NON-HUMAN DNA EVIDENCE

©1998 George Sensabaugh & D.H. Kaye*

CITATION: George Sensabaugh and D.H. Kaye, Non-human DNA Evidence, 38 Jurimetrics J. 1-16 (1998)

ABSTRACT: After a decade of use, the novelty of DNA evidence has worn off. Nonetheless, new methods of DNA analysis are being introduced, new loci within the human genome are being employed, and DNA samples from other organisms are being tested. When evidence made possible by these advances is presented in courts, judges must decide whether the newer procedures are scientifically sound or generally accepted in the scientific community. Focusing on nonhuman DNA evidence, this Article identifies and describes factors that courts should examine in passing on the admissibility of novel DNA evidence. The framework that we construct is neither rigid nor self-applying. Some understanding of and appreciation for the nature, structure, and process of scientific reasoning in general and of the special characteristics of forensic science in particular are necessary to evaluate the scientific quality of the evidence.

Most applications of DNA technology in the forensic setting involve the identification of human beings[2] -- suspects in criminal cases, missing persons, or victims of mass disasters.[3] But nothing limits forensic testing to human DNA. Indeed, the city of Melbourne, Australia, supposedly plans to use DNA analysis of cells in dog droppings to identify those pets that foul the city's pavements, parks and beaches.[4] And non-human DNA has played a major role in far more serious cases, ranging from homicide prosecutions to patent infringement litigation, with organisms as diverse as household pets,[5] livestock,[6] wild animals,[7] insects,[8] plants,[9] bacteria,[10] and viruses.[11]

This Article offers a framework for courts to assess the scientific soundness of non-human DNA testing.[12] We focus on scientific soundness because, in one way or another, the law requires courts to assure that scientific evidence is based on methods and reasoning that have been demonstrated to be capable of reaching correct conclusions.[13] The two most common legal standards for admitting scientific evidence are "scientific soundness" and "general acceptance."[14] Scientific soundness is the essential quality that the Supreme Court demanded in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[15] Under *Daubert*, a court must inquire directly into whether the method that generated the conclusions being offered into evidence is sound. General acceptance within the scientific community is important--but only indirectly--as one indication of soundness. In contrast, under the previously dominant standard of *Frye v. United States*,[16] which remains the applicable standard in many states, admissibility turns on the degree to which the relevant scientific community accepts the method as capable of producing valid and reliable conclusions. Even so, courts in *Frye* states often look to

scientific soundness as an indication of whether there is general acceptance. Thus, soundness is an important consideration in all jurisdictions, and it is the focal point of this Article.

In cases of non-human DNA evidence, the testing is aimed at differentiating among species or at distinguishing individuals within a species. These two tasks can raise somewhat different scientific issues, and no single, mechanically applied legal rule can be formulated to assess the validity of the diversity of applications and methods that might be encountered. However, there are general factors that always should be considered in deciding whether a given application is scientifically respectable. This Article outlines and describes four such factors: 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 testing, determined that the material offered for export actually came from a pig, absolving the suspect of any export law violations.[17]

- In *State v. Bogan*,[18] a woman's body was found in the desert, near several Palo Verde trees. A detective found 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. The jury convicted the defendant of murder, and jurors reportedly found this testimony very persuasive.[19]

- 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.[20]

A. NOVELTY OF THE APPLICATION

The more novel and untested an application is, the more problematic is its introduction into evidence. In many cases, however, an application that is new to the legal system is 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.[21]

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 pig gall bladder, Palo Verde trees, and cat hairs exemplify such applications in that each was devised solely for the case at bar.[22] 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 v. Merrell Dow Pharmaceuticals, Inc.*[23]

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. 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. Current approaches to the detection of genetic variation in humans--RFLP analysis of VNTR polymorphism, PCR, detection of VNTR and STR polymorphisms 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 non-human 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.[24] The result is comparable to what would be obtained with a "cocktail" of single-locus probes--one complex banding pattern sometimes analogized to a bar-code.[25] Probes for DNA fingerprinting are widely used in genetic research in non-human species.[26]

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[27] and the amplified fragment length polymorphism (AFLP) method.[28] Both give bar code-like patterns.[29] In RAPD analysis, a single, arbitrarily constructed, short primer amplifies many DNA fragments of unknown sequence.[30] AFLP analysis begins with a digestion of the sample DNA with a restriction enzyme followed by amplification of selected restriction fragments.[31]

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[32] and hybridization chip technology.[33] 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 hairs--does not seem novel, for there is ample scientific precedent for doing genetic tests on animal

hairs.[34] 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. VALIDITY OF THE UNDERLYING SCIENTIFIC THEORY

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 can be found to be presenting scientific knowledge rather than speculating or dressing up unscientific opinion in the garb of scientific fact.[35] 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?[36] To illustrate the usefulness of these questions, we can return to the cases involving pig gall bladders, cat hairs, and Palo Verde seeds.

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.[37] As for the sample 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.[38] 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 the *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.[39] 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* 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 Palo Verde trees.[40] 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.[41]

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 in *Bogan*, still more was developed expressly for the situation in the Palo Verde tree testing. A court, with the assistance of suitable experts, can make a judgment as to scientific validity in all 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.[42]

C. ESTIMATION OF THE PROBABILITY OF A CHANCE MATCH

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. In *Beamish*, the random match probability was estimated to be one in many millions,[43] and the trial court admitted evidence of this statistic.[44] In *Bogan*, the state's expert estimated the random match probability as one in a million and the defense expert estimated it 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.[45]

Estimating the probability of a random match or related statistics requires a sample of genotypes from the relevant population of organisms. However, because each possible genotype is itself very rare, enormously large samples would be required to estimate the relative frequency of any specific genotype simply by counting how often that genotype occurs. Subtler methods must be used. Each genotype consists of distinct DNA characteristics, called alleles, that occur at each of several locations, or loci, within the genome. Because these alleles are more common, their frequencies can be estimated with reasonable precision.[46] The most accurate estimates therefore 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 multilocus genotype frequency is a simple function of the allele frequencies.[47] 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.[48] 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.

2. How Large Is the Sampling Error?

Assuming that the sampling procedure is reasonably calculated 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.[49] 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.[50]

3. How Was the Random Match Probability or Genotype Frequency Computed?

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.[51] Briefly, for outbreeding sexually reproducing species engaged in "random mating," alleles are presumed to be in "Hardy-Weinberg equilibrium," and genotype frequencies can be estimated using the appropriate term of the Hardy-Weinberg equation.[52] Under the assumption that alleles at different loci are independent (linkage equilibrium), genotype frequencies at different loci can be multiplied together using the "product rule" to obtain a multilocus profile frequency.[53] If a species is sexually reproducing but given to inbreeding, then the genotype frequencies calculated according to the basic product rule 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.[54] Ideally, large population databases can be analyzed to verify independence of alleles.[55] Tests for deviations from the single-locus genotype 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.[56] 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 is 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.[57]

The foregoing discussion of efforts to estimate the random match probability with very limited information pertains to sexually reproducing organisms. Many plants, some simple animals, and all bacteria reproduce asexually.[58] 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. 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.

D. 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. 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.[59] 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 also scientists who have not previously dealt with forensic samples can be unaware of case-specific factors that can confound the interpretation of test results.

* * * * *

After a decade of debate, the novelty of DNA evidence has worn off, but the power and scope of forensic DNA testing continues to expand. As new methods of DNA analysis come into use, as new loci within the human genome become targets for analysis, and as DNA samples from other organisms are tested, new evidentiary issues will arise. We have attempted to identify and describe the factors that courts should attend to in passing on the admissibility of non-human DNA evidence. The framework that we have constructed looks to the nature of the testing, the source of the DNA, and the state of scientific knowledge about these matters. It can give structure and texture to judicial inquiries into the scientific soundness of non-human DNA evidence.

NOTES

* George F. Sensabaugh is Professor, School of Public Health, University of California at Berkeley. D.H. Kaye is Regents' Professor, Arizona State University College of Law. This article is adapted from the forthcoming second edition of the U.S. Federal Judicial Center's Reference Manual on Scientific Evidence. [BACK]

2. . Judicial acceptance of certain methods of DNA profiling to help establish whether an individual is the source of a bloodstain or other material containing DNA has become almost universal. Debates over population genetics have largely subsided. In concluding that the "product rule" for estimating genotype frequencies is generally accepted or scientifically sound in typical situations, appellate courts have relied on scientific publications appearing since 1993 and on the opinions of courts in other jurisdictions. Three publications frequently cited are (1) a five-volume compilation of data entitled *VNTR Population Data: A Worldwide Survey*, released by the FBI in 1993; (2) a paper boldly declaring that the "DNA fingerprinting wars are over" (Eric S. Lander & Bruce Budowle, *Commentary: DNA Fingerprinting Dispute Laid to Rest*, 371 Nature 735, 735 (1994)); and (3) the 1996 report of a committee of the National Academy of Sciences charged with updating the Academy's 1992 report. National Research Council Committee on Forensic DNA Science: An Update, The Evaluation of Forensic DNA Evidence 159 (1996) [hereinafter cited as NRC II] ("the interim ceiling principle is not needed and can be abandoned"). For a discussion of early cases relying on these publications, see D.H. Kaye, *DNA Identification in Criminal Cases: Lingering and Emerging Evidentiary Issues, in* Proceedings of the Seventh International Symposium on Human Identification (1997). *See also* Commonwealth v. Blasioli, 713 A.2d 1117 (Pa. 1998). Some recent opinions also express doubts about the depth of knowledge of the scientists who often testify for defendants. *E.g.*, State v. Copeland, 922 P.2d 1304, 1318 n.5 (Wash. 1996). With the perception that statistical and population genetics issues are not generally a major barrier to admissibility, argument has shifted to issues involving the presentation of evidence and the use of new loci for human DNA profiling. [BACK]

3. . *See, e.g.*, Joseph L. Galloway, *Identifying the Missing*, U.S. News & World Rep., Jan. 17, 1994, at 50; Steven Lee Myers, *'Unknown' Vietnam Soldier Now Has a Name,* N.Y. Times, June 30, 1998, at A1. [BACK]

4. . *Australians Scoop up Dog DNA*, Ariz. Republic, Sept. 26, 1998, at A27, *available in* 1998 WL 7799554. [BACK]

5. . Roger Croteau, *Experts ID Dog DNA in Mauling*, San Antonio Express-News, May 14, 1998, at 01B, *available in* 1998 WL 5091881 (reporting that a district attorney is considering filing misdemeanor charges against the owner of two dogs believed to have seriously mauled an elderly woman because the saliva on the victim's clothes was matched to one of the dogs); Patrick McMahon, *Dog DNA Contributes to Murder Convictions*, USA Today, Sept. 18, 1998, at 13A, *available in* 1998 WL 5736395 (reporting that the trial court admitted testimony that DNA found on the jackets of two men came from a pit bull that the men shot and killed, along with its owners).[BACK]

6. . Diego Beaumonte et al., *Microsatellite DNA Polymorphism Analysis in a Case of Illegal Cattle Purchase*, 40 J. Forensic Sci. 692 (1995) (paternity testing of cattle). [BACK]

7. . 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. Elizabeth A. Guglich et al., *Application of DNA Fingerprinting to Enforcement of Hunting Regulations in Ontario*, 38 J. Forensic Sci. 48 (1993); 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). For a few other applications involving wild animals, see C.S. Baker & S.R. Palumbi, *Which Whales Are Hunted? A Molecular Genetic Approach to Monitoring Whaling*, 265 Science 1538 (1994) (reporting on mitochondrial DNA analyses of commercial whale products to ascertain the extent to which protected species are still being hunted); Robert S. Blackett & Paul Keim, *Big Game Species Identification by Deoxyribonucleic Acid (DNA) Probes*, 37 J. Forensic Sci. 590 (1992); Elizabeth A. Guglich et al., *Forensic Application of Repetitive DNA Markers to the Species Identification of Animal Tissues*, 39 J. Forensic Sci. 353 (1994); J.L. Ramet al., *Authentication of Canned Tuna and Bonito by Sequence and Restriction Site Analysis of Polymerase Chain Reaction Products of Mitochondrial DNA*, 44 J. Agric. & Food Chemistry 2460 (1996) (mitochondrial DNA analysis can be used to distinguish canned tuna from canned bonito, aiding in the enforcement of a EEC rule that requires separate labeling for tuna and bonito). [BACK]

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

9. . DNA profiling has been used to establish the identity of plant cultivars, supporting claims of patent infringement. *E.g.*, B. Koller et al., *Identification of Apple Cultivars Using RAPD Markers*, 85 Theoretical & Applied Genetics 901 (1993); X. Yang & C. Quiros, *Identification and Classification of Celery Cultivars Using RAPD Markers*, 86 Theoretical & Applied Genetics 205 (1993). [BACK]

10. . DNA testing of bacteria in food can help establish the source of outbreaks of food poisoning and thereby facilitate recalls of contaminated foodstuffs. *See* James Brooke, *25 Million Pounds of Beef Recalled; Plant Is Closing over a Danger of Bacteria,* N.Y. Times, Aug. 22, 1997, at A1; *cf.* William J. Broad, *Norway's 1918 Lump of Sugar Yields Clues to Anthrax in War,* N.Y. Times, June 25, 1998, at A11. [BACK]

11. . 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*); Chin-Yih Ou et al., *Molecular Epidemiology of HIV Transmission in a Dental Practice*, 256 Science 1165 (1992) (a dentist allegedly transmitted the AIDS virus to his patients, and viral strains isolated from the dentist and from five of the seven infected patients were found to be very similar in DNA sequence); Gretchen Vogel, *Phylogenetic Analysis: Getting Its Day in Court*, 275 Science 1559 (1997) (HIV-infected blood). Although the DNA sequences used to establish the source of the infections were not identical with the sequences in the infected individuals, HIV is known to mutate rapidly. Even within an infected individual, the virus changes over time. The biological or statistical question, therefore, is how the viral sequence variation in individuals infected from a common source compares to the extent of variation among individuals infected from disparate sources. In the dental practice case, greater sequence variation was found in viruses in a random sample of AIDS patients than in the dentist's patients. These data were interpreted to support the allegation that the dentist had infected some of his patients. [BACK]

12. . We presuppose a basic familiarity with the nomenclature and ideas of DNA technology as it has been employed in criminal cases over the past decade. [BACK]

13. . In a trivial sense, all methods are at least *capable* of reaching correct conclusions. Psychics, tea leaves, dowsing rods, phrenology, graphology, voice stress analyzers, and all the other paraphernalia or practitioners of pseudoscience *sometimes* hit upon the truth. But their accuracy is coincidental (or based on other cues). In contrast, a scientific procedure acquires and uses information in a way that systematically produces correct answers. [BACK]

14. . *See generally* 1 Modern Scientific Evidence: The Law and Science of Expert Testimony (David L. Faigman et al. eds., 1997); 1 McCormick on Evidence § 203 (John W. Strong ed., 5th ed. forthcoming 1999). [BACK]

15. . 509 U.S. 579 (1993). The Court used the term "reliability," but its "reliability" is what scientists would recognize as an amalgam of validity and reliability. To avoid any possible confusion in the use of these terms, we use the broad, nontechnical phrase "scientific soundness." [BACK]

16. . 293 F. 1013 (D.C. Cir. 1923). [BACK]

17. . Interview with Dr. Edgard Espinoza, Deputy Director, National Fish and Wildlife Forensic Laboratory, Ashland, Oregon (June, 1998). Although FDA regulations do not prohibit mislabeling of pig gall bladder, this finding suggests that the exporter may have been defrauding its customers. [BACK]

18. . 905 P.2d 515 (Ariz. Ct. App. 1995). [BACK]

19. . Brent Whiting, *Tree's DNA "Fingerprint" Splinters Killer's Defense*, Ariz. Republic, May 28, 1993, at A1; C.K. Yoon, *Forensic Science--Botanical Witness for the Prosecution*, 260 Science 894 (1993). [BACK]

20. . *DNA Testing on Cat Hairs Helped Link Man to Slaying*, Boston Globe, Apr. 24, 1997, at A21, *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). [BACK]

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

22. . Of course, such evidence hardly is unique to DNA technology. *See, e.g.*, Coppolino v. State, 223 So.2d 68 (Fla. Ct. App.), *app. 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). [BACK]

23. . 509 U.S. 579, 590 (1993) ("to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method"). [BACK]

24. . 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). [BACK]

25. . As with RFLP analysis in general, this RFLP fingerprinting approach requires relatively good quality sample DNA. Degraded DNA results a loss of some of the bands. [BACK]

26. . *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 et al., *DNA Fingerprints of Dogs and Cats*, 18 Animal Genetics 1 (1987). [BACK]

27. . J. Welsh & M. McCelland; *Fingerprinting Genomes Using PCR with Arbitrary Primers*, 18 Nucleic Acids Res. 7213 (1990); J.G.K. Williams et al., *DNA Polymorphisms Amplified by Random Primers are Useful as Genetic Markers*, 18 Nucleic Acids Res. 6531 (1990). [BACK]

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

29. . 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. [BACK]

30. . Primers must be validated in advance to determine which ones give highly discriminating patterns for a particular species in question. More than one can be used. [BACK]

31. . Both the RAPD and AFLP methods provide reproducible results within a laboratory, but AFLP is regarded as being more reproducible at the inter-laboratory level. *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. [BACK]

32. . J.A. Monforte & C.H. Becker, *High-throughput DNA Analysis by Time-of-Flight Mass Spectrometry*, 3 Nature Medicine 360 (1997). [BACK]

33. . David G. Wang et al., *Large-scale Identification, Mapping, and Genotyping of Single-nucleotide Polymorphisms in the Human Genome*, 280 Science 1077 (1998). [BACK]

34. . *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.* at 545. [BACK]

35. . *See Daubert*, 509 U.S. at 590 ("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."). [BACK]

36. . *But cf.* National Research Council Committee on DNA Technology in Forensic Science, DNA Technology in Forensic Science 72 (1992) (listing seven "requirements" for new forensic DNA tests to achieve "the highest standards of scientific rigor"). [BACK]

37. . 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 to add it to the database, where it could be checked by other researchers. [BACK]

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

39. . 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. *See DNA Test on Parents' Cat Helps Put Away Murderer*, Chi. Trib., Apr. 24, 1997, *available in* 1997 WL 3542042. [BACK]

40. . 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.* [BACK]

41. . A question might still be raised regarding the propriety of the RAPD analysis to the particular sample. As noted in the previous section, RAPD analysis yields meaningful results only on DNA originating from a single individual source. If the Palo Verde seed pods were contaminated with another DNA source, a mold for example, the RAPD analysis would reflect the contribution of both the Palo Verde DNA and the contaminant DNA. For such samples, RAPD analysis is clearly not the best approach to use. Thus, a foundation should be laid for the appropriateness of the test to the sample. [BACK]

42. . *Cf. supra* note 39. 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 § 203, at 867 (John Strong ed., 4th ed. 1992). [BACK]

43. . 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). [BACK]

44. . *See also* Tim Klass, *DNA Tests Match Dog, Stains in Murder Case,* Portland Oregonian, Aug. 7, 1998, at D06, *available in* 1998 WL 4227728 (reporting expert testimony in a Washington murder case that "the likelihood of finding 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"). [BACK]

45. . *Bogan*, 905 P.2d at 520. The Arizona case law on this subject is criticized in D.H. Kaye, *Bible Reading: DNA Evidence in Arizona*, 28 Ariz. St. L.J. 1035 (1996). [BACK]

46. . For a rough analogy, consider the outcomes of rolling ten fair dice. Each outcome is like the genotype, and the numbers on the face the dice are like the alleles. [BACK]

47. . 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. [BACK]

48. . Few people would worry, for example, that the sample of blood cells taken from their veins 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 1 Modern Scientific Evidence, *supra* note 13, § 15-4.5.1; NRC II, *supra* note 1, at 126-27, 186. [BACK]

49. . The 1996 NRC Report, *supra* note 1, at 114, refers to "at least several hundred persons." (However, it has been suggested that relatively small databases, consisting of fifty or so individuals, allow statistically acceptable frequency estimation for the common alleles and that rare 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 Hum. Biology 141, 156-57 (1992)). Later, the NRC report 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. [BACK]

50. . *See* Bruce S. Weir, *Forensic Population Genetics and the NRC*, 52 Am. J. Hum. Genetics 437 (1993) (proposing interval estimate of genotype frequency); NRC II, *supra* note 1, at 148 ("calculation of confidence intervals is desirable"). For discussions of the statistical issues, see David J. Balding, *Estimating Products in Forensic Identification Using DNA Profiles*, 90 J. Am. Stat. Ass'n 839, 840 (1995); Ranajit Chakraborty et al., *Evaluation of Standard Error and Confidence Interval of Estimated Multilocus Genotype Probabilities, and Their Implications in DNA Forensics*, 52 Am. J. Hum. Genetics 60, 69 (1993); NRC II, *supra* note 1, at 146-47. [BACK]

51. . The genetic uniqueness of individuals in a large, sexually reproducing, outbreeding species is a central tenet of genetics, but as with humans, identical siblings are not genetically unique. Some species are prone to twin births, and this fact must be considered if the testing has been done on such a species. [BACK]

52. . For explanations of these assumptions and the corresponding equations, see, *e.g.*, NRC II, *supra* note 1. [BACK]

53. . *See*, *e.g.*, *id.* [BACK]

54. . *See, e.g., Bogan*, 905 P.2d at 523-24 (reporting that the biologist who testified for the prosecution consulted with botanists who assured him that Palo Verde trees were an outcrossing species). [BACK]

55. . However, large, pre-existing databases rarely may be available for the populations of interest in 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. *See supra* note 38 and accompanying text. 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. [BACK]

56. . A standard correction for population structure is to incorporate a population structure parameter F_{ST} (or) into the calculation. *See* NRC II, *supra* note 1. However, appropriate values for F_{ST} may not be known for unstudied species. [BACK]

57. . 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. [BACK]

58. . Bacteria can also exchange DNA through several mechanisms unrelated to cell division, including conjugation, transduction, and transformation. Bacterial species differ in the extent to which they undergo these forms of gene transfer. [BACK]

59. . A benefit to using such individuals in preference to narrow specialists is that specialists whose expertise is limited to a very narrow area are often unaware of aspects of an application outside their view. [BACK]

PUBLICATIONS HOME

last updated 11/24/98