of the effects of environmental insult readily can be accounted for in terms of basic DNA chemistry. For example, some agents produce degradation or damaging chemical modifications. Other environmental contaminants inhibit restriction enzymes or PCR. (This effect sometimes can be reversed by cleaning the DNA extract to remove the inhibitor.) But environmental insult does not result in the selective loss of an allele at a locus or in the creation of a new allele at that locus.

[105] Michael B.T. Webb et al., *Microbial DNA Challenge Studies of Variable Number Tandem Repeat (VNTR) Probes Used for DNA Profiling Analysis*, 38 J. Forensic Sci. 1172 (1993).

[106] Autoradiograms sometimes show many bands that line up with the molecular weight sizing ladder bands. (The "ladder" is a set of DNA fragments of known lengths that are placed by themselves in one or more lanes of the gel. The resulting set of bands provides a benchmark for determining the weights of the unknown bands in the samples.) These extra bands can result from contamination of the sample DNA with ladder DNA at the time the samples are loaded onto the electrophoresis gel. Alternatively, the original sample may have been contaminated with a microbe infected with lambda phage, the virus that is used for the preparation of the sizing ladder.

[107] Although PCR primers designed to amplify human gene sequences would not be expected to recognize microbial DNA sequences, much less amplify them, such amplification has been reported with the D1S80 typing system. A. Fernández-Rodríguez et al., *Microbial DNA Challenge Studies of PCR-based Systems in Forensic Genetics*, *in* 6 Advances in Forensic Haemogenetics 177 (A. Carracedo et al., eds., 1996).

[108] Whatever the explanation for the extra sizing bands mentioned *supra* note 106, the lambda origin of the bands can be demonstrated by an additional probing with the ladder probe alone or with a human specific probe without the ladder probe. Likewise, the spurious PCR products observed by Fernández-Rodríguez et al., *supra* note 107, can be differentiated from the true human PCR products, and the same authors have described a modification to the D1S80 typing system that removes all question of the non-human origin of the spurious PCR products. A. Fernández-Rodríguez et al., *D1S80 Typing in Casework: A Simple Strategy to Distinguish Non-specific Microbial PCR Products from Human Alleles*, 7 Progress in Forensic Genetics 18 (1998).

systems should include tests for interference with a variety of microbes to see if artifacts occur; if artifacts are observed, then control tests should be applied to distinguish between the artifactual and the true results.

C. Does a Sample Contain DNA from More Than One Person?

DNA from a single individual can have no more than two alleles at each locus. This follows from the fact that individuals inherit chromosomes in pairs, one from each parent.[109] An individual who inherits the same allele from each parent (a homozygote) can contribute only that one allele to a sample, and an individual who inherits a different allele from each parent (a heterozygote) will contribute those two alleles.[110] Finding three or more alleles at a locus therefore indicates a mixture of DNA from more than one person.[111]

Some kinds of samples, such as post-coital vaginal swabs and blood stains from scenes where several persons are known to have bled, are expected to be mixtures. Sometimes, however, the first indication the sample has multiple contributors comes from the DNA testing. The chance of detecting a mixture by finding extra alleles depends on the proportion of DNA from each contributor as well as the chance that the contributors have different genotypes at one or more loci. As a rule, a minor contributor to a mixture must provide at least 5% of the DNA for the mixture to be recognized.[112] In addition, the various contributors must have some different alleles. The chance that multiple contributors will differ at one or more locus increases with the number of loci tested and the genetic diversity at each locus. Unless many loci are examined, genetic markers with low to moderate diversities do not have much power to detect multiple contributors. Genetic markers that are highly polymorphic are much better at detecting mixtures. Thus, STRs and especially VNTRs are sensitive to mixtures.

---

[109] *See supra* § II.

[110] Loci on the sex chromosomes constitute a special case. Females have two X chromosomes, one from each parent; as with loci on the other chromosomes, they can be either homozygous or heterozygous at the X-linked loci. Males, on the other hand, have one X and one Y chromosome; hence, they have only one allele at the X-linked loci and one allele at the Y-linked loci. In cases of trisomy, such as XXY males, multiple copies of loci on the affected chromosome will be present, but this condition is rare and often lethal.

[111] On very rare occasions, an individual exhibits a phenotype with three alleles at a locus. This can be the result of a chromosome anomaly (such as a duplicated gene on one chromosome or a mutation). A sample from such an individual is usually easily distinguished from a mixed sample. The three-allele variant is seen at only the affected locus, whereas with mixtures, more than two alleles typically are evident at several loci.

[112] With RFLP testing, alleles from a contributor of as little as one percent can be detected at the price of overexposing the pattern from the major contributor. Studies in which DNA from different individuals is combined in differing proportions show that the intensity of the bands reflects the proportions of the mixture. Thus, if bands in a crime-scene sample have different intensities, it may be possible to assign alleles to major and minor contributors. However, if bands are present in roughly equal

VI. Laboratory Performance

A. Quality Control and Assurance

DNA profiling is valid and reliable, but confidence in a particular result depends on the quality control and quality assurance procedures in the laboratory. Quality control refers to measures to help ensure that a DNA-typing result (and its interpretation) meets a specified standard of quality. Quality assurance refers to monitoring, verifying, and documenting laboratory performance.[113] A quality assurance program helps demonstrate that a laboratory is meeting its quality control objectives and thus justifies confidence in the quality of its product.

Professional bodies within forensic science have described procedures for quality assurance. Guidelines have been prepared by two FBI-appointed groups—the Technical Working Group on DNA Analysis Methods (TWGDAM)[114] and the DNA Advisory Board (DAB).[115] The DAB also has encouraged forensic DNA laboratories to seek accreditation,[116] and at least two states require forensic DNA laboratories to be accredited.[117] The American Society of Crime Laboratory Directors–Laboratory Accreditation Board (ASCLD–LAB) accredits forensic laboratories.[118]

---

proportions, this allocation cannot be made, and the statistical interpretation of the observed results must include all possible combinations. *See infra* note 220.

[113] For general descriptions of quality assurance programs, see NRC II, *supra* note 1, at ch. 3 ("Ensuring High Standards of Laboratory Performance"); NRC I, *supra* note 1, at ch. 4.

[114] *See* Technical Working Group on DNA Analysis Methods, *Guidelines for a Quality Assurance Program for DNA Analysis*, 22 Crime Laboratory Dig. 21 (1995) [hereinafter TWGDAM Guidelines], 18 Crime Laboratory Dig. 44 (1991).

[115] *See* Federal Bureau of Investigation, Quality Assurance Standards for Forensic DNA Testing Laboratories, July 15, 1998 [hereinafter DAB Standards]; *see also Recommendations of the DNA Commission of the International Society for Forensic Haemogenetics Relating to the Use of PCR-based Polymorphisms,* 64 Vox Sang. 124 (1993); *1991 Report Concerning Recommendations of the DNA Commission of the International Society for Forensic Haemogenetics Relating to the Use of DNA Polymorphism*, 63 Vox Sang. 70 (1992).

Under the DNA Identification Act of 1994, Pub. L. No. 103-322, 108 Stat. 2065 (codified at 42 U.S.C. § 13701 (1994)), to qualify for federal laboratory improvement funds, a forensic DNA laboratory must meet the quality assurance standards recommended by the DAB and issued by the director of the FBI. The DAB membership includes molecular geneticists, population geneticists, an ethicist, and representatives from federal, state, and local forensic DNA laboratories, private sector DNA laboratories, the National Institute of Standards and Technology, and the judiciary. Its recommendations closely follow the 1995 TWGDAM Guidelines.

[116] DAB Standards, *supra* note 115, at 1 (preface).

[117] *N.Y. Executive Law* § 995-b (McKinney 1999); Cal. DNA and Forensic Identification Data Base and Data Bank Act of 1998, Cal. Penal Code § 297 (West 1999).

[118] *See* American Society of Crime Laboratory Directors—Laboratory Accreditation Board, ASCLD-LAB Accreditation Manual, Jan. 1997. As of mid-1998, ASCLD-LAB had accredited laboratories in Australia, New Zealand, and Hong Kong as well as laboratories in the United States and Canada. The ASCLD-LAB accreditation program does not allow laboratories to obtain accreditation only for particular services—a laboratory seeking accreditation must qualify for the full range of services it offers. This constraint has slowed some forensic DNA labs from seeking accreditation. As an interim solution,

Documentation. The quality assurance guidelines promulgated by TWGDAM, the DAB, and ASCLD-LAB call for laboratories to document laboratory organization and management, personnel qualifications and training, facilities, evidence control procedures, validation of methods and procedures, analytical procedures, equipment calibration and maintenance, standards for case documentation and report writing, procedures for reviewing case files and testimony, proficiency testing, corrective actions, audits, safety programs, and review of sub-contractors. Of course, maintaining even such extensive documentation and records does not guarantee the correctness of results obtained in any particular case. Errors in analysis or interpretation might occur as a result of a deviation from an established procedure, analyst misjudgement, or an accident. Although case-review procedures within a laboratory should be designed to detect errors before a report is issued, it is always possible that some incorrect result will slip through. Accordingly, determination that a laboratory maintains a strong quality assurance program does not eliminate the need for case-by-case review.

Validation. The validation of procedures is central to quality assurance. "Developmental" validation is undertaken to determine the applicability of a new test to crime-scene samples; it defines conditions that give reliable results and identifies the limitations of the procedure. For example, a new genetic marker being considered for use in forensic analysis will be tested to determine if it can be typed reliably in both fresh samples and in samples typical of those found at crime scenes. The validation would include testing samples originating from different tissues—blood, semen, hair, bone, samples containing degraded DNA, samples contaminated with microbes, samples containing DNA mixtures, and so on. Developmental validation of a new marker also includes the generation of population databases and the testing of allele and genotype distributions for independence. Developmental validation normally results in publication in the scientific literature, but a new procedure can be validated in multiple laboratories well ahead of publication.

"Internal" validation, on the other hand, involves the verification by a laboratory that it can reliably perform an established procedure that already has undergone developmental validation. Before adopting a new procedure, the laboratory should verify its ability to use the system in a proficiency trial.

Both forms of validation build on the accumulated body of knowledge and experience. Thus, some aspects of validation testing need be repeated only to the extent required to verify that previously established principles apply. One

---

the National Forensic Science Technology Center (NFSTC) has an agreement with ASCLD-LAB to perform certification audits on DNA sections of laboratories for compliance with DAB and ASCLD-LAB standards; this service is available to private sector DNA laboratories as well as government laboratories.

need not validate the principle of the internal combustion engine every time one brings out a new model of automobile.

Proficiency Testing. Proficiency testing in forensic genetic testing is designed to ascertain whether an analyst can correctly determine genetic types in a sample the origin of which is unknown to the analyst but is known to a tester. Proficiency is demonstrated by making correct genetic typing determinations in repeated trials, and not by opining on whether the sample originated from a particular individual. Proficiency tests also require laboratories to report random-match probabilities to determine if proper calculations are being made.

An internal proficiency trial is conducted within a laboratory. One person in the laboratory prepares the sample and administers the test to another person in the laboratory. An external trial is one in which the test sample originates from outside the laboratory—from another laboratory, a commercial vendor, or a regulatory agency. In a declared (or open) proficiency trial the analyst knows the sample is a proficiency sample. In contrast, in a blind (or more properly "full-blind") trial, the sample is submitted so that the analyst does not recognize it as a proficiency sample.[119] It has been argued that full-blind trials provide a better indication of proficiency because the analyst will not give the trial sample any special attention.[120] On the other hand, full-blind proficiency trials for forensic DNA analysis entail considerably more organizational effort and expense than open proficiency trials. Obviously, the "evidence" samples prepared for the trial have to be sufficiently realistic that the laboratory does not suspect the legitimacy of the submission. A police agency and prosecutor's office have to submit the "evidence" and respond to laboratory inquiries with information about the "case." Finally, the genetic profile from a proficiency test must not be entered into regional and national databases.[121]

---

[119] There is potential confusion over nomenclature with regard to open and blind trials. All proficiency tests are blind in the sense that the analyst does not know the composition of the test sample. In some disciplines, any trial in which the analyst receives "unknowns" from a tester is referred to as a blind trial. With regard to proficiency testing in the forensic area, however, the convention is to distinguish "open" and "blind" trials as described here.

[120] *See, e.g.*, Scheck, *supra* note 69, at 1980. Another argument for the full-blind trial is that it tests a broader range of laboratory operations, from submission of the evidence to the laboratory through the analysis and interpretation stages to the reporting out to the submitting agency. However, these aspects of laboratory operations also can be evaluated, at much less cost, by mechanisms such as laboratory audits and random review of case files.

[121] The feasibility of mounting a national, full-blind proficiency trial program is under study as a part of the DNA Identification Act of 1994, Pub. L. No. 103-322, 108 Stat. 2065 (codified at 42 U.S.C. § 13701 (1994)). The results of this study, funded by the National Institute of Justice, are to be reported to the DAB with subsequent recommendations made to the director of the FBI.

The DAB recommends that every analyst undergo regular external, open proficiency testing[122] and that the laboratory take "corrective action whenever proficiency testing discrepancies [or] casework errors are detected."[123] Certification by the American Board of Criminalistics as a specialist in forensic biology DNA analysis requires one proficiency trial per year. Accredited laboratories must maintain records documenting compliance with required proficiency test standards.[124]

B. Handling Samples

Sample mishandling, mislabeling, or contamination, whether in the field or in the laboratory, is more likely to compromise a DNA analysis than an error in genetic typing. For example, a sample mixup due to mislabeling reference blood samples taken at the hospital could lead to incorrect association of crime-scene samples to a reference individual or to incorrect exclusions. Similarly, packaging two items with wet blood stains into the same bag could result in a transfer of stains between the items, rendering it difficult or impossible to determine whose blood was originally on each item. Contamination in the laboratory may result in artifactual typing results or in the incorrect attribution of a DNA profile to an individual or to an item of evidence. Accordingly, it is appropriate to look at the procedures that have been prescribed and implemented to guard against such error.

Mislabeling or mishandling can occur when biological material is collected in the field, when it is transferred to the laboratory, when it is in the analysis stream in the laboratory,[125] when the analytical results are recorded, or when the recorded results are transcribed into a report. Mislabeling and mishandling can happen with any kind of physical evidence and are of great concern in all fields of forensic science. Because forensic laboratories often have little or no control over the handling of evidence prior to its arrival in the laboratory, checkpoints should be established to detect mislabeling and mishandling along the line of

---

[122] Standard 13.1 specifies that these tests are to be performed at least as frequently as every 180 days. DAB Standards, *supra* note 115, at 16. TWGDAM recommended two open proficiency tests per year per analyst. TWGDAM Guidelines, *supra* note 114.

[123] DAB Standards, *supra* note 115, at 17 (standard 14.1).

[124] Proficiency test results from laboratories accredited by ASCLD-LAB are reported also to an ASCLD-LAB Proficiency Review Committee. The committee independently reviews test results and verifies compliance with accreditation requirements. ASCLD-LAB specifies the vendors whose proficiency tests it accepts for accreditation purposes. Since accreditation can be suspended or withdrawn by unacceptable proficiency trial performance, the proficiency test vendors must meet high standards with respect to test-sample preparation and documentation. Yet, in some instances vendors have provided mislabeled or contaminated test samples. *See* TWGDAM & ASCLD-LAB Proficiency Review Comm., *Guidelines for DNA Proficiency Test Manufacturing and Reporting*, 21 Crime Laboratory Dig. 27–32 (1994).

[125] *E.g.*, United States v. Cuff, 37 F. Supp. 2d 279, 283 (S.D.N.Y. 1999).

evidence flow.[126] Investigative agencies should have guidelines for evidence collection and labeling so that a chain of custody is maintained. Similarly, there should be guidelines, produced with input from the laboratory, for handling biological evidence in the field. These principles remain the same as in the pre-DNA era.[127]

TWGDAM guidelines and DAB recommendations require documented procedures to ensure sample integrity and to avoid sample mixups, labeling errors, recording errors, and the like. They also mandate case review to identify inadvertent errors before a final report is released. Finally, laboratories must retain, when feasible, portions of the crime-scene samples and extracts to allow reanalysis.[128] However, retention is not always possible. For example, retention of original items is not to be expected when the items are large or immobile (for example, a wall or sidewalk). In such situations, a swabbing or scraping of the stain from the item would typically be collected and retained. There also are situations where the sample is so small that it will be consumed in the analysis.[129]

Assuming appropriate chain-of-custody and evidence-handling protocols are in place, the critical question is whether there are deviations in the particular case. This may require a review of the total case documentation as well as the laboratory findings.[130]

As the 1996 NRC Report emphasizes, an important safeguard against error due to mislabeling and mishandling is the opportunity to retest original evidence items or the material extracted from them.[131] Should mislabeling or mishandling have occurred, reanalysis of the original sample and the intermediate extracts should detect not only the fact of the error but also the point at which

---

[126] NRC II, *supra* note 1, at 80–82.

[127] Samples (particularly those containing wet stains) should not be packaged together, and samples should be dried or refrigerated as soon as possible. Storage in the dry state and at low temperatures stabilizes biological material against degradation. George F. Sensabaugh, *Biochemical Markers of Individuality, in* 1 Forensic Science Handbook 338, 385 (Richard Saferstein ed., 1982). The only precaution to have gained force in the DNA era is that evidence items should be handled with gloved hands to protect against handling contamination and inadvertent sample-to-sample transfers.

[128] Forensic laboratories have a professional responsibility to preserve retained evidence so as to minimize degradation. *See* TWGDAM Guidelines, *supra* note 114, at 30 para. 6.3. Furthermore, failure to preserve potentially exculpatory evidence has been treated as a denial of due process and grounds for suppression. People v. Nation, 604 P.2d 1051 (Cal. 1980). In *Arizona v. Youngblood*, 488 U.S. 51 (1988), however, the Supreme Court held that a police agency's failure to preserve evidence not known to be exculpatory does not constitute a denial of due process unless "bad faith" can be shown.

[129] When small samples are involved, whether it is necessary to consume the entire sample is a matter of scientific judgment.

[130] Such a review is best undertaken by someone familiar with police procedures, forensic DNA analysis, and forensic laboratory operations. Case review by an independent expert should be held to the same scientific standard as the work under review. Any possible flaws in labeling or in evidence handling should be specified in detail, with consideration given to the consequence of the possible error.

[131] NRC II, *supra* note 1, at 81.

it occurred. It is even possible in some cases to detect mislabeling at the point of sample collection if the genetic typing results on a particular sample are inconsistent with an otherwise consistent reconstruction of events.[132]

Contamination describes any situation in which foreign material is mixed with a sample of DNA. Contamination by non-biological materials, such as gasoline or grit, can cause test failures, but they are not a source of genetic typing errors. Similarly, contamination with non-human biological materials, such as bacteria, fungi, or plant materials, is generally not a problem. These contaminants may accelerate DNA degradation, but they do not contribute spurious genetic types.[133]

Consequently, the contamination of greatest concern is that resulting from the addition of human DNA. This sort of contamination can occur three ways:[134]

1. The crime-scene samples by their nature may contain a mixture of fluids or tissues from different individuals. Examples include vaginal swabs collected as sexual assault evidence[135] and blood stain evidence from scenes where several individuals shed blood.[136]
2. The crime-scene samples may be inadvertently contaminated in the course of sample handling in the field or in the laboratory. Inadvertent contamination of crime-scene DNA with DNA from a reference sample could lead to a false inclusion.[137]

---

[132] For example, a mislabeling of husband and wife samples in a paternity case might result in an apparent maternal exclusion, a very unlikely event. The possibility of mislabeling could be confirmed by testing the samples for gender and ultimately verified by taking new samples from each party under better controlled conditions.

[133] Validation of new genetic markers includes testing on a variety of non-human species. The probes used in VNTR analysis and the PCR-based tests give results with non-human primate DNA samples (apes and some monkeys). This is not surprising given the evolutionary proximity of the primates to humans. As a rule, the validated test systems give no results with DNA from animals other than primates, from plants, or from microbes. An exception is the reaction of some bacterial DNA samples in testing for the marker D1S80. Fernández-Rodríguez et al., *supra* note 107. However, this could be an artifact of the particular D1S80 typing system, since other workers have not been able to replicate fully their results, and an alternative D1S80 typing protocol gave no spurious results. Shamsah Ebrahim et al., Investigation of the Specificity of STR and D1S80 Primers on Microbial DNA Samples, Presentation B84, 50th Annual Meeting of the American Academy of Forensic Sciences, San Francisco (Feb. 1998).

[134] NRC II, *supra* note 1, at 82–84; NRC I, *supra* note 1, at 65–67; George F. Sensabaugh & Edward T. Blake, *DNA Analysis in Biological Evidence: Applications of the Polymerase Chain Reaction*, *in* 3 Forensic Science Handbook 416, 441 (Richard Saferstein ed., 1993); Sensabaugh & von Beroldingen, *supra* note 97, at 63, 77.

[135] These typically contain DNA in the semen from the assailant and in the vaginal fluid of the victim. The standard procedure for analysis allows the DNA from sperm to be separated from the vaginal epithelial cell DNA. It is thus possible not only to recognize the mixture but also to assign the DNA profiles to the different individuals.

[136] Such mixtures are detected by genetic typing that reveals profiles of more than one DNA source. *See supra* § V.C.

[137] This source of contamination is a greater concern when PCR-based typing methods are to be used due to the capacity of PCR to detect very small amounts of DNA. However, experiments de-

3. Carry-over contamination in PCR-based typing can occur if the amplification products of one typing reaction are carried over into the reaction mix for a subsequent PCR reaction. If the carry-over products are present in sufficient quantity, they could be preferentially amplified over the target DNA.[138] The primary strategy used in most forensic laboratories to protect against carry-over contamination is to keep PCR products away from sample materials and test reagents by having separate work areas for pre-PCR and post-PCR sample handling, by preparing samples in controlled air-flow biological safety hoods, by using dedicated equipment (such as pipetters) for each of the various stages of sample analysis, by decontaminating work areas after use (usually by wiping down or by irradiating with ultraviolet light), and by having a one-way flow of sample from the pre-PCR to post-PCR work areas.[139] Additional protocols are used to detect any carry-over contamination.[140]

In the end, whether a laboratory has conducted proper tests and whether it conducted them properly depends both on the general standard of practice and on the questions posed in the particular case. There is no universal checklist, but the selection of tests and the adherence to the correct test procedures can be reviewed by experts and by reference to professional standards, such as the TWGDAM and DAB guidelines.

---

signed to introduce handling contamination into samples have been unsuccessful. *See* Catherine Theisen Comey & Bruce Budowle, *Validation Studies on the Analysis of the HLA DQa Locus Using the Polymerase Chain Reaction*, 36 J. Forensic Sci. 1633 (1991). Of course, it remains important to have evidence-handling procedures to safeguard against this source of contamination. Police agencies should have documented procedures for the collection, handling, and packaging of biological evidence in the field and for its delivery to the laboratory that are designed to minimize the chance of handling contamination. Ideally, these procedures will have been developed in coordination with the laboratory, and training in the use of these procedures will have been provided. Similarly, laboratories should have procedures in place to minimize the risk of this kind of contamination. *See* DAB Standards, *supra* note 115; TWGDAM Guidelines, *supra* note 114. In particular, these procedures should specify the safeguards for keeping evidence samples separated from reference samples.

[138] Carry-over contamination is not an issue in RFLP analysis, which involves no amplification steps.

[139] Some laboratories with space constraints separate pre-PCR and post-PCR activities in time rather than space. The other safeguards can be used as in a space-separated facility.

[140] Standard protocols include the amplification of blank control samples—those to which no DNA has been added. If carry-over contaminants have found their way into the reagents or sample tubes, these will be detected as amplification products. Outbreaks of carry-over contamination can also be recognized by monitoring test results. Detection of an unexpected and persistent genetic profile in different samples indicates a contamination problem. When contamination outbreaks are detected, appropriate corrective actions should be taken, and both the outbreak and the corrective action should be documented. *See* DAB Standards, *supra* note 115; TWGDAM Guidelines, *supra* note 114.

VII. Interpretation of Laboratory Results

The results of DNA testing can be presented in various ways. With discrete allele systems, it is natural to speak of "matching" and "non-matching" profiles. If the genetic profile obtained from the biological sample taken from the crime scene or the victim (the "trace evidence sample") matches that of a particular individual, then that individual is included as a possible source of the sample. But other individuals also might possess a matching DNA profile. Accordingly, the expert should be asked to provide some indication of how significant the match is. If, on the other hand, the genetic profiles are different, then the individual is excluded as the source of the trace evidence. Typically, proof tending to show that the defendant is the source incriminates the defendant, while proof that someone else is the source exculpates the defendant.[141]

This section elaborates on these ideas, indicating issues that can arise in connection with an expert's testimony interpreting the results of a DNA test.

A. Exclusions, Inclusions, and Inconclusive Results

When the DNA from the trace evidence clearly does not match the DNA sample from the suspect, the DNA analysis demonstrates that the suspect's DNA is not in the forensic sample. Indeed, if the samples have been collected, handled, and analyzed properly, then the suspect is excluded as a possible source of the DNA in the forensic sample. Even a single allele that cannot be explained as a laboratory artifact or other error can suffice to exclude a suspect.[142] As a practical matter, such exclusionary results normally would keep charges from being filed against the excluded suspect.[143]

In some cases, however, DNA testing is inconclusive, in whole or in part. The presence or absence of a discrete allele can be in doubt, or the existence or location of a VNTR band may be unclear.[144] For example, when the trace evidence sample is extremely degraded, VNTR profiling might not show all the

---

[141] Whether being the source of the forensic sample is incriminating depends on other facts in the case. *See infra* note 155. Likewise, whether someone else being the source is exculpatory depends on the circumstances. For example, a suspect who might have committed the offense without leaving the trace evidence sample still could be guilty. In a rape case with several rapists, a semen stain could fail to incriminate one assailant because insufficient semen from that individual is present in the sample.

[142] Due to heteroplasmy, a single sequence difference in mtDNA samples would not be considered an exclusion. *See supra* note 46. With testing at many polymorphic loci, however, it would be unusual to find two unrelated individuals whose DNA matches at all but one locus.

[143] *But see* State v. Hammond, 604 A.2d 793 (Conn. 1992).

[144] *E.g.*, State v. Fleming, 698 A.2d 503, 506 (Me. 1997) ("The fourth probe was declared uninterpretable."); People v. Leonard, 569 N.W.2d 663, 666–67 (Mich. Ct. App. 1997) ("There was a definite match of defendant's DNA on three of the probes, and a match on the other two probes could not be excluded."). In some cases, experts have disagreed as to whether extra bands represented a mixture or resulted from partial digestion of the forensic sample. *E.g.*, State v. Marcus, 683 A.2d 221 (N.J. Super. Ct. App. Div. 1996).

alleles that would be present in a sample with more intact DNA. If the quantity of DNA to be amplified for sequence-specific tests is too small, the amplification might not yield enough product to give a clear signal. Thus, experts sometimes disagree as to whether a particular band is visible on an autoradiograph or whether a dot is present on a reverse dot blot.[145]

Furthermore, even when RFLP bands are clearly visible, the entire pattern of bands can be displaced from its true location in a systematic way (a phenomenon known as band-shifting).[146] Recognizing this phenomenon, analysts might deem some seemingly matching patterns as inconclusive.[147]

---

[145] *E.g.*, People v. Leonard, 569 N.W.2d 663, 667 (Mich. Ct. App.) (prosecution's academic expert concluded that there was a match at all bands rather than just the three that the state laboratory considered to match), *app. denied*, 570 N.W.2d 659 (Mich. 1997); State v. Jobe, 486 N.W.2d 407 (Minn. 1992) (one FBI examiner found a match on the basis of two of four probes, with the other two being inconclusive; another examiner found no match; another scientist called the profiles a "very, very, very significant match"); State v. Marcus, 683 A.2d 221 (N.J. Super. Ct. App. 1996) (defendant's academic expert questioned the results of one probe); State v. Gabriau, 696 A.2d 290, 292 n.3 (R.I. 1997) ("According to [a university geneticist] the laboratory technician had not considered two loci as matches where he himself would have."). In *United States v. Perry*, No. CR 91-395-SC (D.N.M. Sept. 7, 1995), the district court found a defense expert's suggestions of "lab technicians manipulating samples to achieve false matches" and of an analyst's sizing a band "when no band existed" to be "particularly unprincipled," "the stuff of mystery novels, not science." But bona fide disagreements of this sort would certainly go to the weight of the evidence and might bear on its admissibility through Federal Rule of Evidence 403.

It also can be argued that such disagreements pertain to admissibility under *Daubert*—to the extent that "adequate scientific care" necessitates "an objective and quantitative procedure for identifying the pattern of a sample," and that "[p]atterns must be identified separately and independently in suspect and evidence samples." The quoted language appears in NRC I, *supra* note 1, at 53, and it refers to VNTR profiles. Because the lengths of the VNTRs cannot be determined precisely, statistical criteria must be used if a statement as to whether bands "match" is to be made. Such criteria are discussed below, and they might be all that the committee had in mind when it called for an "objective and quantitative procedure." *Cf.* NRC II, *supra* note 1, at 142 ("the use of visual inspection other than as a screen before objective measurement . . . usually should be avoided"). In any event, courts have not been inclined to treat procedures that allow for subjective judgment in ascertaining the location of VNTR bands as fatal to admissibility. *E.g.*, United States v. Perry, No. CR 91-395-SC (D.N.M. Sept. 7, 1995) (stating that "the autorad is a permanent record, and anyone, including defense experts, can conduct an independent measurement of band size . . . "); State v. Jobe, 486 N.W.2d 407, 420 (Minn. 1992) (observing that "each sample is also examined by a second trained examiner and ultimately the 'match' is confirmed or rejected through computer analysis using wholly objective criteria"); State v. Copeland, 922 P.2d 1304, 1323 (Wash. 1996) (suggesting that "complaints about the analyst's ability to override the computer in placing the cursor at the center of a band . . . would be the type of human error going to weight, not admissibility"); *cf.* NRC II, *supra* ("if for any reason the analyst by visual inspection overrides the conclusion from the measurements, that should be clearly stated and reasons given").

[146] *See* NRC II, *supra* note 1, at 142 ("[D]egraded DNA sometimes migrates farther on a gel than better quality DNA. . . . "). Band-shifting produces a systematic error in measurement. Random error is also present. *See infra* § VII.A.4.

[147] *See* NRC II, *supra* note 1, at 142 ("[A]n experienced analyst can notice whether two bands from a heterozygote are shifted in the same or in the opposite direction from the bands in another lane containing the DNA being compared. If the bands in the two lanes shift a small distance in the same direction, that might indicate a match with band-shifting. If they shift in opposite directions, that is

At the other extreme, the genotypes at a large number of loci can be clearly identical, and the fact of a match not in doubt. In these cases, the DNA evidence is quite incriminating, and the challenge for the legal system lies in explaining just how probative it is. Naturally, as with exclusions, inclusions are most powerful when the samples have been collected, handled, and analyzed properly. But there is one logical difference between exclusions and inclusions. If it is accepted that the samples have different genotypes, then the conclusion that the DNA in them came from different individuals is essentially inescapable. In contrast, even if two samples have the same genotype, there is a chance that the forensic sample came—not from the defendant—but from another individual who has the same genotype. This complication has produced extensive arguments over the statistical procedures for assessing this chance or related quantities. This problem of describing the significance of an unequivocal match is taken up later in this section.

The classification of patterns into the two mutually exclusive categories of exclusions and inclusions is more complicated for VNTRs than for discrete alleles. Determining that DNA fragments from two different samples are the same size is like saying that two people are the same height. The height may well be similar, but is it identical? Even if the same person is measured repeatedly, we expect some variation about the true height due to the limitations of the measuring device. A perfectly reliable device gives the same measurements for all repeated measurements of the same item, but no instrument can measure a quantity like height with both perfect precision and perfect reproducibility. Consequently, measurement variability is a fact of life in ascertaining the sizes of VNTRs.[148]

The method of handling measurement variation that has been adopted by most DNA profilers is statistically inelegant,[149] but it has the virtue of simplic-

---

probably not a match, but a simple match rule or simple computer program might declare it as a match.").

At least one laboratory has reported matches of bands that lie outside its match window but exhibit a band-shifting pattern. It uses monomorphic probes to adjust for the band-shifting. *Compare* Caldwell v. State, 393 S.E.2d 436, 441 (Ga. 1990) (admissible as having reached the "scientific stage of verifiable certainty") *and* State v. Futch, 860 P.2d 264 (Or. Ct. App. 1993) (admissible under a *Daubert*-like standard), *with* Hayes v. State, 660 So. 2d 257 (Fla. 1995) (too controversial to be generally accepted), State v. Quatrevingt, 670 So. 2d 197 (La. 1996) (not shown to be valid under *Daubert*), *and* People v. Keene, 591 N.Y.S.2d 733 (N.Y. Sup. Ct. 1992) (holding that the procedure followed in the case, which did not use the nearest monomorphic probe to make the corrections, was not generally accepted).

[148] In statistics, this variability often is denominated "measurement error." The phrase does not mean that a mistake has been made in performing the measurements, but rather that even measurements that are taken correctly fluctuate about the true value of the quantity being measured.

[149] *See* NRC II, *supra* note 1, at 139 ("[T]he most accurate statistical model for the interpretation of VNTR analysis would be based on a continuous distribution. . . . If models for measurement uncertainty become available that are appropriate for the wide range of laboratories performing DNA analy-

ity.[150] Analysts typically are willing to declare that two fragments match if the bands appear to match visually, and if they fall within a specified distance of one another. For example, the FBI laboratory declares matches within a ±5% match window—if two bands are within ±5% of their average length, then the alleles can be said to match.[151]

Whether the choice of ±5% (or any other figure) as an outer limit for matches is scientifically acceptable depends on how the criterion operates in classifying pairs of samples of DNA.[152] The ±5% window keeps the chance of a false exclusion for a single allele quite small, but at a cost. The easier it is to declare a match between bands at different positions, the easier it is to declare a match between two samples with *different* genotypes. Therefore, deciding whether a match window is reasonable involves an examination of the probability not merely of a false exclusion but also of a false inclusion: "[t]he match window should not be set so small that true matches are missed. At the same time, the window should not be so wide that bands that are clearly different are declared to match."[153]

---

ses and if those analyses are sufficiently robust with respect to departures from the models, we would recommend such methods. Indeed, . . . we expect that any problems in the construction of such models will be overcome, and we encourage research on those models."). Forcing a continuous variable like the positions of the bands on an autoradiogram into discrete categories is not statistically efficient. It results in more matching bands being deemed inconclusive or non-matching than more sophisticated statistical procedures. *See, e.g.*, D.A. Berry et al., *Statistical Inference in Crime Investigations Using Deoxyribonucleic Acid Profiling*, 41 Applied Stat. 499 (1992); I.W. Evett et al., *An Illustration of the Advantages of Efficient Statistical Methods for RFLP Analysis in Forensic Science*, 52 Am. J. Hum. Genetics 498 (1993). Also, it treats matches that just squeak by the match windows as just as impressive as perfect matches.

[150] NRC II, *supra* note 1, at 139.

[151] The FBI arrived at this match window by experiments involving pairs of measurements of the same DNA sequences. It found that this window was wide enough to encompass all the differences seen in the calibration experiments. Other laboratories use smaller percentages for their match windows, but comparisons of the percentage figures can be misleading. *See* D.H. Kaye, Science in Evidence 192 (1997). Because different laboratories can have different standard errors of measurement, profiles from two different laboratories might not be considered inconsistent even though some corresponding bands are outside the match windows of both laboratories. The reason: there is more variability in measurements on different gels than on the same gel, and still more in different gels from different laboratories. *See* Satcher v. Netherland, 944 F. Supp. 1222, 1265 (E.D. Va. 1996).

[152] The use of this window was attacked unsuccessfully in *United States v. Yee*, 134 F.R.D. 161 (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 (D. Vt. 1990), *aff'd*, 955 F.2d 786 (2d Cir. 1992); and *United States v. Perry*, No. CR 91-395-SC (D.N.M. Sept. 7, 1995). For assessments of these arguments, see David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts*, 7 Harv. J.L. & Tech. 101 (1993); D.H. Kaye, *The Relevance of "Matching" DNA: Is the Window Half Open or Half Shut?*, 85 J. Crim. L. & Criminology 676 (1995); William C. Thompson, *Evaluating the Admissibility of New Genetic Tests: Lessons from the "DNA War,"* 84 J. Crim. L. & Criminology 22 (1993); Hans Zeisel & David Kaye, Prove It with Figures: Empirical Methods in Law and Litigation 204–06 (1997).

[153] NRC II, *supra* note 1, at 140. Assuming that the only source of error is the statistical uncertainty in the measurements, this error probability is simply the chance that the two people whose DNA is tested have profiles so similar that they satisfy the matching criterion. With genotypes consisting of four or five VNTR loci, that probability is much smaller than the chance of a false exclusion. *Id.* at 141.

Viewed in this light, the ±5% match window is easily defended—it keeps the probabilities of *both* types of errors very small.[154]

B. Alternative Hypotheses

If the defendant is the source of DNA of sufficient quantity and quality found at a crime scene, then a DNA sample from the defendant and the forensic sample should have the same profile. The inference required in assessing the evidence, however, runs in the opposite direction. The forensic scientist reports that the sample of DNA from the crime scene and a sample from the defendant have the same genotype. To what extent does this tend to prove that the defendant is the source of the forensic sample?[155] Conceivably, other hypotheses could account for the matching profiles. One possibility is laboratory error—the genotypes are not actually the same even though the laboratory thinks that they are. This situation could arise from mistakes in labeling or handling samples or from cross-contamination of the samples.[156] As the 1992 NRC report cautioned, "[e]rrors happen, even in the best laboratories, and even when the analyst is certain that every precaution against error was taken."[157] Another possibility is that the laboratory analysis is correct—the genotypes are truly identical—but the forensic sample came from another individual. In general, the true source might be a close relative of the defendant[158] or an unrelated person who, as luck would have it, just happens to have the same profile as the defendant. The former hypothesis we shall refer to as kinship, and the latter as coincidence. To infer that the defendant is the source of the crime scene DNA, one must reject these alternative hypotheses of laboratory error, kinship, and coincidence. Table 1 summarizes the logical possibilities.

---

[154] NRC II, *supra* note 1, at 140–41; Bernard Devlin & Kathryn Roeder, *DNA Profiling: Statistics and Population Genetics, in* 1 Modern Scientific Evidence, *supra* note 53, § 18-3.1.2, at 717–18.

[155] That the defendant is the source does not necessarily mean that the defendant is guilty of the offense charged. Aside from issues of intent or knowledge that have nothing to do with DNA, there remains, for instance, the possibility that the two samples match because someone framed the defendant by putting a sample of defendant's DNA at the crime scene or in the container of DNA thought to have come from the crime scene. *See generally* United States v. Chischilly, 30 F.3d 1144 (9th Cir. 1994) (dicta on "source probability"); Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 Judicature 222 (1993). For reports of state police planting fingerprint and other evidence to incriminate arrestees, see John Caher, *Judge Orders New Trial in Murder Case*, Times Union (Albany), Jan. 8, 1997, at B2; John O'Brien & Todd Lightly, *Corrupt Troopers Showed No Fear*, The Post-Standard (Syracuse), Feb. 4, 1997, at A3 (an investigation of 62,000 fingerprint cards from 1983–1992 revealed 34 cases of planted evidence among one state police troop).

[156] *See supra* § VI.

[157] NRC I, *supra* note 1, at 89.

[158] A close relative, for these purposes, would be a brother, uncle, nephew, etc. For relationships more distant than second cousins, the probability of a chance match is nearly as small as for persons of the same ethnic subgroup. Devlin & Roeder, *supra* note 154, § 18-3.1.3, at 724. For an instance of the "evil twin" defense, see Hunter v. Harrison, No. 71723, 1997 WL 578917 (Ohio Ct. App. Sept. 18, 1997) (unpublished paternity case).

Table 1. Hypotheses that Might Explain a Match Between Defendant's DNA and DNA at a Crime Scene[159]

IDENTITY:	same genotype, defendant's DNA at crime scene
NON-IDENTITY:	
lab error	different genotypes mistakenly found to be the same
kinship	same genotype, relative's DNA at crime scene
coincidence	same genotype, unrelated individual's DNA

Some scientists have urged that probabilities associated with false positive error, kinship, or coincidence be presented to juries. While it is not clear that this goal is feasible, scientific knowledge and more conventional evidence can help in assessing the plausibility of these alternative hypotheses. If laboratory error, kinship, and coincidence can be eliminated as explanations for a match, then only the hypothesis of identity remains. We turn, then, to the considerations that affect the chances of a reported match when the defendant is not the source of the trace evidence.

1. Error

Although many experts would concede that even with rigorous protocols, the chance of a laboratory error exceeds that of a coincidental match,[160] quantifying the former probability is a formidable task. Some commentary proposes using the proportion of false positives that the particular laboratory has experienced in blind proficiency tests or the rate of false positives on proficiency tests averaged across all laboratories.[161] Indeed, the 1992 NRC Report remarks that "proficiency tests provide a measure of the false-positive and false-negative rates of a laboratory."[162] Yet, the same report recognizes that "errors on proficiency tests do not necessarily reflect permanent probabilities of false-positive or false-negative results,"[163] and the 1996 NRC report suggests that a probability of a false-positive error that would apply to a specific case cannot be estimated objectively.[164] If the false-positive probability were, say, 0.001, it would take tens of thousands of proficiency tests to estimate that probability accurately, and the application of an historical industry-wide error rate to a particular laboratory at a later time would be debatable.[165]

---

[159] *Cf.* N.E. Morton, *The Forensic DNA Endgame*, 37 Jurimetrics J. 477, 480 tbl. 1 (1997).

[160] *E.g.*, Devlin & Roeder, *supra* note 154, § 18-5.3, at 743.

[161] *E.g.*, Jonathan J. Koehler, *Error and Exaggeration in the Presentation of DNA Evidence at Trial*, 34 Jurimetrics J. 21, 37–38 (1993); Scheck, *supra* note 69, at 1984 n.93.

[162] NRC I, *supra* note 1, at 94.

[163] *Id.* at 89.

[164] NRC II, *supra* note 1, at 85–87.

[165] *Id.* at 85–86; Devlin & Roeder, *supra* note 154, § 18-5.3, at 744–45. Such arguments have not persuaded the proponents of estimating the probability of error from industry-wide proficiency testing.

Most commentators who urge the use of proficiency tests to estimate the probability that a laboratory has erred in a particular case agree that blind proficiency testing cannot be done in sufficient numbers to yield an accurate estimate of a small error rate. However, they maintain that proficiency tests, blind or otherwise, should be used to provide a conservative estimate of the false-positive error probability.[166] For example, if there were no errors in 100 tests, a 95% confidence interval would include the possibility that the error rate could be almost as high as 3%.[167]

Instead of pursuing a numerical estimate, the second NAS committee and individual scientists who question the value of proficiency tests for estimating case-specific laboratory-error probabilities suggest that each laboratory document all the steps in its analyses and reserve portions of the DNA samples for independent testing whenever feasible. Scrutinizing the chain of custody, examining the laboratory's protocol, verifying that it adhered to that protocol, and conducting confirmatory tests if there are any suspicious circumstances can help to eliminate the hypothesis of laboratory error,[168] whether or not a case-specific probability can be estimated.[169] Furthermore, if the defendant has had a meaningful opportunity to retest a sample but has been unable or unwilling to obtain an inconsistent result, the relevance of a statistic based on past proficiency tests might be questionable.

2. *Kinship*

With enough genetic markers, all individuals except for identical twins should be distinguishable, but this ideal is not always attainable with the limited number of loci typically used in forensic testing.[170] Close relatives have more genes in common than unrelated individuals, and various procedures have been pro-

---

E.g., Jonathan J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, 37 Jurimetrics J. 425 (1997).

[166] *E.g.*, Koehler, *supra* note 155, at 228; Richard Lempert, *After the DNA Wars: Skirmishing with NRC II*, 37 Jurimetrics J. 439, 447–48, 453 (1997).

[167] *See* NRC II, *supra* note 1, at 86 n.1. For an explanation of confidence intervals, see David H. Kaye & David A. Freedman, Reference Guide on Statistics, § IV.A.2, in this manual.

[168] *E.g.*, Jonathan J. Koehler, *On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates*, 67 U. Colo. L. Rev. 859, 866 (1996) ("In the *Simpson* case, [l]aboratory error was unlikely because many blood samples were tested at different laboratories using two different DNA typing methods."); William C. Thompson, *DNA Evidence in the O.J. Simpson Trial*, 67 U. Colo. L. Rev. 827, 827 (1996) ("the extensive use of duplicate testing in the *Simpson* case greatly reduced concerns (that are crucial in most other cases) about the potential for false positives due to poor scientific practices of DNA laboratories").

[169] *See* Berger, *supra* note 69.

[170] *See, e.g.*, B.S. Weir, *Discussion of "Inference in Forensic Identification,"* 158 J. Royal Stat. Soc'y Ser. A 49, 50 (1995) ("the chance that two unrelated individuals in a population share the same 16-allele [VNTR] profile is vanishingly small, and even for full sibs the chance is only 1 in very many thousands").

posed for dealing with the possibility that the true source of the forensic DNA is not the defendant but a close relative.[171] Often, the investigation, including additional DNA testing, can be extended to all known relatives.[172] But this is not feasible in every case, and there is always the chance that some unknown relatives are included in the suspect population.[173] Formulae are available for computing the probability that any person with a specified degree of kinship to the defendant also possesses the incriminating genotype.[174] For example, the probability that an untested brother (or sister) would match at four loci (with alleles that each occur in 5% of the population) is about 0.006; the probability that an aunt (or uncle) would match is about 0.0000005.[175]

---

[171] *See* Thomas R. Belin et al., *Summarizing DNA Evidence When Relatives are Possible Suspects*, 92 J. Am. Stat. Ass'n 706, 707–08 (1997). Recommendation 4.4 of the 1996 NRC report reads:

> If possible contributors of the evidence sample include relatives of the suspect, DNA profiles of those relatives should be obtained. If these profiles cannot be obtained, the probability of finding the evidence profile in those relatives should be calculated with [specified formulae].

NRC II, *supra* note 1, at 6.

[172] NRC II, *supra* note 1, at 113.

[173] When that population is very large, however, the presence of a few relatives will have little impact on the probability that a suspect drawn at random from that population will have the incriminating genotype. *Id.* Furthermore, it has been suggested that the effect of relatedness is of practical importance only for very close relatives, such as siblings. JFY Brookfield, *The Effect of Relatives on the Likelihood Ratio Associated with DNA Profile Evidence in Criminal Cases*, 34 J. Forensic Sci. Soc'y 193 (1994).

[174] *E.g.*, Brookfield, *supra* note 173; David J. Balding & Peter Donnelly, *Inference in Forensic Identification*, 158 J. Royal Stat. Soc'y Ser. A 21 (1995); Ian W. Evett & Bruce S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists 108–18 (1998); Morton, *supra* note 159, at 484; NRC II, *supra* note 1, at 113. *But see* NRC I, *supra* note 1, at 87 (giving an incorrect formula for siblings). Empirical measures that are not directly interpretable as probabilities also have been described. Belin et al., *supra* note 171.

[175] The large discrepancy between two siblings on the one hand, and an uncle and nephew on the other, reflects the fact that the siblings have far more shared ancestry. All their genes are inherited through the same two parents. In contrast, a nephew and an uncle inherit from two unrelated mothers, and so will have few maternal alleles in common. As for paternal alleles, the nephew inherits not from his uncle, but from his uncle's brother, who shares by descent only about one-half of his alleles with the uncle.

One commentator has proposed that unless the police can eliminate all named relatives as possible culprits, "the defendant should be allowed to name any close relative whom he thinks might have committed the crime," and the state should use the probability "that at least one named relative has DNA like the defendant's" as the sole indication of the plausibility of the hypothesis of kinship. Lempert, *supra* note 166, at 461. For example, if the defendant named two brothers and two uncles as possible suspects, then the probability that at least one shares the genotype would be about (2 x .006) + (2 x .0000005), or about 0.012. Whether such numbers should be introduced even when there is no proof that a close relative might have committed the crime is, of course, a matter to be evaluated under Federal Rules of Evidence 104(b), 401, and 403. *See, e.g.*, Taylor v. Commonwealth, No. 1767-93-1, 1995 WL 80189 (Va. Ct. App. Feb. 28, 1995) (unpublished) ("Defendant argues that this evidence did not consider the existence of an identical twin or close relative to defendant, a circumstance which would diminish the probability that he was the perpetrator. While this hypothesis is conceivable, it has no basis in the record and the Commonwealth must only exclude hypotheses of innocence that reasonably flow from the evidence, not from defendant's imagination.").

3. Coincidence

Another rival hypothesis is coincidence: The defendant is not the source of the crime scene DNA, but happens to have the same genotype as an unrelated individual who is the true source. Various procedures for assessing the plausibility of this hypothesis are available. In principle, one could test all conceivable suspects. If everyone except the defendant has a non-matching profile, then the conclusion that the defendant is the source is inescapable. But exhaustive, error-free testing of the population of conceivable suspects is almost never feasible. The suspect population normally defies any enumeration, and in the typical crime where DNA evidence is found, the population of possible perpetrators is so huge that even if all its members could be listed, they could not all be tested.[176]

An alternative procedure would be to take a sample of people from the suspect population, find the relative frequency of the profile in this sample, and use that statistic to estimate the frequency in the entire suspect population. The smaller the frequency, the less likely it is that the defendant's DNA would match if the defendant were not the source of trace evidence. Again, however, the suspect population is difficult to define, so some surrogate must be used. The procedure commonly followed is to estimate the relative frequency of the incriminating genotype in a large population. But even this cannot be done directly because each possible multilocus profile is so rare that it is not likely to show up in any sample of a reasonable size.[177] However, the frequencies of most alleles can be determined accurately by sampling the population[178] to construct

---

[176] 176. In the United Kingdom and Europe, mass DNA screenings in small towns have been undertaken. *See, e.g.*, Kaye, *supra* note 151, at 222–26.

[177] 177. NRC II, *supra* note 1, at 89–90 ("A very small proportion of the trillions of possible profiles are found in any database, so it is necessary to use the frequencies of individual alleles to estimate the frequency of a given profile."). The 1992 NRC report proposed reporting the occurrences of a profile in a database, but recognized that "such estimates do not take advantage of the full potential of the genetic approach." NRC I, *supra* note 1, at 76. For further discussion of the statistical inferences that might be drawn from the absence of a profile in a sample of a given size, see NRC II, *supra*, at 159–60 (arguing that "the abundant data make [the direct counting method] unnecessary").

[178] 178. Ideally, a probability sample from the population of interest would be taken. Probability sampling is described in David H. Kaye & David A. Freedman, Reference Guide on Statistics, § II.B, and Shari Seidman Diamond, Reference Guide on Survey Research, § III.C, in this manual. Indeed, a few experts have testified that no meaningful conclusions can be drawn in the absence of random sampling. *E.g.*, People v. Soto, 88 Cal. Rptr. 2d 34 (1999); State v. Anderson, 881 P.2d 29, 39 (N.M. 1994).

Unfortunately, a list of the people who comprise the entire population of possible suspects is almost never available; consequently, probability sampling from the directly relevant population is generally impossible. Probability sampling from a proxy population is possible, but it is not the norm in studies of the distributions of genes in populations. Typically, convenience samples are used. The 1996 NRC report suggests that for the purpose of estimating allele frequencies, convenience sampling should give results comparable to random sampling, and it discusses procedures for estimating the random sampling error. NRC II, *supra* note 1, at 126–27, 146–48, 186. For an analysis of case law on the need for random sampling in this area, see D.H. Kaye, Bible *Reading: DNA Evidence in Arizona*, 28 Ariz. St. L.J. 1035 (1996).

databases that reveal how often each allele occurs.[179] Principles of population genetics then can be applied to combine the estimated allele frequencies into an estimate of the probability that a person born in the population will have the multilocus genotype. This probability often is referred to as the random match probability. Three principal methods for computing the random match probability from allele frequencies have been developed. This section describes these methods; the next section considers other quantities that have been proposed as measures of the probative value of the DNA evidence.

a. The Basic Product Rule

The basic product rule estimates the frequency of genotypes in an infinite population of individuals who choose their mates and reproduce independently of the alleles used to compare the samples. Although population geneticists describe this situation as random mating, these words are terms of art. Geneticists know that people do not choose their mates by a lottery, and they use "random mating" to indicate that the choices are uncorrelated with the specific alleles that make up the genotypes in question.[180]

In a randomly mating population, the expected frequency of a pair of alleles at each locus depends on whether the two alleles are distinct. If a different allele is inherited from each parent, the expected single-locus genotype frequency is twice the product of the two individual allele frequencies.[181] But if the offspring happens to inherit the same allele from each parent, the expected single-locus genotype frequency is the square of the allele frequency.[182] These proportions

---

[179] In the formative years of forensic DNA testing, defendants frequently contended that the size of the forensic databases were too small to give accurate estimates, but this argument generally proved unpersuasive. *E.g.*, United States v. Shea, 937 F. Supp. 331 (D.N.H. 1997); People v. Soto, 88 Cal. Rptr. 2d 34 (1999); State v. Dishon, 687 A.2d 1074, 1090 (N.J. Super. Ct. App. 1997); State v. Copeland, 922 P.2d 1304, 1321 (Wash. 1996).

To the extent that the databases are comparable to random samples, confidence intervals are a standard method for indicating the amount of error due to sample size. *E.g.*, Kaye, *supra* note 152. Unfortunately, the meaning of a confidence interval is subtle, and the estimate commonly is misconstrued. *See* David H. Kaye & David A. Freedman, Reference Guide on Statistics, § IV.A.2, in this manual.

[180] *E.g.*, NRC II, *supra* note 1, at 90:

> In the simplest population structure, mates are chosen at random. Clearly, the population of the United States does not mate at random; a person from Oregon is more likely to mate with another from Oregon than with one from Florida. Furthermore, people often choose mates according to physical and behavioral attributes, such as height and personality. But they do not choose each other according to the markers used for forensic studies, such as VNTRs and STRs. Rather, the proportion of matings between people with two marker genotypes is determined by their frequencies in the mating population. If the allele frequencies in Oregon and Florida are the same as those in the nation as a whole, then the proportion of genotypes in the two states will be the same as those for the United States, even though the population of the whole country clearly does not mate at random.

[181] In more technical terms, when the frequencies of two alleles are p_1 and p_2, the single-locus genotype frequency for the corresponding heterozygotes is expected to be $2p_1p_2$.

[182] The expected proportion is p_1^2 for allele 1, and p_2^2 for allele 2. With VNTRs, a complication arises with apparent homozygotes. A single band on an autoradiogram might really be two bands that

are known as Hardy-Weinberg proportions. Even if two populations with distinct allele frequencies are thrown together, within the limits of chance variation, random mating produces Hardy-Weinberg equilibrium in a single generation. An example is given in this footnote.[183]

Once the proportion of the population that has each of the single-locus genotypes for the forensic profile has been estimated in this way, the proportion of the population that is expected to share the combination of them—the multilocus profile frequency—is given by multiplying the single-locus proportions. This multiplication is exactly correct when the single-locus genotypes are statistically independent. In that case, the population is said to be in linkage equilibrium.

Extensive litigation and scientific commentary have considered whether the occurrences of alleles at each locus are independent events (Hardy-Weinberg equilibrium), and whether the loci are independent (linkage equilibrium). Beginning around 1990, several scientists suggested that the equilibrium frequencies do not follow the simple model of a homogeneous population mating without regard to the loci used in forensic DNA profiling. They suggested that the major racial populations are composed of ethnic subpopulations whose members tend to mate among themselves.[184] Within each ethnic subpopulation, mating still can be random, but if, say, Italian-Americans have allele frequencies that are markedly different than the average for all whites, and if Italian-Americans only mate among themselves, then using the average frequencies for all whites in the basic product formula could understate—or overstate—a multilocus profile frequency for the subpopulation of Italian-Americans.[185] Similarly, using the popu-

---

are close together, or a second band that is relatively small might have migrated to the edge of the gel during the electrophoresis. Forensic laboratories therefore make a "conservative" assumption. They act as if there is a second, unseen band, and they use the excessively large value of p_2 = 100% for the frequency of the presumably unseen allele. With this modification, the genotype frequency for apparent homozygotes becomes $P = 2p_1$. If the single-banded pattern is a true homozygote, this 2p convention overstates the frequency of the single-locus genotype because 2p is greater than p^2 for any possible proportion p. For instance, if $p = 0.05$, then $2p = 0.10$, which is 40 times greater than $p^2 = 0.0025$.

[183] Suppose that 10% of the sperm in the gene pool of the population carry allele 1 (A_1), and 50% carry allele 2 (A_2). Similarly, 10% of the eggs carry A_1, and 50% carry A_2. (Other sperm and eggs carry other types.) With random mating, we expect 10% x 10% = 1% of all the fertilized eggs to be A_1A_1, and another 50% x 50% = 25% to be A_2A_2. These constitute two distinct homozygote profiles. Likewise, we expect 10% x 50% = 5% of the fertilized eggs to be A_1A_2 and another 50% x 10% = 5% to be A_2A_1. These two configurations produce indistinguishable profiles—a band, dot, or the like for A_1 and another mark for A_2. So the expected proportion of heterozygotes A_1A_2 is 5% + 5% = 10%.

Oddly, some courts and commentators have written that the expected heterozygote frequency for this example is only 5%. *E.g.*, William C. Thompson & Simon Ford, *DNA Typing: Acceptance and Weight of the New Genetic Identification Tests*, 75 Va. L. Rev. 45, 81–82 (1989). For further discussion, see Kaye, *supra* note 178; David H. Kaye, *Cross-Examining Science*, 36 Jurimetrics J. vii (Winter 1996).

[184] The most prominent expression of this position is Richard C. Lewontin & Daniel L. Hartl, *Population Genetics in Forensic DNA Typing*, 254 Science 1745 (1991).

[185] On average, the use of population-wide allele frequencies overstates the genotype frequencies within defendant's subpopulation. *See* Dan E. Krane et al., *Genetic Differences at Four DNA Typing Loci*

lation frequencies could understate—or overstate—the profile frequencies in the white population itself.[186]

Consequently, if we want to know the frequency of an incriminating profile among Italian-Americans, the basic product rule applied to the white allele frequencies could be in error; and there is some chance that it will understate the profile frequency in the white population as a whole. One might presume that the extent of the error could be determined by looking to the variations across racial groups,[187] but, for a short time, a few scientists insisted that variations from one ethnic group to another within a race were larger than variations from one race to another.[188] In light of this literature[189] courts had grounds to conclude that the basic product rule, used with broad population frequencies, was not universally accepted for estimating profile frequencies within subpopulations. Yet, few courts recognized that there was much less explicit dissension over the ability of the rule to estimate profile frequencies in a general population.[190] Particularly in *Frye* jurisdictions, a substantial number of appellate courts began to exclude DNA evidence for want of a generally accepted method of estimating profile frequencies in both situations.[191]

---

in Finnish, Italian, and Mixed Caucasian Populations, 89 Proc. Nat'l Acad. Sci. 10583 (1992); Stanley Sawyer et al., *DNA Fingerprinting Loci Do Show Population Differences: Comments on Budowle et al.*, 59 Am. J. Hum. Genetics 272 (1996) (letter). This mean overestimation occurs because (1) the use of population-wide frequencies rather than subpopulation frequencies underestimates homozygote frequencies and overestimates heterozygote frequencies, and (2) heterozygosity far exceeds homozygosity.

[186] The use of the population-wide allele frequencies usually overstates genotype frequencies in the population as a whole, thereby benefitting most defendants. *See* Kaye, *supra* note 152, at 142.

[187] On the problems in defining racial populations, compare C. Loring Brace, *Region Does Not Mean "Race"—Reality Versus Convention in Forensic Anthropology*, 40 J. Forensic Sci. 171 (1995), with Kenneth A.R. Kennedy, *But Professor, Why Teach Race Identification if Races Don't Exist?*, 40 J. Forensic Sci. 797 (1995).

[188] *Compare* Lewontin & Hartl, *supra* note 184, at 1745 ("there is, on average, one-third more genetic variation among Irish, Spanish, Italians, Slavs, Swedes, and other subpopulations than there is, on average, between Europeans, Asians, Africans, Amerindians, and Oceanians"), *with* Richard C. Lewontin, *Discussion*, 9 Stat. Sci. 259, 260 (1994) ("all parties agree that differentiation among [major ethnic groups] is as large, if not larger than, the difference among tribes and national groups [within major ethnic groups]"). Other population geneticists dismissed as obviously untenable the early assertions of greater variability across the ethnic subpopulations of a race than across races. *E.g.*, B. Devlin et al., *NRC Report on DNA Typing*, 260 Science 1057 (1993); N.E. Morton et al., *Kinship Bioassay on Hypervariable Loci in Blacks and Caucasians*, 90 Proc. Nat'l Acad. Sci. USA 1892, 1896 (1993) (Gene frequencies cited by Lewontin & Hartl are atypical, and "[l]ess than 2% of the diversity selected by Lewontin and Hartl is due to the national kinship to which they attribute it, little of which persists in regional forensic samples.").

[189] The literature on genetic differences across the globe is reviewed in, *e.g.*, Devlin & Roeder, *supra* note 154, § 18–3.2.1, at 725–28 (suggesting that this body of research indicates that the extent of the variation across subpopulations is relatively small).

[190] *See* Kaye, *supra* note 152, at 146. The general perception was that ethnic stratification within the major racial categories posed a problem regardless of whether the relevant population for estimating the random match probability was a broad racial group or a narrow, inbred ethnic subpopulation.

[191] *See* cases cited, Kaye, *supra* note 152. Courts applying *Daubert* or similar standards were more

b. The Product Rule with Ceilings

In 1992, the National Academy of Sciences' Committee on DNA Technology in Forensic Science assumed arguendo that population structure was a serious threat to the basic product rule and proposed a variation to provide an upper bound on a profile frequency within any population or subpopulation.[192] The interim ceiling method uses the same general formulas as the basic product rule,[193] but with different values of the frequencies. Instead of multiplying together the allele frequencies from any single, major racial database, the procedure picks, for each allele in the DNA profile, the largest value seen in *any* race.[194] If that value is less than 10%, the procedure inflates it to 10%. Those values are then multiplied as with the basic product rule. Thus, the ceiling method employs a mix-and-match, inflate, and multiply strategy. The result, it is widely believed, is an extremely conservative estimate of the profile frequency that more than compensates for the possibility of any population structure that might undermine the assumptions of Hardy-Weinberg and linkage equilibria in the major racial populations.[195]

---

receptive to the evidence. *E.g.*, United States v. Jakobetz, 955 F.2d 786 (2d Cir. 1992), *aff'g*, 747 F. Supp. 250 (D. Vt. 1990); United States v. Bonds, 12 F.3d 540 (6th Cir. 1993), *aff'g*, United States v. Yee, 134 F.R.D. 161 (N.D. Ohio 1991); United States v. Chischilly, 30 F.3d 1144 (9th Cir. 1994); United States v. Davis, 40 F.3d 1069 (10th Cir. 1994).

[192] *See* NRC I, *supra* note 1, at 91–92; *id.* at 80 ("Although mindful of the controversy, the committee has chosen to assume for the sake of discussion that population substructure may exist and provide a method for estimating population [genotype] frequencies in a manner that adequately accounts for it."). The report was unclear as to whether its "interim ceiling principle" was a substitute for or merely a supplement to the usual basic product rule. Years later, one member of the committee opined that the committee intended the latter interpretation. Eric S. Lander & Bruce Budowle, *Commentary: DNA Fingerprinting Dispute Laid to Rest*, 371 Nature 735 (1994). In any event, the interim ceiling principle was proposed as a stopgap measure, to be supplanted by another ceiling principle that could be used after sampling many "[g]enetically homogeneous populations from various regions of the world." NRC I, *supra*, at 84.

[193] Applied to a single racial group like whites, the basic product rule estimates the frequency of the multilocus genotype as the product of the single-locus frequencies, and it estimates each single-locus frequency as $2p_1p_2$ for heterozygotes or as a quantity exceeding p^2 for homozygotes, where p refers to frequencies estimated from the database for that race.

[194] Actually, an even larger figure is used—the upper 95% confidence limit on the allele frequency estimate for that race. This is intended to account for sampling error due to the limited size of the databases. NRC I, *supra* note 1, at 92.

[195] *See, e.g.*, NRC II, *supra* note 1, at 156 ("sufficiently conservative to accommodate the presence of substructure . . . a lower limit on the size of the profile frequency"); NRC I, *supra* note 1, at 91 ("conservative calculation"). This modification of the basic product rule provoked vociferous criticism from many scientists, and it distressed certain prosecutors and other law enforcement personnel who perceived the 1992 NRC report as contributing to the rejection of DNA evidence in many jurisdictions. *See, e.g.*, Kaye, *supra* note 2, at 396. The judicial impact of the NRC report and the debate among scientists over the ceiling method are reviewed in 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 (1994) (suggesting that because the disagreement about the ceiling principle is a dispute about legal

c. The Product Rule for a Structured Population

The 1996 NRC Report distinguishes between cases in which the suspect population is a broad racial population and those in which that population is a genetically distinct subgroup. In the former situation, Recommendation 4.1 endorses the basic product rule:

> In general, the calculation of a profile frequency should be made with the product rule. If the race of the person who left the evidence-sample DNA is known, the database for the person's race should be used; if the race is not known, calculations for all the racial groups to which possible suspects belong should be made.[196]

"For example," the committee wrote, "if DNA is recovered from semen in a case in which a woman hitchhiker on an interstate highway has been raped by a white man, the product rule with the 2*p* rule can be used with VNTR data from a sample of whites to estimate the frequency of the profile among white males. If the race of the rapist were in doubt, the product rule could still be used and the results given for data on whites, blacks, Hispanics, and east Asians."[197] However, "[w]hen there are partially isolated subgroups in a population, the situation is more complex; then a suitably altered model leads to slightly different estimates of the quantities that are multiplied together in the formula for the frequency of the profile in the population."[198] Thus, the committee's Recommendation 4.2 urges that:

> If the particular subpopulation from which the evidence sample came is known, the allele frequencies for the specific subgroup should be used as described in Recommendation 4.1.

---

policy rather than scientific knowledge, the debate among scientists does not justify excluding ceiling frequencies).

By 1995, however, many courts were concluding that because a consensus that ceiling estimates are conservative had emerged, these estimates are admissible. At the same time, other courts that only a short while ago had held basic product estimates to be too controversial to be admissible decided that there was sufficient agreement about the basic product rule for it to be used. *See* State v. Johnson, 922 P.2d 294, 300 (Ariz. 1996); State v. Copeland, 922 P.2d 1304, 1318 (Wash. 1996) ("Although at one time a significant dispute existed among qualified scientists, from the present vantage point we are able to say that the significant dispute was short-lived."); Kaye, *supra* note 4.

In 1994, a second NAS committee was installed to review the criticism and the studies that had accumulated in the aftermath of the 1992 report. In 1996, it reported that the ceiling method is an unnecessary and extravagant way to handle the likely extent of population structure. NRC II, *supra* note 1, at 158, 162.

[196] NRC II, *supra* note 1, at 5. The recommendation also calls for modifications to the Hardy-Weinberg proportion for apparent homozygotes. The modifications depend on whether the alleles are discrete (as in PCR-based tests) or continuous (as in VNTR testing). *Id.* at 5 n.2.

[197] *Id.* at 5 (note omitted). *See also* C. Thomas Caskey, *Comments on DNA-based Forensic Analysis*, 49 Am. J. Hum. Genetics 893 (1991) (letter). For a case with comparable facts, see *United States v. Jakobetz*, 747 F. Supp. 250 (D. Vt. 1990), *aff'd*, 955 F.2d 786 (2d Cir. 1992).

[198] NRC II, *supra* note 1, at 5.

[199] *Id.* at 5–6.

> If allele frequencies for the subgroup are not available, although data for the full population are, then the calculations should use the population-structure equations 4.10 for each locus, and the resulting values should be multiplied.[199]

The "suitably altered model" is a generalization of the basic product rule. In this affinal model, as it is sometimes called,[200] the "population-structure equations" are similar to those for multiplying single-locus frequencies. However, they involve not only the individual allele frequencies, but also a quantity that measures the extent of population structure.[201] The single-locus frequencies are multiplied together as in the basic product rule to find the multilocus frequency. Although few reported cases have analyzed the admissibility of random match probabilities estimated with the product rule for structured populations, the validity of the affinal model of a structured population has not been questioned in the scientific literature.[202]

The committee recommended that the population-structure equations be used in special situations,[203] but they could be applied to virtually all cases. The report suggests conservative values of the population-structure constant might be used for broad suspect populations as well as values for many partially isolated subpopulations.[204] The population-structure equations always give more conservative probabilities than the basic product rule when both formulae are applied to the same database, and they are usually conservative relative to calculations based on the subpopulation of the defendant.[205]

---

[200] Devlin & Roeder, *supra* note 154, § 18–3.1.3, at 723.

[201] NRC II, *supra* note 1, at 114–15 (equations 4.10a & 4.10b). *See also* papers cited, Devlin & Roeder, *supra* note 154, § 18-3.1.3, at 723 n.37. This quantity usually is designated θ. *See generally* Evett & Weir, *supra* note 174, at 94–107, 118–23, 156–62.

[202] The district court in *United States v. Shea*, 957 F. Supp. 331, 343 (D.N.H. 1997), held that a random match probability using an F_{ST} adjustment satisfies *Daubert. See also* United States v. Gaines, 979 F. Supp. 1429 (S.D. Fla. 1997).

[203] The report explains that the recommendation to use the population-structure equations "deals with the case in which the person who is the source of the evidence DNA is known to belong to a particular subgroup of a racial category." NRC II, *supra* note 1, at 6. It offers this illustration:

> For example, if the hitchhiker was not on an interstate highway but in the midst of, say, a small village in New England and we had good reason to believe that the rapist was an inhabitant of the village, the product rule could still be used (as described in Recommendation 4.1) if there is a reasonably large database on the villagers.
>
> If specific data on the villagers are lacking, a more complex model could be used to estimate the random-match probability for the incriminating profile on the basis of data on the major population group (whites) that includes the villagers.

Id. For further discussion of when Recommendation 4.1 applies, see *infra* note 208.

[204] *Id.* at 115, 116 ("typical values for white and black populations are less than 0.01, usually about 0.002. Values for Hispanics are slightly higher . . . .") ("For urban populations, 0.01 is a conservative value. A higher value—say 0.03—could be used for isolated villages."); *cf.* Devlin & Roeder, *supra* note 154, § 18–3.1.3, at 723–24 ("For [VNTR] markers, θ is generally agreed to lie between 0 and .02 for most populations.").

[205] Devlin & Roeder, *supra* note 154, § 18–3.1.3, at 723.

In a few situations, however, very little data on either the larger population or the specific subpopulation will be available.[206] To handle such cases, Recommendation 4.3 provides:

> If the person who contributed the evidence sample is from a group or tribe for which no adequate database exists, data from several other groups or tribes thought to be closely related to it should be used. The profile frequency should be calculated as described in Recommendation 4.1 for each group or tribe.[207]

Similar procedures have been followed in a few cases where the issue has surfaced.[208]

---

[206] *See, e.g.*, People v. Atoigue, DCA No. CR 91-95A, 1992 WL 245628 (D. Guam App. Div. 1992), *aff'd without deciding whether admission of DNA evidence was error*, No. 92-10589, 1994 WL 477518 (9th Cir. 1994) (unpublished).

[207] NRC II, *supra* note 1, at 6. The committee explained that:

> This recommendation deals with the case in which the person who is the source of the evidence DNA is known to belong to a particular subgroup of a racial category but there are no DNA data on either the subgroup or the population to which the subgroup belongs. It would apply, for example, if a person on an isolated Indian reservation in the Southwest, had been assaulted by a member of the tribe, and there were no data on DNA profiles of the tribe. In that case, the recommendation calls for use of the product rule (as described in Recommendation 4.1) with several other closely related tribes for which adequate databases exist.

Id.

[208] A variation on this procedure was used in *United States v. Chischilly*, 30 F.3d 1144, 1158 n.29 (9th Cir. 1994), to handle the concern that the FBI had insufficient data on VNTR allele frequencies among Navajos. In *Government of the Virgin Islands v. Byers*, 941 F. Supp. 513 (D.V.I. 1996), two black men in St. Thomas engaged in "a four-month crime spree" of rape, robbery, kidnapping, and burglary. *Id.* at 514. After one woman was raped a second time by the pair, she identified one as Byers. Byers pled guilty to various charges and testified against an acquaintance, whom the FBI linked to three victims by a three-locus VNTR profile. *Id.* Random match probabilities for African-Americans, whites, and Hispanics were estimated from the FBI's databases, which did not include inhabitants of St. Thomas. The defendant argued that because the African-American database did not include Afro-Caribbeans, the probabilities were inadmissible. *Id.* at 515. The district court reasoned that:

> [A]s the 1996 NRC Report concluded, population subgrouping is important only if we know that the suspect is a member of a particular subgroup. All that was known about the suspect in this case was his race. The victims did not indicate whether he was a transplanted North American, a native St. Thomian, or an immigrant from one of the other Caribbean islands. As recommended by the 1996 NRC Report, the FBI's database for Blacks was used in comparing the defendant's DNA profile since the suspect's race is known in this case. Because investigators did not know the subgroup to which the suspect belonged, there was no need to compare the defendant's DNA profile with any subgroup. The FBI procedure of giving DNA frequency estimations for several different racial groups was more than adequate under the circumstances.

Id. at 522. In our view, the court's reliance on Recommendation 4.1 of the 1996 report was misplaced. Although the victims could not know with certainty whether their assailants were African-American or Afro-Caribbean, the locale of the crimes indicates that the suspect population was dominated by the latter, and that group is not a subpopulation of the African-American population for which a database is available. Consequently, Recommendation 4.3 would seem to apply. Nevertheless, by crediting FBI testimony that the distribution of VNTR alleles in African-Americans is similar to that in Afro-Caribbeans, the court followed the substance of Recommendation 4.3. *Id.*; *see also* Government of Virgin Islands v. Penn, 838 F. Supp. 1054, 1071 (D.V.I. 1993) ("any concern that the St. Thomas black population's bin frequencies are drastically different from those of the United States' black population is unwarranted").

d. Adjusting for a Database Search

Whatever variant of the product rule might be used to find the probability of the genotype in a population, subpopulation, or relative, the number is useful only insofar as it establishes (1) that the DNA profile is sufficiently discriminating to be probative, and (2) that the same DNA profile in the defendant and the crime-scene stain is unlikely to occur if the DNA came from someone other than the defendant. Yet, unlikely events happen all the time. An individual wins the lottery even though it was very unlikely that the particular ticket would be a winner. The chance of a particular supertanker running aground and producing a massive spill on a single trip may be very small, but the Exxon Valdez did just that.

The apparent paradox of supposedly low-probability events being ubiquitous results from what statisticians call a "selection effect" or "data mining." If we pick a lottery ticket at random, the probability p that we have the winning ticket is negligible. But if we search through all the tickets, sooner or later we will find the winning one. And even if we search through some smaller number N of tickets, the probability of picking a winning ticket is no longer p, but Np.[209]

Likewise, there may be a small probability p that a randomly selected individual who is not the source of the forensic sample has the incriminating genotype. That is somewhat like having a winning lottery ticket.[210] If N people are included in the search for a person with the matching DNA, then the probability of a match in this group is not p, but some quantity that could be as large as Np.[211] This type of reasoning led the second NRC committee to recommend that "[w]hen the suspect is found by a search of DNA databases, the random-match probability should be multiplied by N, the number of persons in the database."[212]

The first NAS committee also felt that "[t]he distinction between finding a match between an evidence sample and a suspect sample and finding a match between an evidence sample and one of many entries in a DNA profile databank

---

[209] If there are T tickets and one winning ticket, then the probability that a randomly selected ticket is the winner is $p = 1/T$, and the probability that a set of N randomly selected tickets includes the winner is $N/T = Np$, where $1 \leq N \leq T$.

[210] The analysis of the DNA database search is more complicated than the lottery example suggests. In the simple lottery, there was exactly one winner. In the database case, we do not know how many "winners" there are, or even if there are any. The situation is more like flipping a coin N times, where the coin has a probability p of heads on each independent toss.

[211] *See* NRC II, *supra* note 1, at 163–65. Assuming that the individual who left the trace evidence sample is not in a database of unrelated people, the probability of at least one match is $1 - (1-p)N$, which is equal to or less than Np.

[212] NRC II, *supra* note 1, at 161 (Recommendation 5.1). The DNA databases that are searched usually consist of profiles of offenders convicted of specified crimes. *See, e.g.*, Boling v. Romer, 101 F.3d 1336 (10th Cir. 1996); Rise v. Oregon, 59 F.3d 1556 (9th Cir. 1995); Jones v. Murray, 962 F.2d 302 (4th Cir. 1992); Landry v. Attorney General, 709 N.E.2d 1085 (Mass. 1999) (all rejecting constitutional challenges to compelling offenders to provide DNA samples for databases).

is important."[213] Rather than proposing a statistical adjustment to the match probability, however, that committee recommended using only a few loci in the databank search, then confirming the match with additional loci, and presenting only "the statistical frequency associated with the additional loci . . . ."[214]

A number of statisticians reject the committees' view that the random match probability should be inflated, either by a factor of N or by ignoring the loci used in the database search.[215] They argue that, if anything, the DNA evidence against the defendant is slightly stronger when not only has the defendant been shown to possess the incriminating profile, but also a large number of other individuals have been eliminated as possible sources of the crime scene DNA.[216] They conclude that no adjustment is required.

At its core, the statistical debate turns on how the problem is framed and what type of statistical reasoning is accepted as appropriate. The NAS committees ask how surprising it would be to find a match in a large database if the database does not contain the true source of the trace evidence. The more surprising the result, the more it appears that the database does contain the source. Because it would be more surprising to find a match in a test of a single innocent suspect than it would be to find a match by testing a large number of innocent suspects, the NAS committees conclude that the single-test match is more convincing evidence than the database search match.

The critics do not deny the mathematical truism that examining more innocent individuals increases the chance of finding a match, but they maintain that the committees have asked the wrong question. They emphasize that the question of interest to the legal system is not whether the database contains the culprit, but whether the one individual whose DNA matches the trace evidence DNA is the source of that trace; and they note that as the size of a database approaches that of the entire population, finding one and only one matching individual should be more, not less, convincing evidence against that person.[217] Thus, instead of looking at how surprising it would be to find a match in a group of innocent suspects, the "no-adjustment" school asks how much the result of the database search enhances the probability that the individual so identified is the source. They reason that the many exclusions in a database search reduce the number of people who might have left the trace evidence if

---

[213] It used the same N*p* formula in a numerical example to show that "[t]he chance of finding a match in the second case is considerably higher, because one . . . fishes through the databank, trying out many hypotheses." NRC I, *supra* note 1, at 124.

[214] *Id.* The second NAS Committee did not object to this procedure. It proposed the N*p* adjustment as an alternative that might be useful when there were very few typable loci in the trace evidence sample.

[215] *E.g.*, Peter Donnelly & Richard D. Friedman, *DNA Database Searches and the Legal Consumption of Scientific Evidence*, 97 Mich. L. Rev. 931 (1999); authorities cited, *id.* at 933 n.13.

[216] *Id.* at 933, 945, 948, 955, 957; Evett & Weir, *supra* note 174, at 219–22.

[217] *See, e.g.*, Donnelly & Friedman, *supra* note 215, at 952–53.

the suspect did not. This additional information, they conclude, increases the likelihood that the defendant is the source, although the effect is indirect and generally small.[218]

C. Measures of Probative Value

Sufficiently small probabilities of a match for close relatives and unrelated members of the suspect population undermine the hypotheses of kinship and coincidence. Adequate safeguards and checks for possible laboratory error make that explanation of the finding of matching genotypes implausible. The inference that the defendant is the source of the crime scene DNA is then secure. But this mode of reasoning by elimination is not the only way to analyze DNA evidence. This section discusses two alternatives that some statisticians prefer—likelihoods and posterior probabilities. In the next section, we review all the statistics that relate to rival hypotheses and probative value and consider the legal doctrine that must be considered in deciding the admissibility of the various types of presentations.

1. Likelihood Ratios

To choose between two competing hypotheses, one can compare how probable the evidence is under each hypothesis. Suppose that the probability of a match in a well-run laboratory is close to 1 when the samples both contain only the defendant's DNA, while the probability of a coincidental match and the probability of a match with a close relative are close to 0. In these circumstances, the DNA profiling result strongly supports the claim that the defendant is the source, for the observed outcome—the match—is many times more probable when the defendant is the source than when someone else is. How many times more probable? Suppose that there is a 1% chance that the laboratory would miss a true match, so that the probability of its finding a match when the defendant is the source is 0.99. Suppose further that $p = 0.00001$ is the random match probability. Then the match is 0.99/0.00001, or 99,000 times more likely to be seen if the defendant is the source than if an unrelated individual is. Such a ratio is called a likelihood ratio, and a likelihood ratio of 99,000 means that the DNA profiling supports the claim of identity 99,000 times more strongly than it supports the hypothesis of coincidence.[219]

Likelihood ratios are particularly useful for VNTRs and for trace evidence samples that contain DNA from more than one person.[220] With VNTRs, the

---

[218] *Id.* at 245.

[219] *See* NRC II, *supra* note 1, at 100; Kaye, *supra* note 152.

[220] *See supra* § V. Mixed samples arise in various ways—blood from two or more persons mingled at the scene of a crime, victim and assailant samples on a vaginal swab, semen from multiple sexual assailants, and so on. In many cases, one of the contributors—for example, the victim—is known, and

procedure commonly used to estimate the allele frequencies that are combined via some version of the product rule is called binning.[221] In the simplest and most accurate version, the laboratory first forms a "bin" that stretches across the range of fragment lengths in the match window surrounding an evidence band. For example, if a 1,000 base-pair (bp) band is seen in the evidence sample, and the laboratory's match window is ±5%, then the bin extends from 950 to 1,050 bp. The laboratory then finds the proportion of VNTR bands in its database that fall within this bin. If 7% of the bands in the database lie in the 950–1,050 bp range, then 7% is the estimated allele frequency for this band. The two-stage procedure of (1) declaring matches between two samples when all the corresponding bands lie within the match window and (2) estimating the frequency of a band in the population by the proportion that lie within the corresponding bin is known as match-binning.[222]

As noted in section VII.A, match-binning is statistically inefficient. It ignores the extent to which two samples match and gives the same coincidence probability to a close match as it does to a marginal one. Other methods obviate the need for matching by simultaneously combining the probability of the observed degree of matching with the probability of observing bands that are that close together. These "similarity likelihood ratios" dispense with the somewhat arbitrary dichotomy between matches and nonmatches.[223] They have been advocated on the ground that they make better use of the DNA data,[224] but they

---

the genetic profile of the unknown portion is readily deduced. In those situations, the analysis of a remaining single-person profile can proceed in the ordinary fashion. "However, when the contributors to a mixture are not known or cannot otherwise be distinguished, a likelihood-ratio approach offers a clear advantage and is particularly suitable." NRC II, *supra* note 1, at 129. *Contra* R.C. Lewontin, *Population Genetic Issues in the Forensic Use of DNA, in* 1 Modern Scientific Evidence, The Law and Science of Expert Testimony, *supra* note 53, § 17–5.0, at 703–05; Thompson, *supra* note 168, at 855–56. For an exposition of this likelihood ratio approach, see Evett & Weir, *supra* note 174, at 188–205.

[221] There are two types of binning in use. *Floating bins* are conceptually simpler and more appropriate than *fixed bins*, but the latter can be justified as an approximation to the former. For the details of binning and suggestions for handling some of the complications that have caused disagreements over certain aspects of fixed bins, see NRC II, *supra* note 1, at 142–45.

[222] Likelihood ratios for match-binning results are identical to those for discrete allele systems. If the bin frequencies reveal that a proportion p of the population has DNA whose bands each fall within the match window of the corresponding evidence bands, then the match-binning likelihood ratio is $1/p$.

[223] The methods produce likelihood ratios tailored to the observed degree of matching. Two more or less "matching" bands would receive less weight when the measured band lengths differ substantially, and more weight when the lengths differ very little. Devlin & Roeder, *supra* note 154, § 18–3.1.4, at 724. And, bands that occur in a region where relatively few people have VNTRs contribute more to the likelihood ratio than if they occur in a zone where VNTRs are common.

[224] *See* NRC II, *supra* note 1, at 161 ("VNTR data are essentially continuous, and, in principle, a continuous model should be used to analyze them."); authorities cited, *id.* at 200; A. Collins & N.E. Morton, *Likelihood Ratios for DNA Identification*, 91 Proc. Nat'l Acad. Sci. USA 6007 (1994); Devlin & Roeder, *supra* note 154, § 18–3.1.4, at 724.

have been attacked, primarily on the ground that they are complicated and difficult for nonstatisticians to understand.[225]

2. Posterior Probabilities

The likelihood ratio expresses the relative strength of an hypothesis, but the judge or jury ultimately must assess a different type of quantity—the probability of the hypothesis itself. An elementary rule of probability theory known as Bayes' theorem yields this probability. The theorem states that the odds in light of the data (here, the observed profiles) are the odds as they were known prior to receiving the data times the likelihood ratio: *posterior odds = likelihood ratio x prior odds*.[226] For example, if the relevant match probability[227] were 1/100,000, and if the chance that the laboratory would report a match between samples from the same source were 0.99, then the likelihood ratio would be 99,000, and the jury could be told how the DNA evidence raises various prior probabilities that the defendant's DNA is in the evidence sample.[228] It would be appropriate to explain that these calculations rest on many premises, including the premise that the genotypes have been correctly determined.[229]

One difficulty with this use of Bayes' theorem is that the computations consider only one alternative to the claim of identity at a time. As indicated in § VII(B), however, several rival hypotheses might apply in a given case. If it is not defendant's DNA in the forensic sample, is it from his father, his brother, his uncle, et cetera? Is the true source a member of the same subpopulation? A member of a different subpopulation in the same general population? In principle the likelihood ratio can be generalized to a likelihood function that takes on suitable values for every person in the world, and the prior probability for each person can be cranked into a general version of Bayes' rule to yield the posterior probability that the defendant is the source. In this vein, a few commentators suggest that Bayes' rule be used to combine the various likelihood

---

[225] *E.g.*, Lewontin, *supra* note 220, § 17–5.0, at 705.

[226] Odds and probabilities are two ways to express chances quantitatively. If the probability of an event is P, the odds are P/(1 – P). If the odds are O, the probability is O/(O + 1). For instance, if the probability of rain is 2/3, the odds of rain are 2 to 1 because (2/3) / (1 – 2/3) = (2/3) / (1/3) = 2. If the odds of rain are 2 to 1, then the probability is 2/(2 + 1) = 2/3.

[227] By "relevant match probability," we mean the probability of a match given a specified type of kinship or the probability of a random match in the relevant suspect population. For relatives more distantly related than second cousins, the probability of a chance match is nearly as small as for persons of the same subpopulation. Devlin & Roeder, *supra* note 154, § 18–3.1.3, at 724.

[228] For further discussion of how Bayes' rule might be used in court with DNA evidence, see, *e.g.*, Kaye, *supra* note 152; NRC II, *supra* note 1, at 201–03.

[229] *See* Richard Lempert, *The Honest Scientist's Guide to DNA Evidence*, 96 Genetica 119 (1995). If the jury accepted these premises and also decided to accept the hypothesis of identity over those of kinship and coincidence, it still would be open to the defendant to offer explanations of how the forensic samples came to include his or her DNA even though he or she is innocent.

ratios for all possible degrees of kinship and subpopulations.[230] However, it is not clear how this ambitious proposal would be implemented.[231]

D. Which Probabilities or Statistics Should Be Presented?

Up to this point, we have described probabilities that can be used in evaluating the extent to which the discovery that the trace evidence sample contains DNA of the same type as the defendant's establishes that this DNA came from the defendant. We have concentrated on the methods that are available to compute the probabilities, and we have examined the concerns that have been voiced about the validity of these methods. This section discusses the legal question regarding which of the various scientifically defensible probabilities should be admissible in court. Assuming that the probabilities are computed according to a method that meets *Daubert*'s demand for scientific validity and reliability and thus satisfies Rule 702, the major issue arises under Rule 403: To what extent will the presentation assist the jury to understand the meaning of a match so that the jury can give the evidence the weight it deserves? This question involves psychology and law, and we summarize the assertions and analyses that have been offered with respect to the various probabilities and statistics that can be used to indicate the probative value of DNA evidence.

1. Should Match Probabilities Be Excluded?

Are small frequencies or probabilities inherently prejudicial? The most common form of expert testimony about matching DNA takes the form of an explanation of how the laboratory ascertained that the defendant's DNA has the profile of the forensic sample plus an estimate of the profile frequency or random match probability. Many arguments have been offered against this entrenched practice. First, it has been suggested that jurors do not understand probabilities in general,[232] and infinitesimal match probabilities[233] will so bedazzle jurors that they will not appreciate the other evidence in the case or any innocent explanations for the

---

[230] *See* Balding & Donnelly, *supra* note 174.

[231] A related proposal in Lempert, *supra* note 166, suffers from the same difficulty of articulating the composition of the suspect population and the prior probabilities for its members. Professor Lempert reasons that "the relevant match statistic, if it could be derived, is an average that turns on the number of people in the suspect population and a likelihood that each has DNA matching the defendant's DNA, weighted by the probability that each committed the crime if the defendant did not." *Id.* at 458. He concludes that although this "weighted average statistic" does not directly state how likely it is "that the defendant and not some third party committed the crime," it is superior to "the 'random man' match statistic" in that it "tells the jury how surprising it would be to find a DNA match if the defendant is innocent." *Id.*

[232] *E.g.*, R.C. Lewontin, *Forensic DNA Typing Dispute*, 372 Nature 398 (1994).

[233] There have been cases in which the reported population frequencies are measured in the billionths or even trillionths. *E.g.*, Perry v. State, 606 So. 2d 224, 225 (Ala. Crim. App. 1992) ("one in 12

match.[234] Empirical research into this hypothesis has been limited and inconclusive,[235] and remedies short of exclusion are available.[236] Thus, no jurisdiction currently excludes all match probabilities on this basis.[237]

A more sophisticated variation on this theme is that the jury will misconstrue the random match probability—by thinking that it gives the probability that the match is random.[238] Suppose that the random match probability p is some very small number such as one in a billion. The words are almost identical, but the probabilities can be quite different. The random match probability is the probability that (A) the requisite genotype is in the sample from the individual tested *if* (B) the individual tested has been selected at random. In contrast, the probability that the match is random is the probability that (B) the individual tested has been selected at random *given that* (A) the individual has the requisite genotype. In general, for two events A and B, P(A *given* B) does not equal P(B *given*

---

billion"); Snowden v. State, 574 So. 2d 960, 960 (Ala. Crim. App. 1990) ("'approximately one in eleven billion,' with a 'minimum value' of one in 2.5 billion and a 'maximum' value of one in 27 trillion"); State v. Bible, 858 P.2d 1152, 1191 (Ariz. 1993) (between one in 60 million and one in 14 billion); State v. Daughtry, 459 S.E.2d 747, 758–59 (N.C. 1995) ("one in 5.5 billion for each of the caucasion, African-American, and Lumbee populations in North Carolina"); State v. Buckner, 890 P.2d 460, 460 (Wash. 1995) ("one Caucasian in 19.25 billion").

[234] *Cf.* Government of the Virgin Islands v. Byers, 941 F. Supp. 513, 527 (D.V.I. 1996) ("Vanishingly small probabilities of a random match may tend to establish guilt in the minds of jurors and are particularly suspect."); Commonwealth v. Curnin, 565 N.E.2d 440, 441 (Mass. 1991) ("evidence of this nature [a random-match probability of 1 in 59 million] . . . , having an aura of infallibility, must have a strong impact on a jury").

[235] *See* NRC II, *supra* note 1, at 197; Jason Schklar & Shari Seidman Diamond, *Juror Reactions to DNA Evidence: Errors and Expectancies*, 23 Law & Hum. Behav. 159, 181–82 (1999).

[236] Suitable cross-examination, defense experts, and jury instructions might reduce the risk that small estimates of the match probability will produce an unwarranted sense of certainty and lead a jury to disregard other evidence. NRC II, *supra* note 1, at 197

[237] *E.g.*, United States v. Chischilly, 30 F.3d 1144 (9th Cir. 1994) (citing cases); Martinez v. State, 549 So. 2d 694, 694–95 (Fla. Dist. Ct. App. 1989) (rejecting the argument that testimony that "one individual in 234 billion" would have the same banding pattern was "so overwhelming as to deprive the jury of its function"); State v. Weeks, 891 P.2d 477, 489 (Mont. 1995) (rejecting the argument that "the exaggerated opinion of the accuracy of DNA testing is prejudicial, as juries would give undue weight and deference to the statistical evidence" and "that the probability aspect of the DNA analysis invades the province of the jury to decide the guilt or innocence of the defendant"); State v. Schweitzer, 533 N.W.2d 156, 160 (S.D. 1995) (reviewing cases).

[238] Numerous opinions or experts present the random match probability in this manner. Compare the problematic characterizations in, *e.g.*, *United States v. Martinez*, 3 F.2d 1191, 1194 (8th Cir. 1993) (referring to "a determination of the probability that someone other than the contributor of the known sample could have contributed the unknown sample"), and *State v. Foster*, 910 P.2d 848 (Kan. 1996) (a DNA analyst testified that "the probability of another person in the Caucasian population having the same banding pattern was 1 in 100,000"), with the more accurate comments of an FBI examiner in *State v. Freeman*, No. A-95-1027, 1996 WL 608328, at *7 (Neb. Ct. App. Oct. 22, 1996), *aff'd*, 571 N.W.2d 276 (Neb. 1997), that "[t]he probability of randomly selecting an unrelated individual from the Caucasian population who would have the same DNA profile as I observed in the K2 sample for Mr. Freeman was approximately one in 15 million." For more examples of mischaracterizations of the random match probability, see cases and authorities cited, NRC II, *supra* note 1, at 198 n.92.

A). The claim that it does is known as the fallacy of the transposed conditional.[239]

To appreciate that the equation is fallacious, consider the probability that a lawyer picked at random from all lawyers in the United States is a federal judge. This "random judge probability" is practically zero. But the probability that a person randomly selected from the current federal judiciary is a lawyer is one. The "random judge probability" P(judge *given* lawyer) does not equal the transposed probability P(lawyer *given* judge). Likewise, the random match probability P(genotype *given* unrelated source) does not necessarily equal P(unrelated source *given* genotype).

To avoid this fallacious reasoning by jurors, some defense counsel have urged the exclusion of random match probabilities, and some prosecutors have suggested that it is desirable to avoid testimony or argument about probabilities, and instead to present the statistic as a simple frequency—an indication of how rare the genotype is in the relevant population.[240] The 1996 NRC report noted that "few courts or commentators have recommended the exclusion of evidence merely because of the risk that jurors will transpose a conditional probability,"[241] and it observed that "[t]he available research indicates that jurors may be more likely to be swayed by the 'defendant's fallacy' than by the 'prosecutor's fallacy.' When advocates present both fallacies to mock jurors, the defendant's fallacy dominates."[242] Furthermore, the committee suggested that "if the initial presentation of the probability figure, cross-examination, and opposing testimony all fail to clarify the point, the judge can counter both fallacies by appropriate instructions to the jurors that minimize the possibility of cognitive errors."[243]

---

[239] It is also called the "inverse fallacy," or the "prosecutor's fallacy." The latter expression is rare in the statistical literature, but it is common in the legal literature on statistical evidence. For an exposition of related errors, see Koehler, *supra* note 161.

[240] George W. Clark, *Effective Use of DNA Evidence in Jury Trials*, Profiles in DNA, Aug. 1997, at 7, 8 ("References to probabilities should normally be avoided, inasmuch as such descriptions are frequently judicially equated with disfavored "probabilities of guilt. . . . [T]he purpose of frequency data is simply to provide the factfinder with a guide to the relative rarity of a DNA match . . . .").

[241] NRC II, *supra* note 1, at 198 (citing McCormick on Evidence, *supra* note 11, § 212).

[242] *Id.* The "defendant's fallacy" consists of dismissing or undervaluing the matches with high likelihood ratios because other matches are to be expected in unrealistically large populations of potential suspects. For example, defense counsel might argue that (1) even with a random match probability of one in a million, we would expect to find ten unrelated people with the requisite genotypes in a population of 10 million; (2) the defendant just happens to be one of these ten, which means that the chances are nine out of ten that someone unrelated to the defendant is the source; so (3) the DNA evidence does nothing to incriminate the defendant. The problem with this argument is that in a case involving both DNA and non-DNA evidence against the defendant, it is unrealistic to assume that there are 10 million equally likely suspects.

[243] *Id.* (footnote omitted). The committee suggested the following instruction to define the random match probability:

> In evaluating the expert testimony on the DNA evidence, you were presented with a number indicating the

To date, no federal court has excluded a random match probability (or, for that matter, an estimate of the small frequency of a DNA profile in the general population) as unfairly prejudicial just because the jury might misinterpret it as a posterior probability that the defendant is the source of the forensic DNA. One court, however, noted the need to have the concept "properly explained,"[244] and prosecutorial misrepresentations of the random match probabilities for other types of evidence have produced reversals.[245]

Are small match probabilities irrelevant? Second, it has been maintained that match probabilities are logically irrelevant when they are far smaller than the probability of a frame-up, a blunder in labeling samples, cross-contamination, or other events that would yield a false positive.[246] The argument is that the jury should concern itself only with the chance that the forensic sample is reported to match the defendant's profile even though the defendant is not the source. Such a report could happen either because another person who is the source of the forensic sample has the same profile or because fraud or error of a kind that falsely incriminates the defendant occurs in the collection, handling, or analysis of the DNA samples. Match probabilities do not express this chance of a match being reported when the defendant is not the source unless the probability of a false-positive report is essentially zero.

Both theoretical and practical rejoinders to this argument about relevance have been given. At the theoretical level, some scientists question a procedure that would prevent the jury from reasoning in a stepwise, eliminative fashion. In their view, a rational juror might well want to know that the chance that another person selected at random from the suspect population has the incriminating genotype is negligible, for this would enable the juror to eliminate the hy-

---

> probability that another individual drawn at random from the [specify] population would coincidentally have the same DNA profile as the [blood stain, semen stain, etc.]. That number, which assumes that no sample mishandling or laboratory error occurred, indicates how distinctive the DNA profile is. It does not by itself tell you the probability that the defendant is innocent.

Id. at 198 n.93. *But see* D.H. Kaye, *The Admissibility of "Probability Evidence" in Criminal Trials—Part II*, 27 Jurimetrics J. 160, 168 (1987) ("Nevertheless, because even without misguided advice from counsel, the temptation to compute the probability of criminal identity [by transposition] seems strong, and because the characterization of the population proportion as a [random match probability] does little to make the evidence more intelligible, it might be best to bar the prosecution from having its expert state the probability of a coincidental misidentification, as opposed to providing [a simpler] estimate of the population proportion.").

[244] United States v. Shea, 957 F. Supp. 331, 345 (D.N.H. 1997).

[245] *E.g.*, United States v. Massey, 594 F.2d 676, 681 (8th Cir. 1979) (in closing argument about hair evidence, "the prosecutor 'confuse[d] the probability of concurrence of the identifying marks with the probability of mistaken identification'").

[246] *E.g.*, Jonathan J. Koehler et al., *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?,* 35 Jurimetrics J. 201 (1995); Lewontin & Hartl, *supra* note 184, at 1749 ("probability estimates like 1 in 738,000,000,000,000 . . . are terribly misleading because the rate of laboratory error is not taken into account").

potheses of kinship or coincidence.[247] If the juror concludes that there is little chance that the same genotype would exist in the forensic sample if the DNA originated from anyone but the defendant, then the juror can proceed to consider whether that genotype is present because someone has tried to frame the defendant, or whether it is not really present but was reported to be there because DNA samples were mishandled or misanalyzed.[248] These probabilities, they add, are not amenable to objective modeling and should not be mixed with probabilities that are derived from verifiable models of genetics.[249]

At the practical level, there is disagreement about the adequacy of the estimates that have been proposed to express the probability of a false positive result. The opponents of match probabilities usually argue that an error rate somewhat higher than that observed in a series of proficiency tests should be substituted for the match probability,[250] but the extent to which any such figure applies to the case at bar has been questioned.[251] No reported cases have excluded statistics on proficiency tests administered at a specific laboratory as too far removed from the case at bar to be relevant,[252] but neither has it been held that these statistics must be used in place of random match or kinship probabilities.[253]

---

[247] *E.g.*, NRC II, *supra* note 1, at 85; NRC I, *supra* note 1, at 88; Russell Higuchi, *Human Error in Forensic DNA Typing*, 48 Am. J. Hum. Genetics 1215 (1991) (letter). Of course, if the defense were to stipulate that a true DNA match establishes identity, there would be no need for probabilities that would help the jury to reject the rival hypotheses of coincidence or kinship.

[248] *E.g.*, Devlin & Roeder, *supra* note 154, § 18–5.3, at 743–44 ("One way to handle the possibility of a laboratory error, which follows the usual presentation of similar types of evidence, is to present the evidence in two stages: Does the evidence suggest that the samples were obtained from the same individual? If so, is there a harmless reason? Either formal calculations or informal analysis could be used to evaluate the possibility of a laboratory error, both of which should be predicated on the facts of the specific case.").

[249] *E.g.*, Morton, *supra* note 159, at 480–81; *cf.* NRC I, *supra* note 1, at 88 ("Coincidental identity and laboratory error are different phenomena, so the two cannot and should not be combined in a single estimate.").

[250] *But see* Thompson, *supra* note 69, at 417 (suggesting that "DNA evidence" should be excluded as "unacceptable scientifically if the probability of an erroneous match cannot be quantified").

[251] *See, e.g.*, David J. Balding, *Errors and Misunderstandings in the Second NRC Report*, 37 Jurimetrics J. 469, 475–76, 476 n.21 (1997) ("report[ing] a match probability which adds error rates to profile frequencies . . . would clearly be unacceptable since overall error rates are not directly relevant: jurors must assess on the basis of the evidence presented to them the chance that an error has occurred in the particular case at hand," but "[e]rror rates observed in blind trials may well be helpful to jurors"); Berger, *supra* note 69. *But cf.* Thompson, *supra* note 69, at 421 ("While it makes little sense to present a single number derived from proficiency tests as *the* error rate in every case, it makes less sense to exclude quantitative estimates of the error altogether.").

[252] *But see* United States v. Shea, 957 F. Supp. 331, 344 n.42 (D.N.H. 1997) ("The parties assume that error rate information is admissible at trial. This assumption may well be incorrect. Even though a laboratory or industry error rate may be logically relevant, a strong argument can be made that such evidence is barred by Fed. R. Evid. 404 because it is inadmissible propensity evidence.").

[253] *See* Armstead v. State, 673 A.2d 221 (Md. 1996) (rejecting the argument that the introduction of a random match probability deprives the defendant of due process because the error rate on proficiency

Are match probabilities unfairly prejudicial when they are smaller than the probability of laboratory error? It can be argued that very small match probabilities are relevant but unfairly prejudicial. Such prejudice could occur if the jury did not simply use a small match probability to reject the hypotheses of coincidence or kinship, but was so impressed with this single number that it neglected or underweighted the probability of a match arising due to a false-positive laboratory error.[254] Some commentators believe that this prejudice is so likely and so serious that "jurors ordinarily should receive *only* the laboratory's false positive rate . . . ."[255] The 1996 NRC report is skeptical of this view, especially when the defendant has had a meaningful opportunity to retest the DNA at a laboratory of his or her choice, and it suggests that judicial instructions can be crafted to avoid this form of prejudice.[256]

Are small match probabilities unfairly prejudicial when not accompanied by an estimated probability of a laboratory error? Rather than excluding small match probabilities entirely, a court might require the expert who presents them also to report a probability that the laboratory is mistaken about the profiles.[257] Of course, some experts would deny that they can provide a meaningful statistic for the case at hand, but they could report the results of proficiency tests and leave it to the jury to use this figure as best it can in considering whether a false-positive error has occurred.[258] To assist the jury in making sense of two num-

---

tests is many orders of magnitude greater than the match probability); Williams v. State, 679 A.2d 1106 (Md. 1996) (reversing because the trial court restricted cross-examination about the results of proficiency tests involving other DNA analysts at the same laboratory).

[254] *E.g.*, Koehler et al., *supra* note 246; Thompson, *supra* note 69, at 421–22.

[255] Richard Lempert, *Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 Cardozo L. Rev. 303, 325 (1991) (emphasis added); *see also* Lempert, *supra* note 166, at 447; Scheck, *supra* note 69, at 1997.

[256] NRC II, *supra* note 1, at 199 (notes omitted):

> The argument that jurors will make better use of a single figure for the probability that an innocent suspect would be reported to match has never been tested adequately. The argument for a single figure is weak in light of this lack of research into how jurors react to different ways of presenting statistical information, and its weakness is compounded by the grave difficulty of estimating a false-positive error rate in any given case. But efforts should be made to fill the glaring gap in empirical studies of such matters.

The district court in *United States v. Shea*, 957 F. Supp. 331, 334–45 (D.N.H. 1997), discussed some of the available research and rejected the argument that separate figures for match and error probabilities are prejudicial. For more recent research, see Schklar & Diamond, *supra* note 235, at 179 (concluding that separate figures are desirable in that "[j]urors . . . may need to know the disaggregated elements that influence the aggregated estimate as well as how they were combined in order to evaluate the DNA test results in the context of their background beliefs and the other evidence introduced at trial").

[257] Koehler, *supra* note 155, at 229 ("A good argument can be made for requiring DNA laboratories to provide fact finders with conservatively high estimates of their false positive error rates when they provide evidence about genetic matches. By the same token, laboratories should be required to divulge their estimated false negative error rate in cases where exclusions are reported."). This argument has prevailed in a few cases. *E.g.*, United States v. Porter, Crim. No. F06277-89, 1994 WL 742297 (D.C. Super. Ct. Nov. 17, 1994) (mem.). Other courts have rejected it. *E.g.*, United States v. Lowe, 954 F. Supp. 401, 415 (D. Mass. 1997), *aff'd*, 145 F.3d 45 (1st Cir. 1998).

[258] *See* NRC I, *supra* note 1, at 94 ("Laboratory error rates should be measured with appropriate

bers, however, it has been suggested that an expert take the additional step of reporting how the probability that a matching genotype would be found coincidentally *or* erroneously changes given the random match probability and various values for the probability of a false-positive error.[259]

2. Should Likelihood Ratios Be Excluded?

Likelihood ratios associated with DNA evidence were discussed in section VII.C.1. The 1996 NRC Report offers the following analysis of their admissibility:

> Although LRs [likelihood ratios] are rarely introduced in criminal cases, we believe that they are appropriate for explaining the significance of data and that existing statistical knowledge is sufficient to permit their computation. None of the LRs that have been devised for VNTRs can be dismissed as clearly unreasonable or based on principles not generally accepted in the statistical community. Therefore, legal doctrine suggests that LRs should be admissible unless they are so unintelligible that they provide no assistance to a jury or so misleading that they are unduly prejudicial. As with frequencies and match probabilities, prejudice might exist because the proposed LRs do not account for laboratory error, and a jury might misconstrue even a modified version that did account for it as a statement of the odds in favor of S [the claim that the defendant is the source of the forensic DNA sample]. [But] the possible misinterpretation of LRs as the odds in favor of identity . . . is a question of jury ability and performance to which existing research supplies no clear answer.[260]

---

proficiency tests and should play a role in the interpretation of results of forensic DNA typing. . . . A laboratory's overall rate of incorrect conclusions due to error should be reported with, but separately from, the probability of coincidental matches in the population. Both should be weighed in evaluating evidence."); NRC II, *supra* note 1, at 87 ("[A] calculation that combines error rates with match probabilities is inappropriate. The risk of error is properly considered case by case, taking into account the record of the laboratory performing the tests, the extent of redundancy, and the overall quality of the results."). The district court in *Government of the Virgin Islands v. Byers*, 941 F. Supp. 513 (D.V.I. 1996), declined to require proficiency test results as a precondition for admissibility. *See also* Berger, *supra* note 69, at 1093 ("the rationale for [requiring the prosecution to introduce a pooled error rate] is weak, and . . . such a shift would be inconsistent with significant evidentiary policies").

[259] *See* Thompson, *supra* note 69, at 421–22 (footnote omitted):

> For example, an expert could say that if the probability of a random match is .00000001 and the probability of an erroneous match is .001, then the overall probability of a false match is approximately .001. . . . If the probability of an erroneous match is unclear or controversial (as it undoubtedly will be in many cases), then illustrative combinations could be performed for a range of hypothetical probabilities.

This procedure could lead to arguments about the relevance of the values for the "probability of an erroneous match." Depending on such factors as the record of the laboratory on proficiency tests, the precautions observed in processing the samples, and the availability of the samples for independent testing, the prosecution could contend that the .001 figure in this example has no foundation in the evidence.

[260] NRC II, *supra* note 1, at 200–01. A footnote adds that:

> Likelihood ratios were used in State v. Klindt, 389 N.W.2d 670 (Iowa 1986) . . . , and are admitted routinely in parentage litigation, where they are known as the 'paternity index' . . . . Some state statutes use them to create a presumption of paternity . . . . The practice of providing a paternity index has been carried over into criminal cases in which genetic parentage is used to indicate the identity of the perpetrator of an offense. . . . .

Id. at 200 n.97.

Notwithstanding the lack of adequate empirical research, other commentators believe that the danger of prejudice (in the form of the transposition fallacy) warrants the exclusion of likelihood ratios.[261]

3. Should Posterior Probabilities Be Excluded?

Match probabilities state the chance that certain genotypes would be present conditioned on specific hypotheses about the source of the DNA (a specified relative, or an unrelated individual in a population or subpopulation). Likelihood ratios express the relative support that the presence of the genotypes in the defendant gives to these hypotheses compared to the claim that the defendant is the source. Posterior probabilities or odds express the chance that the defendant is the source (conditioned on various assumptions). These probabilities, if they are meaningful and accurate, would be of great value to the jury.

Experts have been heard to testify to posterior probabilities. In *Smith v. Deppish*,[262] for example, the state's "DNA experts informed the jury that . . . there was more than a 99 percent probability that Smith was a contributor of the semen,"[263] but how such numbers are obtained is not apparent. If they are instances of the transposition fallacy, then they are scientifically invalid (and objectionable under Rule 702) and unfairly prejudicial (under Rule 403).

However, a meaningful posterior probability can be computed with Bayes' theorem.[264] Ideally, one would enumerate every person in the suspect population, specify the prior odds that each is the source of the forensic DNA and weight those prior odds by the likelihoods (taking into account the familial relationship of each possible suspect to the defendant) to arrive at the posterior odds that the defendant is the source of the forensic sample. But this hardly seems practical. The 1996 NRC Report therefore discusses a somewhat different implementation of Bayes' theorem. Assuming that the hypotheses of kinship and error could be dismissed on the basis of other evidence, the report focuses on "the variable-prior-odds method," by which:

> an expert neither uses his or her own prior odds nor demands that jurors formulate their prior odds for substitution into Bayes's rule. Rather, the expert presents the jury with a

---

[261] *See* Koehler, *supra* note 168, at 880; Thompson, *supra* note 168, at 850; *cf.* Koehler et al., *supra* note 246 (proposing the use of a likelihood ratio that incorporates laboratory error).

[262] 807 P.2d 144 (Kan. 1991).

[263] *See also* Thomas v. State, 830 S.W.2d 546, 550 (Mo. Ct. App. 1992) (a geneticist testified that "the likelihood that the DNA found in Marion's panties came from the defendant was higher than 99.99%"); Commonwealth v. Crews, 640 A.2d 395, 402 (Pa. 1994) (an FBI examiner who at a preliminary hearing had estimated a coincidental-match probability for a VNTR match "at three of four loci" reported at trial that the match made identity "more probable than not").

[264] *See supra* § VII.C.2.

> table or graph showing how the posterior probability changes as a function of the prior probability.[265]

This procedure, it observes, "has garnered the most support among legal scholars and is used in some civil cases."[266] Nevertheless, "very few courts have considered its merits in criminal cases."[267] In the end, the report concludes:

> How much it would contribute to jury comprehension remains an open question, especially considering the fact that for most DNA evidence, computed values of the likelihood ratio (conditioned on the assumption that the reported match is a true match) would swamp any plausible prior probability and result in a graph or table that would show a posterior probability approaching 1 except for very tiny prior probabilities.[268]

E. Which Verbal Expressions of Probative Value Should Be Presented?

Having surveyed various views about the admissibility of the probabilities and statistics indicative of the probative value of DNA evidence, we turn to a related issue that can arise under Rules 702 and 403: Should an expert be permitted to offer a non-numerical judgment about the DNA profiles?

Inasmuch as most forms of expert testimony involve qualitative rather than quantitative testimony, this may seem an odd question. Yet, many courts have held that a DNA match is inadmissible unless the expert attaches a scientifically valid number to the figure.[269] In reaching this result, some courts cite the statement in the 1992 NRC report that "[t]o 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."[270]

---

[265] NRC II, *supra* note 1, at 202 (footnote omitted).

[266] *Id.*

[267] *Id.* (footnote omitted).

[268] *Id.* For arguments said to show that the variable-prior-odds proposal is "a bad idea," see Thompson, *supra* note 69, at 422–23.

[269] *E.g.*, Commonwealth v. Daggett, 622 N.E.2d 272, 275 n.4 (Mass. 1993) (plurality opinion insisting that "[t]he point is not that this court should require a numerical frequency, but that the scientific community clearly does"); State v. Carter, 524 N.W.2d 763, 783 (Neb. 1994) ("evidence of a DNA match will not be admissible if it has not been accompanied by statistical probability evidence that has been calculated from a generally accepted method"); State v. Cauthron, 846 P.2d 502 (Wash. 1993) ("probability statistics" must accompany testimony of a match); *cf.* Commonwealth v. Crews, 640 A.2d 395, 402 (Pa. 1994) ("The factual evidence of the physical testing of the DNA samples and the matching alleles, even without statistical conclusions, tended to make appellant's presence more likely than it would have been without the evidence, and was therefore relevant.").

[270] NRC I, *supra* note 1, at 74. For criticism of this statement, see Kaye, *supra* note 195, at 381–82 (footnote omitted):

> [I]t 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. Nor, when all estimates of the frequency are in the millionths or billionths, would it be meaningless

The 1996 report phrases the scientific question somewhat differently. Like the 1992 report, it states that "[b]efore forensic experts can conclude that DNA testing has the power to help identify the source of an evidence sample, it must be shown that the DNA characteristics vary among people. Therefore, it would not be scientifically justifiable to speak of a match as proof of identity in the absence of underlying data that permit some reasonable estimate of how rare the matching characteristics actually are."[271] However, the 1996 report then explains that "determining whether quantitative estimates should be presented to a jury is a different issue. Once science has established that a methodology has some individualizing power, the legal system must determine whether and how best to import that technology into the trial process."[272]

Since the loci typically used in forensic DNA identification have been shown to have substantial individualizing power, it is scientifically sound to introduce evidence of matching profiles. Nonetheless, even evidence that meets the scientific soundness standard of *Daubert* is not admissible if its prejudicial effect clearly outweighs its probative value. Unless some reasonable explanation accompanies testimony that two profiles match, it is surely arguable that the jury will have insufficient guidance to give the scientific evidence the weight that is deserves.[273]

Instead of presenting frequencies or match probabilities obtained with quantitative methods, however, a scientist would be justified in characterizing every four-locus VNTR profile, for instance, as "rare," "extremely rare," or the like.[274] At least one state supreme court has endorsed this qualitative approach as a substitute to the presentation of more debatable numerical estimates.[275]

The most extreme case of a purely verbal description of the infrequency of a profile arises when that profile can be said to be unique. The 1992 report cautioned that "an expert should—given . . . the relatively small number of loci

---

to inform the jury that there is a match that is known to be extremely rare in the general population. Courts may reach differing results on the legal propriety of qualitative as opposed to quantitative assessments, but they only fool themselves when they act as if scientific opinion automatically dictates the correct answer.

[271] NRC II, *supra* note 1, at 192. As indicated in earlier sections, these "underlying data" have been collected and analyzed for many genetic systems.

[272] *Id.*

[273] *Id.* at 193 ("Certainly, a judge's or juror's untutored impression of how unusual a DNA profile is could be very wrong. This possibility militates in favor of going beyond a simple statement of a match, to give the trier of fact some expert guidance about its probative value.").

[274] *Cf. id.* at 195 ("Although different jurors might interpret the same words differently, the formulas provided . . . produce frequency estimates for profiles of three or more loci that almost always can be conservatively described as 'rare.'").

[275] State v. Bloom, 516 N.W.2d 159, 166–67 (Minn. 1994) ("Since it may be pointless to expect ever to reach a consensus on how to estimate, with any degree of precision, the probability of a random match, and that given the great difficulty in educating the jury as to precisely what that figure means and does not mean, it might make sense to simply try to arrive at a fair way of explaining the significance of the match in a verbal, qualitative, non-quantitative, nonstatistical way."); *see also* Kenneth R. Kreiling, Review-Comment, *DNA Technology in Forensic Science*, 33 Jurimetrics J. 449 (1993).

used and the available population data—avoid assertions in court that a particular genotype is unique in the population."[276] Following this advice in the context of a profile derived from a handful of single-locus VNTR probes, several courts initially held that assertions of uniqueness are inadmissible,[277] while others found such testimony less troublesome.[278]

With the advent of more population data and loci, the 1996 NRC report pointedly observed that "we are approaching the time when many scientists will wish to offer opinions about the source of incriminating DNA."[279] Of course, the uniqueness of any object, from a snowflake to a fingerprint, in a population that cannot be enumerated never can be proved directly. The committee therefore wrote that "[t]here is no 'bright-line' standard in law or science that can pick out exactly how small the probability of the existence of a given profile in more than one member of a population must be before assertions of uniqueness are justified . . . . There might already be cases in which it is defensible for an expert to assert that, assuming that there has been no sample mishandling or laboratory error, the profile's probable uniqueness means that the two DNA samples come from the same person."[280]

---

[276] NRC I, *supra* note 1, at 92.

[277] *See* State v. Hummert, 905 P.2d 493 (Ariz. Ct. App. 1994), *rev'd*, 933 P.2d 1187 (1997); State v. Cauthron, 846 P.2d 502, 516 (Wash. 1993) (experts presented no "probability statistics" but claimed that the DNA could not have come from anyone else on earth), *overruled*, State v. Copeland, 922 P.2d 1304 (Wash. 1996); State v. Buckner, 890 P.2d 460, 462 (Wash. 1995) (testimony that the profile "would occur in only one Caucasian in 19.25 billion" and that because "this figure is almost four times the present population of the Earth, the match was unique" was improper), *aff'd on reconsideration*, 941 P.2d 667 (Wash. 1997).

[278] State v. Zollo, 654 A.2d 359, 362 (Conn. App. Ct. 1995) (testimony that the chance "that the DNA sample came from someone other than the defendant was 'so small that . . . it would not be worth considering'" was not inadmissible as an opinion on an ultimate issue in the case "because his opinion could reasonably have aided the jury in understanding the [complex] DNA testimony"); Andrews v. State, 533 So. 2d 841, 849 (Fla. Ct. App. 1988) (geneticist "concluded that to a reasonable degree of scientific certainty, appellant's DNA was present in the vaginal smear taken from the victim"); People v. Heaton, 640 N.E.2d 630, 633 (Ill. App. Ct. 1994) (an expert who used the product rule to estimate the frequency at 1/52,600 testified over objection to his opinion that the "defendant was the donor of the semen"); State v. Pierce, No. 89-CA-30, 1990 WL 97596, at *2–3 (Ohio Ct. App. July 9, 1990) (affirming admission of testimony that the probability would be one in 40 billion "that the match would be to a random occurrence," and "[t]he DNA is from the same individual"), *aff'd*, 597 N.E.2d 107 (Ohio 1992); *cf.* State v. Bogan, 905 P.2d 515, 517 (Ariz. Ct. App. 1995) (it was proper to allow a molecular biologist to testify, on the basis of a PCR-based analysis that he "was confident the seed pods found in the truck originated from" a palo verde tree near a corpse); Commonwealth v. Crews, 640 A.2d 395, 402 (Pa. 1994) (testimony of an FBI examiner that he did not know of a single instance "where different individuals that are unrelated have been shown to have matching DNA profiles for three or four probes" was admissible under *Frye* despite an objection to the lack of a frequency estimate, which had been given at a preliminary hearing as 1/400).

[279] NRC II, *supra* note 1, at 194.

[280] As an illustration, the committee cited *State v. Bloom*, 516 N.W.2d 159, 160 n.2 (Minn. 1994), a case in which a respected population geneticist was prepared to testify that "in his opinion the nine-locus match constituted 'overwhelming evidence that, to a reasonable degree of scientific certainty, the

The report concludes that "[b]ecause the difference between a vanishingly small probability and an opinion of uniqueness is so slight, courts may choose to allow the latter along with, or instead of the former, when the scientific findings support such testimony."[281] Confronted with an objection to an assertion of uniqueness, a court may need to verify that a large number of sufficiently polymorphic loci have been tested.[282]

DNA from the victim's vaginal swab came from the [defendant], to the exclusion of all others.'" NRC II, *supra* note 1, at 194–95 n.84. *See also* People v. Hickey, 687 N.E.2d 910, 917 (Ill. 1997) (given the results of nine VNTR probes plus PCR-based typing, two experts testified that a semen sample originated from the defendant).

[281] NRC II, *supra* note 1, at 195. If an opinion as to uniqueness were simply tacked on to a statistical presentation, it might be challenged as cumulative. *Cf. id.* ("Opinion testimony about uniqueness would simplify the presentation of evidence by dispensing with specific estimates of population frequencies or probabilities. If the basis of an opinion were attacked on statistical grounds, however, or if frequency or probability estimates were admitted, this advantage would be lost.").

[282] The NAS committee merely suggested that a sufficiently small random match probability compared to the earth's population could justify a conclusion of uniqueness. The committee did not propose any single figure, but asked: "Does a profile frequency of the reciprocal of twice the earth's population suffice? Ten times? One hundred times?" *Id.* at 194. Another approach would be to consider the probability of recurrence in a close relative. *Cf.* Belin et al., *supra* note 171.

The FBI uses a slightly complex amalgam of such approaches. Rather than ask whether a profile probably is unique in the world's population, the examiner focuses on smaller populations that might be the source of the evidentiary DNA. When the surrounding evidence does not point to any particular ethnic group, the analyst takes the random match probability and multiplies it by ten (to account for any uncertainty due to population structure). The analyst then asks what the probability of generating a population of unrelated people as large as that of the entire U.S. (290 million people) that contains no duplicate of the evidentiary profile would be. If that "no-duplication" probability is one percent or less, the examiner must report that the suspect "is the source of the DNA obtained from [the evidentiary] specimen . . . ." Memorandum from Jenifer A.L. Smith to Laboratory, Oct. 1, 1997, at 3. Similarly, the FBI computes the no-duplication probability in each ethnic or racial subgroup that may be of interest. If that probability is 1% or less, the examiner must report that the suspect is the source of the DNA. *Id.* Finally, if the examiner thinks that a close relative could be the source, and these individuals cannot be tested, standard genetic formulae are used to find the probability of the same profile in a close relative, that probability is multiplied by ten, and the resulting no-duplication probability for a small family (generally ten or fewer individuals) is computed. Once again, if the no-duplication probability is no more than 1%, the examiner reports that the suspect is the source. *Id.* at 3–4. In an apparent genuflection to older cases requiring testifying physicians to have "a reasonable degree of medical certainty," the analyst must add the phrase "to a reasonable degree of scientific certainty" to the ultimate opinion that the suspect is the source. *Id.* at 2–4. This type of testimony is questioned in Evett & Weir, *supra* note 174, at 244.

VIII. Novel Applications of DNA Technology

Most routine applications of DNA technology in the forensic setting involve the identification of human beings— suspects in criminal cases, missing persons, or victims of mass disasters. However, inasmuch as DNA technology can be applied to the analysis of any kind of biological evidence containing DNA, and because the technology is advancing rapidly, unusual applications are inevitable. In cases in which the evidentiary DNA is of human origin, new methods of analyzing DNA will come into at least occasional use, and new loci or DNA polymorphisms will be used for forensic work. In other cases, the evidentiary DNA will come from non-human organisms—household pets,[283] wild animals,[284] insects,[285] even bacteria[286] and viruses.[287] These applications are directed either at distinguishing among species or at distinguishing among individuals (or subgroups) within a species. These two tasks can raise somewhat different scientific issues, and no single, mechanically applied test can be formulated to assess the validity of the diversity of applications and methods that might be encountered.

Instead, this section outlines and describes four factors that may be helpful in deciding whether a new application is scientifically sound. These are the novelty of the application, the validity of the underlying scientific theory, the validity of any statistical interpretations, and the relevant scientific community to consult in assessing the application. We illustrate these considerations in the context of three novel, recent applications of DNA technology to law enforcement:

- Although federal law prohibits the export of bear products, individuals in this country have offered to supply bear gall bladder for export to Asia, where it is prized for its supposed medicinal properties. In one investigation, the National Fish and Wildlife Forensic Laboratory, using DNA test-

---

[283] Ronald K. Fitten, *Dog's DNA May Be Key in Murder Trial: Evidence Likely to Set Court Precedent*, Seattle Times, Mar. 9, 1998, at A1, *available in* 1998 WL 3142721 (reporting a trial court ruling in favor of admitting evidence linking DNA found on the jackets of two men to a pit bull that the men allegedly shot and killed, along with its owners).

[284] For example, hunters sometimes claim that they have cuts of beef rather than the remnants of illegally obtained wildlife. These claims can be verified or refuted by DNA analysis. *Cf.* State v. Demers, 707 A.2d 276, 277–78 (Vt. 1997) (unspecified DNA analysis of deer blood and hair helped supply probable cause for search warrant to look for evidence of illegally hunted deer in defendant's home).

[285] Felix A.H. Sperling et al., *A DNA-Based Approach to the Identification of Insect Species Used for Postmortem Interval Estimation*, 39 J. Forensic Sci. 418 (1994).

[286] DNA testing of bacteria in food can help establish the source of outbreaks of food poisoning and thereby facilitate recalls of contaminated foodstuffs. *See* Jo Thomas, *Outbreak of Food Poisoning Leads to Warning on Hot Dogs and Cold Cuts*, N.Y. Times, Dec. 24, 1998.

[287] *See* State v. Schmidt, 699 So. 2d 448 (La. Ct. App. 1997) (where the defendant was a physician accused of murdering his former lover by injecting her with the AIDS virus, the state's expert witnesses established that PCR-based analysis of human HIV can be used to identify HIV strains so as to satisfy *Daubert*).

ing, determined that the material offered for export actually came from a pig, absolving the suspect of any export law violations.[288]

- In *State v. Bogan*,[289] a woman's body was found in the desert, near several palo verde trees. A detective noticed two seed pods in the bed of a truck that the defendant was driving before the murder. A biologist performed DNA profiling on this type of palo verde and testified that the two pods "were identical" and "matched completely with" a particular tree and "didn't match any of the [other] trees," and that he felt "quite confident in concluding that" the tree's DNA would be distinguishable from that of "any tree that might be furnished" to him. After the jury convicted the defendant of murder, jurors reported that they found this testimony very persuasive.[290]
- In *R. v. Beamish*, a woman disappeared from her home on Prince Edward Island, on Canada's eastern seaboard. Weeks later a man's brown leather jacket stained with blood was discovered in a plastic bag in the woods. In the jacket's lining were white cat hairs. After the missing woman's body was found in a shallow grave, her estranged common-law husband was arrested and charged. He lived with his parents and a white cat. Laboratory analysis showed the blood on the jacket to be the victim's, and the hairs were ascertained to match the family cat at ten STR loci. The defendant was convicted of the murder.[291]

A. Is the Application Novel?

The more novel and untested an application is, the more problematic is its introduction into evidence. In many cases, however, an application can be new to the legal system but be well established in the field of scientific inquiry from which it derives. This can be ascertained from a survey of the peer-reviewed scientific literature and the statements of experts in the field.[292]

---

[288] Interview with Dr. Edgard Espinoza, Deputy Director, National Fish and Wildlife Forensic Laboratory, in Ashland, Ore. (June 1998). Also, FDA regulations do not prohibit mislabeling of pig gall bladder.

[289] 905 P.2d 515 (Ariz. Ct. App. 1995).

[290] Brent Whiting, *Tree's DNA "Fingerprint" Splinters Killer's Defense*, Ariz. Republic, May 28, 1993, at A1, *available in* 1993 WL 8186972; *see also* Carol Kaesuk Yoon, *Forensic Science: Botanical Witness for the Prosecution*, 260 Science 894 (1993).

[291] *DNA Testing on Cat Hairs Helped Link Man to Slaying*, Boston Globe, Apr. 24, 1997, *available in* 1997 WL 6250745; Gina Kolata, *Cat Hair Finds Way into Courtroom in Canadian Murder Trial*, N.Y. Times, Apr. 24, 1997, at A5; Marilyn A. Menott-Haymond et al., *Pet Cat Hair Implicates Murder Suspect*, 386 Nature 774 (1997).

[292] Even though some applications are represented by only a few papers in the peer-reviewed literature, they may be fairly well established. The breadth of scientific inquiry, even within a rather specialized field, is such that only a few research groups may be working on any particular problem. A better gauge is the extent to which the genetic typing technology is used by researchers studying related

Applications designed specially to address an issue before the court are more likely to be truly novel and thus may be more difficult to evaluate. The studies of the gall bladder, palo verde trees, and cat hairs exemplify such applications in that each was devised solely for the case at bar.[293] In such cases, there are no published, peer-reviewed descriptions of the particular application to fall back on, but the analysis still could give rise to "scientific knowledge" within the meaning of *Daubert*.[294]

The novelty of an unusual application of DNA technology involves two components—the novelty of the analytical technique, and the novelty of applying that technique to the samples in question.[295] With respect to the analytical method, forensic DNA technology in the last two decades has been driven in part by the development of many new methods for the detection of genetic variation between species and between individuals within a species. The approaches outlined in table A-1 for the detection of genetic variation in humans—RFLP analysis of VNTR polymorphism, PCR, detection of VNTR and STR polymorphism by electrophoresis, and detection of sequence variation by probe hybridization or direct sequence analysis—have been imported from other research contexts. Thus, their use in the detection of variation in nonhuman species and of variation among species involves no new technology. DNA technology transcends organismal differences.

Some methods for the characterization of DNA variation widely used in studies of other species, however, are not used in forensic testing of human DNA. These are often called "DNA fingerprint" approaches. They offer a snapshot characterization of genomic variation in a single test, but they essentially presume that the sample DNA originates from a single individual, and this presumption cannot always be met with forensic samples.

The original form of DNA "fingerprinting" used electrophoresis, Southern blotting, and a multilocus probe that simultaneously recognizes many sites in the genome.[296] The result is comparable to what would be obtained with a

---

problems and the existence of a general body of knowledge regarding the nature of the genetic variation at issue.

[293] Of course, such evidence hardly is unique to DNA technology. *See, e.g.*, Coppolino v. State, 223 So. 2d 68 (Fla. Dist. Ct. App.), *appeal dismissed*, 234 So. 2d 120 (Fla. 1968) (holding admissible a test for the presence of succinylcholine chloride first devised for this case to determine whether defendant had injected a lethal dose of this curare-like anesthetic into his wife).

[294] 509 U.S. 579, 590 (1993) ("to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method").

[295] From its inception, both these aspects of forensic DNA testing have been debated. *See, e.g.*, 1 McCormick on Evidence, *supra* note 11, § 205, at 902; Thompson & Ford, *supra* note 183.

[296] The probes were pioneered by Alec Jeffreys. *See, e.g.*, Alec J. Jeffreys et al., *Individual-specific "Fingerprints" of Human DNA*, 316 Nature 76 (1985). In the 1980s, the "Jeffreys probes" were used for forensic purposes, especially in parentage testing. *See, e.g.*, D.H. Kaye, *DNA Paternity Probabilities*, 24 Fam. L.Q. 279 (1990).

"cocktail" of single-locus probes—one complex banding pattern sometimes analogized to a bar-code.[297] Probes for DNA fingerprinting are widely used in genetic research in non-human species.[298]

With the advent of PCR as the central tool in molecular biology, PCR-based "fingerprinting" methods have been developed. The two most widely used are the random amplified polymorphic DNA (RAPD) method[299] and the amplified fragment length polymorphism (AFLP) method.[300] Both give bar code-like patterns.[301] In RAPD analysis, a single, arbitrarily constructed, short primer amplifies many DNA fragments of unknown sequence.[302] AFLP analysis begins with a digestion of the sample DNA with a restriction enzyme followed by amplification of selected restriction fragments.[303]

Although the DNA fingerprinting procedures are not likely to be used in the analysis of samples of human origin, new approaches to the detection of genetic variation in humans as well as other organisms are under development. On the horizon are methods based on mass spectrometry and hybridization chip technology. As these or other methods come into forensic use, the best measure of scientific novelty will be the extent to which the methods have found their way into the scientific literature. Use by researchers other than those who developed them indicates some degree of scientific acceptance.

The second aspect of novelty relates to the sample analyzed. Two questions are central: Is there scientific precedent for testing samples of the sort tested in the particular case? And, what is known about the nature and extent of genetic variation in the tested organism and in related species? *Beamish*, the Canadian case involving cat hairs, illustrates both points. The nature of the sample—cat

---

[297] As with RFLP analysis in general, this RFLP fingerprinting approach requires a relatively good quality sample DNA. Degraded DNA results in a loss of some of the bars in the barcode-like pattern.

[298] *E.g.*, DNA Fingerprinting: State of the Science (S.D.J. Pena et al. eds., 1993). The discriminating power of a probe must be determined empirically in each species. The probes used by Jeffreys for human DNA fingerprinting, for instance, are less discriminating for dogs. A.J. Jeffreys & D.B. Morton, *DNA Fingerprints of Dogs and Cats*, 18 Animal Genetics 1 (1987).

[299] John Welsh & Michael McClelland, *Fingerprinting Genomes Using PCR with Arbitrary Primers*, 18 Nucleic Acids Res. 7213 (1990); John G.K. Williams et al., *DNA Polymorphisms Amplified by Arbitrary Primers Are Useful as Genetic Markers*, 18 Nucleic Acids Res. 6531 (1990).

[300] Pieter Vos et al., *AFLP: A New Technique for DNA Fingerprinting*, 23 Nucleic Acids Res. 4407 (1995).

[301] The identification of the seed pods in *State v. Bogan*, 905 P.2d 515 (Ariz. Ct. App. 1995), was accomplished with RAPD analysis. The general acceptance of this technique in the scientific community was not seriously contested. Indeed, the expert for the defense conceded the validity of RAPD in genetic research and testified that the state's expert had correctly applied the procedure. *Id.* at 520.

[302] Primers must be validated in advance to determine which give highly discriminating patterns for a particular species in question.

[303] Both the RAPD and AFLP methods provide reproducible results within a laboratory, but AFLP is more reproducible across laboratories. *See, e.g.*, C.J. Jones et al., *Reproducibility Testing of RAPD, AFLP and SSR Markers in Plants by a Network of European Laboratories*, 3 Molecular Breeding 381 (1997). This may be an issue if results from different laboratories must be compared.

hairs—does not seem novel, for there is ample scientific precedent for doing genetic tests on animal hairs.[304] But the use of STR testing to identify a domestic cat as the source of particular hairs was new. Of course, this novelty does not mean that the effort was scientifically unsound; indeed, as explained in the next section, the premise that cats show substantial microsatellite polymorphism is consistent with other scientific knowledge.

B. Is the Underlying Scientific Theory Valid?

Daubert does not banish novel applications of science from the courtroom, but it does demand that trial judges assure themselves that the underlying science is sound, so that the scientific expert is presenting scientific knowledge rather than speculating or dressing up unscientific opinion in the garb of scientific fact.[305] The questions that might be asked to probe the scientific underpinnings extend the line of questions asked about novelty: What is the principle of the testing method used? What has been the experience with the use of the testing method? What are its limitations? Has it been used in applications similar to those in the instant case—for instance, for the characterization of other organisms or other kinds of samples? What is known of the nature of genetic variability in the organism tested or in related organisms? Is there precedent for doing any kind of DNA testing on the sort of samples tested in the instant case? Is there anything about the organism, the sample, or the context of testing that would render the testing technology inappropriate for the desired application?[306] To illustrate the usefulness of these questions, we can return to the cases involving pig gall bladders, cat hairs, and palo verde seed pods.

Deciding whether the DNA testing is valid is simplest in the export case. The question there was whether the gall bladders originated from bear or from some other species. The DNA analysis was based on the approach used by evolutionary biologists to study relationships among vertebrate species. It relies on sequence variation in the mitochondrial cytochrome b gene. DNA sequence analysis is a routine technology, and there is an extensive library of cytochrome b sequence data representing a broad range of vertebrate species.[307] As for the sample

---

[304] *E.g.*, Russell Higuchi et al., *DNA Typing from Single Hairs*, 332 Nature 543, 545 (1988). Collection of hair is non-invasive and is widely used in wildlife studies where sampling in the field would otherwise be difficult or impossible. Hair also is much easier to transport and store than blood, a great convenience when working in the field. *Id.*

[305] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 590 (1993) ("The adjective 'scientific' implies a grounding in the methods and procedures of science. Similarly, the word 'knowledge' connotes more than subjective belief or unsupported speculation.").

[306] *But cf.* NRC I, *supra* note 1, at 72 (listing seven "requirements" for new forensic DNA tests to achieve "the highest standards of scientific rigor").

[307] If the bear cytochrome b gene sequence were not in the database, it would be obligatory for the proponents of the application to determine it and add it to the database, where it could be checked by other researchers.

material—the gall bladder—such cells may not have been used before, but gall bladder is simply another tissue from which DNA can be extracted.[308] Thus, although the application was novel in that an approach had to be devised to address the question at hand, each segment of the application rests on a solid foundation of scientific knowledge and experience. No great inferential leap from the known to the unknown was required to reach the conclusion that the gall bladder was from a pig rather than a bear.

The DNA analysis in *Beamish* required slightly more extrapolation from the known to the unknown. As indicated in the previous section, the use of cat hairs as a source of DNA was not especially novel, and the very factors that reveal a lack of novelty also suggest that it is scientifically valid to test the DNA in cat hairs. But we also observed that the use of STR typing to distinguish among cats was novel. Is such reasoning too great a leap to constitute scientific knowledge? A great deal is known about the basis and extent of genetic variation in cats and other mammals. In particular, microsatellite polymorphism is extensive in all mammalian species that have been studied, including other members of the cat family. Furthermore, by testing small samples from two cat populations, the researchers verified the loci they examined were highly polymorphic.[309] Thus, the novelty in using STR analysis to identify cats is not scientifically unsettling; rather, it extends from and fits with everything else that is known about cats and mammals in general. However, as one moves from well-studied organisms to ones about which little is known, one risks crossing the line between knowledge and speculation.

The DNA testing in *State v. Bogan*[310] pushes the envelope further. First, the genetic variability of palo verde trees had not been previously studied. Second, it was not known whether enough DNA could be extracted from seed pods to perform a genetic analysis. Both of these questions had to be answered by new testing. RAPD analysis, a well-established method for characterizing genetic variation within a species, demonstrated that palo verde trees were highly variable. Seed pods were shown to contain adequate DNA for RAPD analysis. Finally, a blind trial showed that RAPD profiles correctly identified individual

---

[308] There is a technical concern that the DNA extracted from a gall bladder might contain inhibitors that would interfere with the subsequent sequence analysis; however, this merely affects whether the test will yield a result, and not the accuracy of any result.

[309] One sample consisted of nineteen cats in Sunnyside, Prince Edward Island, where the crime occurred. *See* Commentary, *Use of DNA Analysis Raises Some Questions* (CBS radio broadcast, Apr. 24, 1997), *transcript available in* 1997 WL 5424082 ("19 cats obtained randomly from local veterinarians on Prince Edward Island"); Marjorie Shaffer, *Canadian Killer Captured by a Whisker from Parents' Pet Cat*, Biotechnology Newswatch, May 5, 1997, *available in* 1997 WL 8790779 ("the Royal Canadian Mounted Police rounded up 19 cats in the area and had a veterinarian draw blood samples"). The other sample consisted of nine cats from the United States. *DNA Test on Parents' Cat Helps Put Away Murderer*, Chi. Trib., Apr. 24, 1997, *available in* 1997 WL 3542042.

[310] 905 P.2d 515 (Ariz. Ct. App. 1995).

palo verde trees.[311] In short, the lack of pre-existing data on DNA fingerprints of palo verde trees was bridged by scientific experimentation that established the validity of the specific application.

The DNA analyses in all three situations rest on a coherent and internally consistent body of observation, experiment, and experience. That information was mostly pre-existing in the case of the gall bladder testing. Some information on the population genetics of domestic cats on Prince Edward's Island had to be generated specifically for the analysis in *Beamish*, and still more was developed expressly for the situation in the palo verde tree testing in *Bogan*. A court, with the assistance of suitable experts, can make a judgment as to scientific validity in these cases because the crucial propositions are open to critical review by others in the scientific community and are subject to additional investigation if questions are raised. Where serious doubt remains, a court might consider ordering a blind trial to verify the analytical laboratory's ability to perform the identification in question.[312]

C. Has the Probability of a Chance Match Been Estimated Correctly?

The significance of a human DNA match in a particular case typically is presented or assessed in terms of the probability that an individual selected at random from the population would be found to match. A small random match probability renders implausible the hypothesis that the match is just coincidental.[313] In *Beamish*, the random match probability was estimated to be one in many millions,[314] and the trial court admitted evidence of this statistic.[315] In

---

[311] The DNA in the two seed pods could not be distinguished by RAPD testing, suggesting that they fell from the same tree. The biologist who devised and conducted the experiments analyzed samples from the nine trees near the body and another nineteen trees from across the county. He "was not informed, until after his tests were completed and his report written, which samples came from" which trees. *Bogan*, 905 P.2d at 521. Furthermore, unbeknownst to the experimenter, two apparently distinct samples were prepared from the tree at the crime scene that appeared to have been abraded by the defendant's truck. The biologist correctly identified the two samples from the one tree as matching, and he "distinguished the DNA from the seed pods in the truck bed from the DNA of all twenty-eight trees except" that one. *Id.*

[312] *Cf. supra* note 311. The blind trial could be devised and supervised by a court-appointed expert, or the parties could be ordered to agree on a suitable experiment. *See* 1 McCormick on Evidence, *supra* note 11, § 203, at 867.

[313] *See supra* § VII.

[314] David N. Leff, *Killer Convicted by a Hair: Unprecedented Forensic Evidence from Cat's DNA Convinced Canadian Jury*, Bioworld Today, Apr. 24, 1997, *available in* 1997 WL 7473675 ("the frequency of the match came out to be on the order of about one in 45 million," quoting Steven O'Brien); *All Things Considered: Cat DNA* (NPR broadcast, Apr. 23, 1997), *available in* 1997 WL 12832754 ("it was less than one in two hundred million," quoting Steven O'Brien).

[315] *See also* Tim Klass, *DNA Tests Match Dog, Stains in Murder Case*, Portland Oregonian, Aug. 7, 1998, at D06 (reporting expert testimony in a Washington murder case that "the likelihood of finding

State v. Bogan,[316] the random match probability was estimated by the state's expert as one in a million and by the defense expert as one in 136,000, but the trial court excluded these estimates because of the then-existing controversy over analogous estimates for human RFLP genotypes.[317]

Estimating the probability of a random match or related statistics requires a sample of genotypes from the relevant population of organisms. As discussed in section VII, the most accurate estimates combine the allele frequencies seen in the sample according to formulae that reflect the gene flow within the population. In the simplest model for large populations of sexually reproducing organisms, mating is independent of the DNA types under investigation, and each parent transmits half of his or her DNA to the progeny at random. Under these idealized conditions, the basic product rule gives the multilocus genotype frequency as a simple function of the allele frequencies.[318] The accuracy of the estimates thus depends on the accuracy of the allele frequencies in the sample database and the appropriateness of the population genetics model.

1. How Was the Database Obtained?

Since the allele frequencies come from sample data, both the method of sampling and the size of the sample can be crucial. The statistical ideal is probability sampling, in which some objective procedure provides a known chance that each member of the population will be selected. Such random samples tend to be representative of the population from which they are drawn. In wildlife biology, however, the populations often defy enumeration, and hence strict random sampling rarely is possible. Still, if the method of selection is uncorrelated with the alleles being studied, then the sampling procedure is tantamount to random sampling with respect to those alleles.[319] Consequently, the key question about the method of sampling for a court faced with estimates based on a database of cats, dogs, or any such species, is whether that sample was obtained in some biased way—a way that would systematically tend to include (or exclude) organisms with particular alleles or genotypes from the database.

---

a 10-for-10 match in the DNA of a randomly chosen dog of any breed or mix would be one in 3 trillion, and the odds for a nine-of-10 match would be one in 18 billion").

[316] 905 P.2d 515 (Ariz. Ct. App. 1995).

[317] *Id.* at 520. The Arizona case law on this subject is criticized in Kaye, *supra* note 178.

[318] More complicated models account for the population structure that arises when inbreeding is common, but they require some knowledge of how much the population is structured. *See supra* § VII.

[319] Few people would worry, for example, that the sample of blood cells taken from their vein for a test of whether they suffer from anemia is not, strictly speaking, a random sample. The use of convenience samples from human populations to form forensic databases is discussed in, *e.g.*, NRC II, *supra* note 1, at 126–27, 186. Case law is collected *supra* note 179.

2. How Large Is the Sampling Error?

Assuming that the sampling procedure is reasonably structured to give representative samples with respect to those genotypes of forensic interest, the question of database size should be considered. Larger samples give more precise estimates of allele frequencies than smaller ones, but there is no sharp line for determining when a database is too small.[320] Instead, just as pollsters present their results within a certain margin of error, the expert should be able to explain the extent of the statistical error that arises from using samples of the size of the forensic database.[321]

3. How Was the Random Match Probability Computed?

As we have indicated, the theory of population genetics provides the framework for combining the allele frequencies into the final profile frequency. The frequency estimates are a mathematical function of the genetic diversity at each locus and the number of loci tested. The formulas for frequency estimates depend on the mode of reproduction and the population genetics of the species. For outbreeding sexually reproducing species,[322] under conditions that give rise to Hardy-Weinberg and linkage equilibrium, genotype frequencies can be estimated with the basic product rule.[323] If a species is sexually reproducing but given to inbreeding, or if there are other impediments to Hardy-Weinberg or linkage equilibrium, such genotype frequencies may be incorrect. Thus, the reasonableness of assuming Hardy-Weinberg equilibrium and linkage equilibrium depends on what and how much is known about the population genetics of the species.[324] Ideally, large population databases can be analyzed to verify independence of alleles.[325] Tests for deviations from the single-locus genotype

---

[320] The 1996 NRC Report refers to "at least several hundred persons," but it has been suggested that relatively small databases, consisting of fifty or so individuals, allow statistically acceptable frequency estimation for the common alleles. NRC II, *supra* note 1, at 114. A new, specially constructed database is likely to be small, but alleles can be a assigned a minimum value, resulting in conservative genotype frequency estimates. Ranajit Chakraborty, *Sample Size Requirements for Addressing the Population Genetic Issues of Forensic Use of DNA Typing*, 64 Human Biology 141, 156–57 (1992). Later, the NAS committee suggests that the uncertainty that arises "[i]f the database is small . . . can be addressed by providing confidence intervals on the estimates." NRC II, *supra* note 1, at 125.

[321] Bruce S. Weir, *Forensic Population Genetics and the NRC*, 52 Am. J. Hum. Genetics 437 (1993) (proposing interval estimate of genotype frequency); *cf.* NRC II, *supra* note 1, at 148 (remarking that "calculation of confidence intervals is desirable," but also examining the error that could be associated with the choice of a database on an empirical basis).

[322] Outbreeding refers to the propensity for individuals to mate with individuals who are not close relations.

[323] *See supra* § VII.

[324] In *State v. Bogan*, 905 P.2d 515 (Ariz. Ct. App. 1995), for example, the biologist who testified for the prosecution consulted with botanists who assured him that palo verde trees were an outcrossing species. *Id.* at 523–24.

[325] However, large, pre-existing databases may not be available for the populations of interest in

frequencies expected under Hardy-Weinberg equilibrium will indicate if population structure effects should be accorded serious concern. These tests, however, are relatively insensitive to minor population structure effects, and adjustments for possible population structure might be appropriate.[326] For sexually reproducing species believed to have local population structure, a sampling strategy targeting the relevant population would be best. If this is not possible, estimates based on the larger population might be presented with appropriate caveats. If data on the larger population are unavailable, the uncertainty implicit in basic product rule estimates should not be ignored, and less ambitious alternatives to the random match probability as a means for conveying the probative value of a match might be considered.[327]

A different approach may be called for if the species is not an outbreeding, sexually reproducing species. For example, many plants, some simple animals, and bacteria reproduce asexually. With asexual reproduction, most offspring are genetically identical to the parent. All the individuals that originate from a common parent constitute, collectively, a clone. The major source of genetic variation in asexually reproducing species is mutation.[328] When a mutation occurs, a new clonal lineage is created. Individuals in the original clonal lineage continue to propagate, and two clonal lineages now exist where before there was one. Thus, in species that reproduce asexually, genetic testing distinguishes clones, not individuals, and the product rule cannot be applied to estimate genotype frequencies for individuals. Rather, the frequency of a particular clone in a population of clones must be determined by direct observation. For example, if a rose thorn found on a suspect's clothing were to be identified as originating from a particular cultivar of rose, the relevant question becomes how common that variety of rose bush is and where it is located in the community.

---

these more novel cases. Analyses of the smaller, ad hoc databases are unlikely to be decisive. In *Beamish*, for instance, two cat populations were sampled. The sample of nineteen cats from Sunnyside, in Prince Edward Island, and the sample of nine cats from the United States revealed considerable genetic diversity; moreover, most of the genetic variability was between individual cats, not between the two populations of cats. There was no statistically significant evidence of population substructure, and there was no statistically significant evidence of linkage disequilibrium in the Sunnyside population. The problem is that with such small samples, the statistical tests for substructure are not very sensitive; hence, the failure to detect it is not strong proof that either the Sunnyside or the North American cat population is unstructured.

326. A standard correction for population structure is to incorporate a population structure parameter F_{ST} into the calculation. Such adjustments are described *supra* § VII. However, appropriate values for F_{ST} may not be known for unstudied species.

327. The "tree lineup" in *Bogan* represents one possible approach. Adapting it to *Beamish* would have produced testimony that the researchers were able to exclude all the other (28) cats presented to them. This simple counting, however, is extremely conservative.

328. Bacteria also can exchange DNA through several mechanisms unrelated to cell division, including conjugation, transduction, and transformation. Bacterial species differ in their susceptibility to undergo these forms of gene transfer.

In short, the approach for estimating a genotype frequency depends on the reproductive pattern and population genetics of the species. In cases involving unusual organisms, a court will need to rely on experts with sufficient knowledge of the species to verify that the method for estimating genotype frequencies is appropriate.

D. What Is the Relevant Scientific Community?

Even the most scientifically sophisticated court may find it difficult to judge the scientific soundness of a novel application without questioning appropriate scientists.[329] Given the great diversity of forensic questions to which DNA testing might be applied, it is not possible to define specific scientific expertises appropriate to each. If the technology is novel, expertise in molecular genetics or biotechnology might be necessary. If testing has been conducted on a particular organism or category of organisms, expertise in that area of biology may be called for. If a random match probability has been presented, one might seek expertise in statistics as well as the population biology or population genetics that goes with the organism tested. Given the penetration of molecular technology into all areas of biological inquiry, it is likely that individuals can be found who know both the technology and the population biology of the organism in question. Finally, where samples come from crime scenes, the expertise and experience of forensic scientists can be crucial. Just as highly focused specialists may be unaware of aspects of an application outside their field of expertise, so too scientists who have not previously dealt with forensic samples can be unaware of case-specific factors that can confound the interpretation of test results.

---

329. *See supra* § I.C.

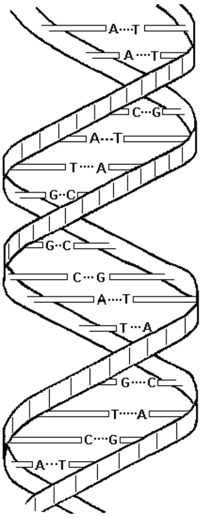
Appendix

A. Structure of DNA

DNA is a complex molecule made of subunits known as nucleotides that link together to form a long, spiraling strand. Two such strands are intertwined around each other to form a double helix as shown in Figure A-1. Each strand has a "backbone" made of sugar and phosphate groups and nitrogenous bases attached to the sugar groups.[330] There are four types of bases, abbreviated A, T, G, and C, and the two strands of DNA in the double helix are linked by weak chemical bonds such that the A in one strand is always paired to a T in the other strand and the G in one strand is always paired to a C in the other.[331] The A:T and G:C complementary base pairing means that knowledge of the sequence of one strand predicts the sequence of the complementary strand. The sequence of the nucleotide base pairs carries the genetic information in the DNA molecule—it is the genetic "text." For example, the sequence ATT on one strand (or TAA on the other strand) "means" something different than GTT (or CAA).

Figure A-1. A Schematic Diagram of the DNA Molecule

The bases in the nucleotide (denoted C, G, A, and T) are arranged like the rungs in a spiral staircase.



---

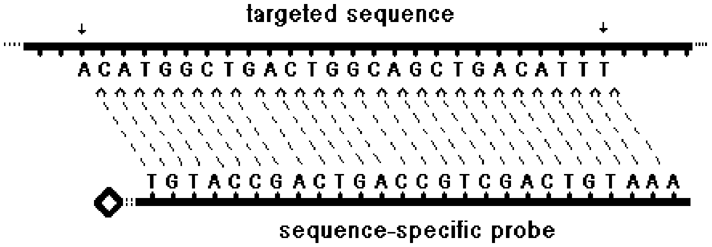
[330] For more details about DNA structure, see, *e.g.*, Anthony J.F. Griffiths et al., An Introduction to Genetic Analysis (6th ed. 1996); Mange & Mange, *supra* note 23, at 95.

[331] The bonds that connect the complementary bases are known as *hydrogen bonds*.

B. DNA Probes

A sequence specific oligonucleotide (SSO) probe is a short segment of single-stranded DNA with bases arranged in a particular order. The order is chosen so that the probe will bind to the complementary sequence on a DNA fragment, as sketched in Figure A-2.

Figure A-2. A Sequence-Specific Probe Links (Hybridizes) to the Targeted Sequence on a Single Stand of DNA



C. Examples of Genetic Markers in Forensic Identification

Table A-1 offers examples of the major types of genetic markers used in forensic identification.[332] As noted in the table, simple sequence polymorphisms, some variable number tandem repeat (VNTR) polymorphisms, and nearly all short tandem repeat (STR) polymorphisms are detected using polymerase chain reaction (PCR) as a starting point. Most VNTRs containing long core repeats are too large to be amplified reliably by PCR and are instead characterized by restriction fragment length polymorphism (RFLP) analysis using Southern blotting. As a result of the greater efficiency of PCR-based methods, VNTR typing by RFLP analysis is fading from use.

---

[332] The table is adapted from NRC II, *supra* note 1, at 74.

Table A-1. Genetic Markers Used in Forensic Identification

Nature of variation at locus		
Locus example	Method of detection	Number of alleles
Variable number tandem repeat (VNTR) loci contain repeated core sequence elements, typically 15–35 base pairs (bp) in length. Alleles differ in the number of repeats and are distinguished on the basis of size.		
D2S44 (core repeat 31 bp)	Intact DNA digested with restriction enzyme, producing fragments that are separated by gel electrophoresis; alleles detected by Southern blotting followed by probing with locus-specific radioactive or chemiluminescent probe	At least 75 (size range is 700–8500 bp); allele size distribution is essentially continuous
D1S80 (core repeat 16 bp)	Amplification of allelic sequences by PCR; discrete allelic products separated by electrophoresis and visualized directly	About 30 (size range is 350–1000 bp); alleles can be discretely distinguished
Short tandem repeat (STR) loci are VNTR loci with repeated core sequence elements 2–6 bp in length. Alleles differ in the number of repeats and are distinguished on the basis of size.		
HUMTHO1 (tetranucleotide repeat)	Amplification of allelic sequences by PCR; discrete allelic products separated by electrophoresis on sequencing gels and visualized directly, by capillary electrophoresis, or by other methods	8 (size range 179–203 bp); alleles can be discretely distinguished
Simple sequence variation (nucleotide substitution in a defined segment of a sequence)		
DQA (an expressed gene in the histocompatibility complex)	Amplification of allelic sequences by PCR; discrete alleles detected by sequence-specific probes	8 (6 used in DQA kit)
Polymarker (a set of five loci)	Amplification of allelic sequences by PCR; discrete alleles detected by sequence-specific probes	Loci are bi- or triallelic; 972 genotypic combinations
Mitochondrial DNA control region (D-loop)	Amplification of control-sequence and sequence determination	Hundreds of sequence variants are known

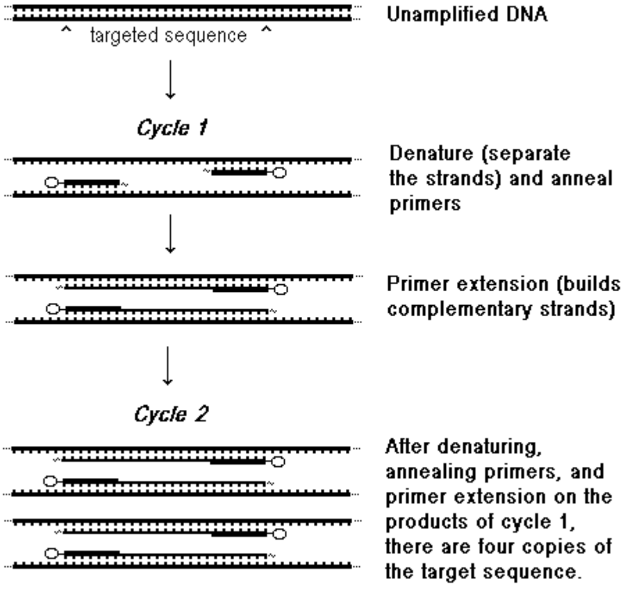
D. Steps of PCR Amplification

The second National Research Council report provides a concise description of how PCR "amplifies" DNA:

> First, each double-stranded segment is separated into two strands by heating. Second, these single-stranded segments are hybridized with primers, short DNA segments (20–30 nucleotides in length) that complement and define the target sequence to be amplified. Third, in the presence of the enzyme DNA polymerase, and the four nucleotide building blocks (A, C, G, and T), each primer serves as the starting point for the replication of the target sequence. A copy of the complement of each of the separated strands is made, so that there are two double-stranded DNA segments. The three-step cycle is repeated, usually 20–35 times. The two strands produce four copies; the four, eight copies; and so on until the number of copies of the original DNA is enormous. The main difference between this procedure and the normal cellular process is that the PCR process is limited to the amplification of a small DNA region. This region is usually not more than 1,000 nucleotides in length, so PCR methods cannot, at least at present, be used [to amplify] large DNA regions, such as most VNTRs. [333]

Figure A-3 illustrates the steps in the PCR process for two cycles.[334]

Figure A-3. The PCR Process


Unamplified DNA
^ targeted sequence ^
Cycle 1
Denature (separate the strands) and anneal primers
Primer extension (builds complementary strands)
Cycle 2
After denaturing, annealing primers, and primer extension on the products of cycle 1, there are four copies of the target sequence.

---

[333] NRC II, *supra* note 1, at 69–70.

In principle, PCR amplification doubles the number of double-stranded DNA fragments each cycle. Although there is some inefficiency in practice, the yield from a 30-cycle amplification is generally about one million to ten million copies of the targeted sequence.

E. Quantities of DNA in Forensic Samples

Amounts of DNA present in some typical kinds of evidence samples are indicated in Table A-2. These are approximate, and the quantities of DNA extracted from evidence in particular cases may vary somewhat.[335]

Table A-2. DNA Content of Biological Samples[336] and Genetic Testing Success Rates

Type of Sample	DNA Content	PCR Success Rate
Blood	20,000–40,000 ng/mL	
stain 1 cm x 1 cm	ca. 200 ng	> 95%
stain 1 mm x 1 mm	ca. 2 ng	
Semen	150,000–300,000 ng/mL	
on post-coital vaginal swab	0–3000 ng	>95%
Saliva	1000–10,000 ng/mL	
on a cigarette butt	0–25ng	50–70%
Hair		
root end of pulled hair	1–750 ng	>90%
root end of shed hair	1–12 ng	<20%
hair shaft	0.001–0.040 ng/cm	
Urine	1–20 ng/mL	
Skin cells		
from socks, gloves, or clothing repeatedly used		30–60%
from handled objects (e.g., a doorknob)		<20%

ng = nanogram, or 1/1,000,000,000th of a gram; mL = milliliter; cm = centimeter; mm = millimeter

---

[334] The figure is adapted from NRC I, *supra* note 1, at 41, fig. 1-6.

[335] The amounts in the table are given in nanograms (ng) or ng per milliliter (ng/mL). A nanogram is one billionth (1/1,000,000,000) of a gram.

[336] Adapted from NRC I, *supra* note 1, at 28 (with additions); PCR genetic test success rate estimates from the New York City Office of the Chief Medical Examiner, Department of Forensic Biology.

Glossary of Terms

adenine (A). One of the four bases, or nucleotides, that make up the DNA molecule. Adenine only binds to thymine. See nucleotide.

affinal method. A method for computing the single locus profile probabilities for a theoretical subpopulation by adjusting the single locus profile probability, calculated with the product rule from the mixed population database, by the amount of heterogeneity across subpopulations. The model is appropriate even if there is no database available for a particular subpopulation, and the formula always gives more conservative probabilities than the product rule applied to the same database.

allele. In classical genetics, an allele is one of several alternative forms of a gene. A biallelic gene has two variants; others have more. Alleles are inherited separately from each parent, and for a given gene, an individual may have two different alleles (heterozygosity) or the same allele (homozygosity). In DNA analysis, the term is applied to any DNA region (whether or not it constitutes a gene) used for analysis.

Alu sequences. A family of short interspersed elements (SINEs) distributed throughout the genomes of primates.

amplification. Increasing the number of copies of a DNA region, usually by PCR.

amplified fragment length polymorphism (AMP-FLP). A DNA identification technique that uses PCR-amplified DNA fragments of varying lengths. The DS180 locus is a VNTR whose alleles can be detected with this technique.

antibody. A protein (immunoglobulin) molecule, produced by the immune system, that recognizes a particular foreign antigen and binds to it; if the antigen is on the surface of a cell, this binding leads to cell aggregation and subsequent destruction.

antigen. A molecule (typically found in the surface of a cell) whose shape triggers the production of antibodies that will bind to the antigen.

autoradiograph (autoradiogram, autorad). In RFLP analysis, the x-ray film (or print) showing the positions of radioactively marked fragments (bands) of DNA, indicating how far these fragments have migrated, and hence their molecular weights.

autosome. A chromosome other than the X and Y sex chromosomes.

band. See autoradiograph.

band shift. Movement of DNA fragments in one lane of a gel at a different rate than fragments of an 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.

base pair (bp). Two complementary nucleotides bonded together at the matching bases (A and T or C and G) along the double helix "backbone" of the DNA molecule. 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.

Bayes' theorem. An elementary formula that relates certain conditional probabilities. It can be used to describe the impact of new data on the probability that a hypothesis is true.

bin, fixed. In VNTR profiling, a bin is a range of base pairs (DNA fragment lengths). When a database is divided into fixed bins, the proportion of bands within each bin is determined and the relevant proportions are used in estimating the profile frequency.

bins, floating. In VNTR profiling, a bin is a range of base pairs (DNA fragment lengths). In a floating bin method of estimating a profile frequency, the bin is centered on the base pair length of the allele in question, and the width of the bin can be defined by the laboratory's matching rule (e.g., ±5% of band size).

binning. Grouping VNTR alleles into sets of similar sizes because the alleles' lengths are too similar to differentiate.

blind proficiency test. See proficiency test.

capillary electrophoresis. A method for separating DNA fragments (including STRs) according to their lengths. A long, narrow tube is filled with an entangled polymer or comparable sieving medium, and an electric field is applied to pull DNA fragments placed at one end of the tube through the medium. The procedure is faster and uses smaller samples than gel electrophoresis, and it can be automated.

ceiling principle. A procedure for setting a minimum DNA profile frequency proposed in 1992 by a committee of the National Academy of Science. One hundred persons from each of 15–20 genetically homogeneous populations spanning the range of racial groups in the United States are sampled. For each allele, the higher frequency among the groups sampled (or 5%, whichever is larger) is used in calculating the profile frequency. Compare interim ceiling principle.

chip. A miniaturized system for genetic analysis. One such chip mimics capillary electrophoresis and related manipulations. DNA fragments, pulled by small voltages, move through tiny channels etched into a small block of glass, silicon, quartz, or plastic. This system should be useful in analyzing STRs.

Another technique mimics reverse dot blots by placing a large array of oligonucleotide probes on a solid surface. Such hybridization arrays should be useful in identifying SNPs and in sequencing mitochondrial DNA.

chromosome. A rod-like structure composed of DNA, RNA, and proteins. Most normal human cells contain 46 chromosomes, 22 autosomes and a sex chromosome (X) inherited from the mother, and another 22 autosomes and one sex chromosome (either X or Y) inherited from the father. The genes are located along the chromosomes. See also homologous chromosomes.

coding DNA. A small fraction of the human genome contains the "instructions" for assembling physiologically important proteins. The remainder of the DNA is "non-coding."

CODIS (combined DNA index system). A collection of databases on STR and other loci of convicted felons maintained by the FBI.

complementary sequence. The sequence of nucleotides on one strand of DNA that corresponds to the sequence on the other strand. For example, if one sequence is CTGAA, the complementary bases are GACTT.

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

database. A collection of DNA profiles.

degradation. The breaking down of DNA by chemical or physical means.

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). The molecule that contains genetic information. 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.

diploid number. See haploid number.

D-loop. A portion of the mitochrondrial genome known as the "control region" or "displacement loop" instrumental in the regulation and initiation of mtDNA gene products.

DNA polymerase. The enzyme that catalyzes the synthesis of double-stranded DNA.

DNA probe. See probe

DNA profile. The alleles at each locus. For example, a VNTR profile is the pattern of band lengths on an autorad. A multilocus profile represents the combined results of multiple probes. *See* genotype.

DNA sequence. The ordered list of base pairs in a duplex DNA molecule or of bases in a single strand.

DQA. The gene that codes for a particular class of Human Leukocyte Antigen (HLA). This gene has been sequenced completely and can be used for forensic typing. See human leukocyte antigen.

DQ. The antigen that is the product of the DQA gene. *See* DQA, human leukocyte antigen.

EDTA. A preservative added to blood samples.

electrophoresis. See capillary electrophoresis, gel electrophoresis.

endonuclease. An enzyme that cleaves the phosphodiester bond within a nucleotide chain.

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.

enzyme. A protein that catalyzes (speeds up or slows down) a reaction.

ethidium bromide. A molecule that can intercalate into DNA double helices when the helix is under torsional stress. Used to identify the presence of DNA in a sample by its fluorescence under ultraviolet light.

fallacy of the transposed conditional. See transposition fallacy.

false match. Two samples of DNA that have different profiles could be declared to match if, instead of measuring the distinct DNA in each sample, there is an error in handling or preparing samples such that the DNA from a single sample is analyzed twice. The resulting match, which does not reflect the true profiles of the DNA from each sample, is a false match. Some people use "false match" more broadly, to include cases in which the true profiles of each sample are the same, but the samples come from different individuals. Compare true match. See also match, random match.

gel, agarose. A semisolid medium used to separate molecules by electrophoresis.

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

gene. A set of nucleotide base pairs on a chromosome that contains the "instructions" for controlling some cellular function such as making an enzyme. The gene is the fundamental unit of heredity; each simple gene "codes" for a specific biological characteristic.

gene frequency. The relative frequency (proportion) of an allele in a population.

genetic drift. Random fluctuation allele frequencies from generation to generation.

genetics. The study of the patterns, processes, and mechanisms of inheritance of biological characteristics.

genome. The complete genetic makeup of an organism, comprising roughly 100,000 genes in humans.

genotype. The particular forms (alleles) of a set of genes possessed by an organism (as distinguished from phenotype, which refers to how the genotype expresses itself, as in physical appearance). In DNA analysis, the term is applied to the variations within all DNA regions (whether or not they constitute genes) that are analyzed.

genotype, single locus. The alleles that an organism possesses at a particular site in its genome.

genotype, multilocus. The alleles that an organism possesses at several sites in its genome.

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

Hae III. A particular restriction enzyme.

haploid number. Human sex cells (egg and sperm) contain 23 chromosomes each. This is the haploid number. When a sperm cell fertilizes an egg cell, the number of chromosomes doubles to 46. This is the diploid number.

haplotype. A specific combination of linked alleles at several loci.

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.

heteroplasty. The condition in which some copies of mitochondrial DNA in the same individual have different base pairs at certain points.

heterozygous. Having a different allele at a given locus on each of a pair of homologous chromosomes. See allele. Compare homozygous.

homologous chromosomes. The 44 autosomes (non-sex chromosomes) in the normal human genome are in homologous pairs (one from each parent) that share an identical set of genes, but may have different alleles at the same loci.

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 DQA.

homozygous. Having the same allele at a given locus on each of a pair of homologous chromosomes. See allele. Compare heterozygous.

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.

independence. Two events are said to be independent if one is neither more nor less likely to occur when the other does.

interim ceiling principle. A procedure proposed in 1992 by a committee of the National Academy of Sciences for setting a minimum DNA profile frequency. For each allele, the highest frequency (adjusted upward for sampling error) found in any major racial group (or 10%, whichever is higher), is used in product-rule calculations. Compare ceiling principle.

kilobase (kb). One thousand bases.

linkage. The inheritance together of two or more genes on the same chromosome.

linkage equilibrium. A condition in which the occurrence of alleles at different loci is independent.

locus. A location in the genome, i.e., a position on a chromosome where a gene or other structure begins.

mass spectroscopy. The separation of elements or molecules according to their molecular weight. In the version being developed for DNA analysis, small quantities of PCR-amplified fragments are irradiated with a laser to form gaseous ions that traverse a fixed distance. Heavier ions have longer times of flight, and the process is known as "matrix-assisted laser desorption-ionization time-of-flight mass spectroscopy." MALDI-TOF-MS, as it is abbreviated, may be useful in analyzing STRs.

match. The presence of the same allele or alleles in two samples. Two DNA profiles are declared to match when they are indistinguishable in genetic type. For loci with discrete alleles, two samples match when they display the same set of alleles. For RFLP testing of VNTRs, two samples match when the pattern of the bands is similar and the positions of the corresponding bands at each locus fall within a preset distance. See match window, false match, true match.

match window. If two RFLP bands lie with a preset distance, called the match window, that reflects normal measurement error, they can be declared to match.

microsatellite. Another term for an STR.

minisatellite. Another term for a VNTR.

mitochondria. A structure (organelle) within nucleated (eukaryotic) cells that is the site of the energy producing reactions within the cell. Mitochondria contain their own DNA (often abbreviated as mtDNA), which is inherited only from mother to child.

molecular weight. The weight in grams of one mole of a pure, molecular substance.

monomorphic. A gene or DNA characteristic that is almost always found in only one form in a population.

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. Such probes rarely are used now in forensic applications in the United States.

multilocus profile. See profile.

multiplexing. Typing several loci simultaneously.

mutation. The process that produces a gene or chromosome set differing from the type already in the population; the gene or chromosome set that results from such a process.

nanogram (ng). A billionth of a gram.

nucleic acid. RNA or DNA.

nucleotide. A unit of DNA consisting of a base (A, C, G, or T) and attached to a phosphate and a sugar group; the basic building block of nucleic acids. See deoxyribonucleic acid.

nucleus. The membrane-covered portion of a eukaryotic cell containing most of the DNA and found within the cytoplasm.

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

paternity index. A number (technically, a likelihood ratio) that indicates the support that the paternity test results lend to the hypothesis that the alleged father is the biological father as opposed to the hypothesis that another man selected at random is the biological father. Assuming that the observed phenotypes correctly represent the phenotypes of the mother, child, and alleged father tested, the number can be computed as the ratio of the probability of the phenotypes under the first hypothesis to the probability under the second hypothesis. Large values indicate substantial support for the hypothesis of paternity; values near zero indicate substantial support for the hypothesis that someone other than the alleged father is the biological father; and values near unity indicate that the results do not help in determining which hypothesis is correct.

pH. A measure of the acidity of a solution.

phenotype. A trait, such as eye color or blood group, resulting from a genotype.

polymarker. A commercially marketed set of PCR-based tests for protein polymorphisms.

polymerase chain reaction (PCR). A process that mimics DNA's own replication processes to make up to millions of copies of short strands of genetic material in a few hours.

polymorphism. The presence of several forms of a gene or DNA characteristic in a population.

point mutation. See SNP.

population genetics. The study of the genetic composition of groups of individuals.

population structure. When a population is divided into subgroups that do not mix freely, that population is said to have structure. Significant structure can lead to allele frequencies being different in the subpopulations.

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 PCR-based analyses).

product rule. When alleles occur independently at each locus (Hardy-Weinberg equilibrium) and across loci (linkage equilibrium), the proportion of the population with a given genotype is the product of the proportion of each allele at each locus, times factors of two for heterozygous loci.

proficiency test. A test administered at a laboratory to evaluate its performance. In a blind proficiency study, the laboratory personnel do not know that they are being tested.

prosecutor's fallacy. See transposition fallacy.

protein. A class of biologically important molecules made up of a linear string of building blocks called amino acids. The directions for the synthesis of any particular protein are encoded in the DNA sequence of its gene.

quality assurance. A program conducted by a laboratory to ensure accuracy and reliability.

quality audit. A systematic and independent examination and evaluation of a laboratory's operations.

quality control. Activities used to monitor the ability of DNA typing to meet specified criteria.

random match. A match in the DNA profiles of two samples of DNA, where one is drawn at random from the population. See also random match probability.

random match probability. The chance of a random match. As it is usually used in court, the random match probability refers to the probability of a true match when the DNA being compared to the evidence DNA comes from a person drawn at random from the population. This random true match probability reveals the probability of a true match when the samples of DNA come from different, unrelated people.

random mating. The members of a population are said to mate randomly with respect to particular genes of DNA characteristics when the choice of mates is independent of the alleles.

recombination. In general, any process in a diploid or partially diploid cell that generates new gene or chromosomal combinations not found in that cell or in its progenitors.

reference population. The population to which the perpetrator of a crime is thought to belong.

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). Variation among people in the length of a segment of DNA cut at two restriction sites.

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.

Reverse Dot Blot. A detection method used to identify SNPs in which DNA probes are affixed to a membrane, and amplified DNA is passed over the probes to see if it contains the complementary sequence.

sequence-specific oligonucleotide (SSO) probe. Also, allele-specific oligonucleotide (ASO) probe. 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 electrophoretograms associated with RFLP analysis.

sequencing. Determining the order of base pairs in a segment of DNA.

short tandem repeat (STR). See variable number tandem repeat.

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.

SNP (single nucleotide polymorphism). A substitution, insertion, or deletion of a single base pair at a given point in the genome.

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.

transposition fallacy. Confusing the conditional probability of A given B [P(A|B)] with that of B given A [P(B|A)]. Few people think that the probability that a person speaks Spanish (A) given that he or she is a citizen of Chile (B) equals the probability that a person is a citizen of Chile (B) given that he or she speaks Spanish (A). Yet, many court opinions, newspaper articles, and even some expert witnesses speak of the probability of a matching DNA genotype (A) given that someone other than the defendant is the source of the crime scene DNA (B) as if it were the probability of someone else being the source (B) given the matching profile (A). Transposing conditional probabilities correctly requires Bayes' Theorem.

true match. Two samples of DNA that have the same profile should match when tested. If there is no error in the labeling, handling, and analysis of the samples and in the reporting of the results, a match is a true match. A true match establishes that the two samples of DNA have the same profile. Unless the profile is unique, however, a true match does not conclusively prove that the two samples came from the same source. Some people use "true match" more narrowly, to mean only those matches among samples from the same source. Compare false match. See also match, random match.

variable number tandem repeat (VNTR). A class of RFLPs due to 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. VNTRs are longer than STRs.

window. See match window.

X chromosome. See chromosome.

Y chromosome. See chromosome.

References on DNA

Ian W. Evett & Bruce S. Weir, Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists (1998).

Elaine Johnson Mange & Arthur P. Mange, Basic Human Genetics (2d ed. 1999).

National Research Council Committee on DNA Forensic Science: An Update, The Evaluation of Forensic DNA Evidence (1996).

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