Volume 3, Spring Issue, 1990

DNA FINGERPRINTING: THE *CASTRO* CASE

*Stephen M. Patton**

INTRODUCTION

In recent years there has been much discussion on the use of scientific evidence in the courtroom. Parties increasingly ask the courts to admit the testimony of research scientists, physicians, psychologists, and other technically-trained people. Paralleling this increased demand for the admission of scientific evidence is a growing awareness that current legal methods of reviewing and weighing such evidence are insufficient and should be reconsidered.

The increasing complexity and persuasive force of scientific evidence is exemplified by the use of "DNA fingerprinting." The possibility of identifying a human being by a tiny shred of tissue or drop of blood has a strong appeal for its potential to revolutionize rape, paternity, and murder cases.[1] However, courts must be careful to define the procedures used and insure that they are reliable.

This Recent Development examines one of the most significant cases involving DNA fingerprinting, *People v. Castro*.[2] *Castro* is one of the first cases in the relatively short history of DNA fingerprinting in which a court conducted an exhaustive evaluation of both the DNA procedure and the application of traditional admissibility rules.

I. THE ADMISSIBILITY OF NOVEL SCIENTIFIC EVIDENCE

The traditional common law standard of admissibility is the relevancy test. Any evidence must illuminate a fact at issue and must not be outweighed by its tendency to confuse the jury. *Frye v. United States*[3]

---

* Harvard Law School, Class of 1991. The author is indebted to P. L. Newman, LL.M. candidate, Harvard Law School.

1. "[DNA typing is] the single greatest advance in the 'search for truth' . . . since the advent of cross-examination." People v. Wesley, 140 Misc. 2d 306, 533 N.Y.S.2d 643, 656 (Albany County Ct. 1988). *See also* Thompson & Ford, *DNA Typing: Acceptance and Weight of the New Genetic Identification Tests*, 75 VA. L. REV. 45, 45 (1989).

2. 545 N.Y.S.2d 985 (Sup. Ct. 1989).

3. 293 F. 1013 (D.C. Cir. 1923). The "general acceptance" standard adopted in *Frye* has achieved something less than universal acceptance. As one commentator explains:

> "General scientific acceptance" is a proper condition on the court's taking judicial notice of scientific facts, but not a criterion for the admissibility of scientific evidence. Any relevant conclusions which are supported by a qualified expert witness should be received unless there are other reasons for

proposed an additional qualification on the admissibility of novel scientific evidence, which is that it must be generally accepted by the scientific community. New evidence, the court held, should be approved by the people who use it:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, *the thing* from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[4]

This standard is of limited utility without further definition.[5] What shall be accepted? If accepted, then for what purpose? Who shall accept it? How many experts constitute general acceptance and in what field should they specialize?

Of all the ambiguities, the most troubling is the primary definition: what did the court mean by "the thing"? Courts have interpreted this phrase to mean the "underlying principle," the "technique," the "procedure," the "technology," and the "scientific technique" among other permutations, but "it is doubtful that these preferences reflect a conscious choice with regard to this issue."[6]

DNA fingerprinting also suffers from a lack of precise definition. Although the legal literature describes a handful of common procedures used by geneticists, the ways in which these procedures are compiled

---

> exclusion. Particularly, probative value may be overborne by the familiar dangers of prejudicing or misleading the jury, and undue consumption of time.

C. MCCORMICK, HANDBOOK OF THE LAW OF EVIDENCE § 203, at 491 (1972) (footnote omitted). Although this Recent Development focuses on the introduction of DNA fingerprinting evidence under the *Frye* standard, somewhat different considerations may apply in states that have adopted the Federal Rules of Evidence approach for dealing with scientific evidence. C. MCCORMICK, EVIDENCE 554 (1987).

4. *Frye*, 293 F. at 1014. *See generally* Giannelli, Frye v. United States, *a Half-Century Later*, 80 COLUM. L. REV. 1197 (1980). *See also* P. GIANNELLI & E. IMWINKELRIED, SCIENTIFIC EVIDENCE § 1–5 (1986).

5. *See* United States v. Ziegler, 350 F. Supp. 685, 687 (D.D.C. 1972), *rev'd per curiam*, 475 F.2d 1280 (D.C. Cir. 1972) ("A preliminary task is to define the phrase 'general acceptance.' The cases following the *Frye* rationale have been carefully considered and they offer little guidance.").

6. Giannelli, *supra* note 4, at 1211–12.

and applied vary widely.[7] These variations in procedure can create variations in reliability.[8] Indeed, with all the writing on the use of DNA as identification evidence, even a common terminology eludes the experts.[9]

Definitional issues are difficult to resolve in the best of circumstances, but when DNA evidence, with its extreme probabilities, diverse terminology, and multiple procedures is evaluated, the ability of the *Frye* test to function effectively is questionable. The salient question is: How carefully does the *Castro* opinion define "the thing" that it decides to accept?

II. DNA "FINGERPRINTING"

DNA fingerprinting, a type of DNA forensic technology, is a technique used to identify persons by analyzing DNA[10] from their tissues. DNA fingerprinting is not a single process, but a collection of procedures for separating DNA from the cells in which it is found, slicing it up into various lengths, separating the resulting fragments by length, and finally identifying the resulting fragments by the use of radioactive "probes" which recognize specific sequences of nucleotides.[11] Because this analysis examines differences among sets of DNA fragments obtained by digestion with restriction enzymes, it is called restriction fragment

---

7. *See* Thompson & Ford, *supra* note 1, at 64–81; Burk, *DNA Fingerprinting: Possibilities and Pitfalls of a New Technique*, 28 JURIMETRICS J. 455, 458–63 (1988).

8. Thompson & Ford, *supra* note 1, at 57–59.

9. "DNA fingerprinting," "DNA typing," "DNA profiling," "DNA mini-satellite analysis," "DNA forensic identification tests," "DNA identification tests," "DNA RFLP analysis" and "DNA tests" are all terms that have been used in recent commentaries and cases. Reliance on names that are used in legal literature should be limited because the same name is sometimes used to describe different tests and different names are used to describe the same test. Moreover, even tests performed by the same company may be different at different times. As Thompson & Ford note, "[w]hen Cellmark first opened . . . the company relied exclusively on 'multi-locus' probes. . . . In early 1988 the company abandoned the use of multi-locus probes for criminal identification in favor of single-locus probes similar to those used by Lifecodes, though Cellmark still uses multi-locus probes in paternity cases." Thompson & Ford, *supra* note 1, at 49.

10. "DNA molecules contain the genetic information that makes every person unique. The structure of DNA resembles a twisted ladder or 'double helix,' with the side rails composed of long chains of sugar molecules. The rungs of the ladder consist of pairs of 'nucleotides' or 'bases.' There are only four possible bases and the sequence of these bases defines the genetic information contained in the DNA." *Castro*, 545 N.Y.S.2d at 988.

11. Two companies currently employing this method of analysis for DNA identification are Lifecodes Corporation of Valhalla, NY, and Cellmark Diagnostics of Germantown, MD. Cetus Corporation of Emeryville, CA uses similar procedures for digesting the DNA, but does not separate resulting fragments by size. Their process is termed "allele-specific probe analysis." *See* Note, *The Dark Side of DNA Profiling: Unreliable Scientific Evidence Meets the Criminal Defendant*, 42 STAN. L. REV. 465, 471 (1990).

length polymorphism ("RFLP") analysis.[12]

DNA identification[13] is a multi-stage process. First, DNA is extracted from the rest of the cellular material. The DNA is then digested by a molecule called a "restriction enzyme" which slices the DNA at specific points identified by particular base sequences. This creates a set of DNA fragments of varying size, each fragment having an identical base series at its ends.[14]

The next step is gel electrophoresis. The short fragments of DNA are placed in small wells along one end of a slab of agarose gel and an electrical field is applied across the gel. DNA has an electrical charge so the electrical field causes the fragments to move through the gel, from one end to the other. The small fragments move more quickly than the longer fragments. The electrical field is shut off after a certain amount of time, so the DNA fragments will have migrated to resting positions based on their size.

The court in *Castro* includes in its description an extra step that Lifecodes performs which is valuable as a control on the process. In this step the entire gel is washed with ethidium bromide which bonds to the DNA, thus making the smears of DNA fragments visible so the scientist can be assured there is enough DNA across the entire gel to make a reliable comparison.[15]

The DNA fragments are then blotted onto a nitrocellulose filter at exactly the same position as they are in the gel, in a process known as "Southern Blotting." The pattern of the DNA fragment lengths will only be of value if the pattern can be made visible. This is achieved by binding molecules called "probes" to the fragments. Probes are actually short lengths of single-stranded DNA (one half of the double helix) with a radioactive "label" attached. The blotter is washed with a solution of many probe molecules and the probes bond or "hybridize" with

---

12. The fragments produced by digestion of the DNA are "restriction fragments," and variations in size ("polymorphisms") of these restriction fragments are identified by the probes. These polymorphisms are rare and therefore distinctive. "Of the 3,000 million nucleotides we inherit from each parent, about 1 in 1,000 is a site of variation, or polymorphism, in the population." Lander, *DNA Fingerprinting on Trial*, 339 NATURE 501, 501 (1989). The restriction fragments valuable for identification purposes are those that vary from person to person, rather than the ones that are common to all people.

13. This Recent Development considers primarily RFLP analysis, since it was the process used by Lifecodes in the *Castro* case. Comments about statistical data will apply to all DNA identification tests.

14. Although every individual's DNA changes gradually throughout life, these changes are too minor to be detected by RFLP analysis. Accordingly, the RFLP analysis yields identical results on every sample taken from an individual, provided the experimental conditions are the same.

15. 545 N.Y.S.2d at 991.

fragments on the blotter that have a base sequence complementary to the pattern of bases of the probe. All excess probe material is then washed away. Since only some of the DNA fragments will have a base sequence that complements a given probe, the probe will bond only at certain locations on the blotter. To see the places where the probe has bonded, a piece of photographic film is placed in contact with the blotter and is exposed to the radioactivity from the probe molecules. This film is then developed and black spots or bars appear wherever a probe bonded to a DNA fragment. These black areas that show the presence of "alleles," or variations in a gene, are known as "bands." The photographic film itself is known as an "autoradiograph" or "autorad."[16]

Each band reveals three bits of information about the fragment: first, the base sequence at the end points of the DNA fragments complements the sequence of the restriction enzyme; second, the fragment contains at least one sequence of bases that matched the sequence of the probe; and third, the approximate length of the DNA fragment.[17] All the other bases on the DNA fragment between the fragment endpoints and the relatively short sequences that matched the probe are unknown.

Two samples taken from the same person will have identical patterns of bands,[18] but the fact that two samples match does not mean that they came from the same person, since there may be many people who have similar bands. As an extreme example, there are probes that identify "non-polymorphic"[19] regions of human DNA that are common to all people and hence produce bands that do not vary from person to person.

---

16. In polymorphic regions of DNA the genes vary in length. When DNA of several individuals is digested and run through gels, these variations in genes, called "alleles," appear as a family of bands in similar locations on the autorads. If the gene does not vary greatly from person to person, there will only be a few possible bands. *Castro*, 545 N.Y.S.2d at 991.

17. Fragment length may be determined from the position of the band on the gel. The actual length of the fragments may be calculated by their position on the gel, the type and consistency of the gel itself, the strength of the electrical field applied, and the amount of time the field was applied to move the DNA fragments. *Id.* Alternatively, the lengths of the fragments may be determined by comparison to the positions of markers of known length that run simultaneously on a different portion of the gel. The DNA fragments used for identification are typically 2,000 to 40,000 bases long.

18. Since the agarose gel may vary in quality from one gel slab to another, identical fragments may move at different speeds in different gels and hence may be deposited at slightly different locations. Thompson and Ford, *supra* note 1, at 70 n.118. *See also supra* note 8 and accompanying text.

19. These probes bond to DNA fragments that do not vary in length from person to person; hence, such probes are of little value in making identifications. However they may be of value in determining the quality of the DNA. If a common band is not produced using a non-polymorphic probe, it implies that the DNA was degraded. *Castro*, 545 N.Y.S.2d at 994.

For the evidence to have any incriminating value, statistics specifying the rate of occurrence of these bands in the general population must be consulted. Such statistics may be produced by making autorads of many different people. The rate of occurrence for each band is estimated, and used to calculate the odds of a random occurrence for each band. If two samples have a single matching band for an allele that occurs in only one out of a hundred people, typical testimony would be that there is a one-in-a-hundred chance that both the suspect and the perpetrator have that band in common. The more bands that appear in both samples, the smaller the odds of a coincidental match. Of course, in the absence of error, the lack of a match for any band does indicate that two samples came from different people. Thus, DNA fingerprinting has unquestionable exculpatory power, but caution must be employed when using it as an inculpatory device.

III. THE *CASTRO* DECISION

On February 5, 1987, Vilma Ponce and her two-year-old daughter were stabbed to death in the Bronx.[20] Police questioned José Castro, a handyman in the neighborhood, and noticed a bloodstain on his watch. Samples taken from the deceased victims and the watch were sent to Lifecodes for analysis. DNA fingerprinting analysis was performed, and a match between the two samples was reported which mentioned no difficulties or ambiguities in the processing or analysis.[21] The prosecution attempted to have the results of this testing admitted as evidence.

Over a twelve-week period in the spring of 1989, a "*Frye-Middleton*"[22] hearing was held before Judge Scheindlin of the Superior Court of Bronx County, New York, with expert witnesses appearing for both the defense and prosecution.[23] On August 14, 1989, the court delivered its decision, finding in general that "DNA forensic identification tests" are acceptable for both inculpatory and exculpatory purposes.[24] However, the court found that in this specific case the testing laboratory had not applied approved procedures; hence, evidence of guilt was not admitted.

---

20. 545 N.Y.S.2d at 986.

21. *See* Lander, *supra* note 12, at 501–02.

22. New York State uses a modified *Frye* standard known as the *Frye-Middleton* rule. This standard requires novel scientific evidence to be "generally accepted as reliable" to be admitted. People v. Middleton, 54 N.Y.2d 42, 49 (1981). The *Frye-Middleton* and the *Frye* tests will be referred to collectively as "the *Frye* test."

23. 545 N.Y.S.2d at 985.

24. *Id.* at 995.

The court recognized the difficulty of applying the *Frye* test to complex evidence and advanced a three-part refinement of the *Frye* test. A court should address the following questions in ruling on the admissibility of DNA fingerprinting:

> *Prong I.* Is there a theory which is generally accepted in the scientific community, which supports the conclusion that DNA forensic testing can produce reliable results?
> *Prong II.* Are there techniques or experiments that currently exist that are capable of producing reliable results in DNA identification and which are generally accepted in the scientific community?
> *Prong III.* Did the testing laboratory perform the accepted scientific techniques in analyzing the forensic samples in this particular case?[25]

In its Prong I analysis the court reviewed the theory of DNA analysis in general terms. The court explained that DNA is unique to the individual, and that although it is not possible to disassemble a person's DNA and hence uniquely identify the individual, it is possible to make quantitative estimates of relative uniqueness by fragmenting the DNA and comparing the fragments obtained to a statistical database of DNA fragment probabilities. The court deemed this theoretical proposition to be generally accepted as reliable, and hence admissible under the *Frye* rule.

In Prong II, the court reviewed the RFLP analysis technique in general. This technique is commonly used, and the court accepted it as "generally accepted as reliable." The court concluded the Prong II analysis by stating that "DNA forensic identification tests to determine inclusions are reliable and meet the *Frye* standard of admissibility," and that "DNA forensic identification tests to determine exclusions are reliable and meet the *Frye* standard of admissibility."[26]

In Prong III, the court looked at the reliability of the specific tests performed by Lifecodes. The court found that Lifecodes did not follow accepted scientific procedures because it failed to perform certain experiments, techniques, and controls necessary to produce reliable results.[27]

---

25. *Id.* at 987. Presumably, future courts in this jurisdiction will not have to perform the Prong I and II investigation, but will merely address the Prong III issues.

26. *Id.* at 995.

27. *Id.* at 996–98. Many of the problems that arise in DNA fingerprinting are due to the lack of control over the sample. A sample recovered from a crime scene may contain bacteria that have grown in the sample, and thus display both bacterial and human DNA bands in the autorad. Samples that have dried, aged, been in contact with chemicals, or have simply decayed may be broken at random points, thus causing higher molecular weight bands to disappear. Finally, samples recovered from a crime scene may have concentrations of

Thus, while DNA fingerprinting in general was accepted as admissible by the court, the specific DNA fingerprinting analysis performed by Lifecodes was not. The court suggested that every case involving DNA fingerprinting evidence should include a Prong III hearing to insure that the specific techniques employed measure up to the *Frye* standard. Although Prong III relates more to the weight of the evidence than to traditional concerns over admissibility, the court suggested that the Prong III analysis should also be carried out in a pre-trial hearing.

IV. COMMENDATIONS AND CRITICISMS

A. General Acceptance and Reliability of Data Interpretation

The *Castro* case is significant as an application of the *Frye* standard to evidence produced by a complex scientific technique. The *Frye* rule, calling for evidence to be "generally accepted as reliable," before it can be admitted, appears to stress both the reliability and the general acceptability of the evidence by combining the two into a single test. In practice, however, this combination obscures the separable question of reliability.

Reliability does not exist in a vacuum but implies a fitness for a particular purpose. A test may be reliable for one particular set of circumstances, yet not reliable for another. In *Castro* the court stated that the procedures were generally accepted, but were unreliable as applied in *Castro*.

An inquiry into a DNA test's "general acceptance as reliable" should direct us to the particular purpose that the test is to serve. In this case, Lifecodes compared DNA samples to see whether they had the same bands. Since there was a close match, Lifecodes further analyzed the band patterns using population statistics to ensure that the matches were sufficiently significant to constitute evidence of criminal guilt. To convict, it is necessary to show, using population statistics for the various match loci, that two random samples are unlikely to demonstrate the same match.

This situational reliability review requires a deeper inspection of the method of data interpretation than was carried out. In *Castro*, the court described the generic procedures for interpreting autorads and found them acceptable. But this analysis alone is not enough; the court should have also evaluated the matching rule implicit in Lifecodes' testimony

---

DNA so low that some bands will appear quite faint on the autorad and be ignored.

that the bands matched. Such testimony implies some standard of matching, some rule of band comparison. The standard of comparison is part of "the thing from which the deduction is made" and should also be evaluated for general acceptance just like the procedures involved in producing the physical evidence.[28] Instead, the court merely stated that the matching rules should be consistently applied within "a permissible margin of error."[29]

B. Probes as Individual Tests

The court's approval of a "DNA forensic identification test" in general marks a significant departure from previous applications of the *Frye* test. Previously, courts considered each comparison of a specific chemical compound to be a separate technique requiring separate evaluation.

The court reviewed evidence presented by Lifecodes in detail, and identified problems specific to both the general procedures and the particular probes used by Lifecodes. One probe-specific problem identified by the court is the significant risk of environmental degradation of high molecular weight bands present in some probes.[30] If not accounted for, such degradation can lead to false positives and false negatives.

A second problem recognized by the court is the possibility of misreading autorads when the relative allele intensities for a particular probe vary widely.[31] Allele intensity is a function of the particular probe,

---

28. *See Frye*, 293 F. at 1013. Indeed, in *Frye* itself the issue was not questionable physical evidence or its method of production, but the *interpretation* of that evidence. The evidence challenged in *Frye* was an early analogue of the polygraph. It measured systolic blood pressure while the suspect was asked probing questions. The medical profession generally accepted systolic blood pressure measurement as reliable at that time, and accepted that blood pressure changed under stress. Nevertheless, the court found there was no general acceptance of interpreting certain changes in blood pressure as evidence of lying.

29. 545 N.Y.S.2d at 992.

30. *Id.* at 996. For example, in *Castro* a perfect match was found using probe D2S44 for a single band at 10.25 Kilobases (Kb). Defense experts noted that 90 percent of the bands revealed when using this probe on samples from Hispanics are of higher molecular weight than 10.25 Kb. The existence of no bands larger than 10.25 Kb. suggested that the DNA in the *Castro* samples had degraded and the higher molecular weight bands had disappeared. Prosecution expert Michael Baird, Lifecodes' director of paternity and forensics, rationalized that the ethidium bromide stain "indicated that enough material was present to get a signal in the 12–15 Kb range," if there were any bands there, though he could not be certain. Lander, *supra* note 12, at 503.

31. 545 N.Y.S.2d at 994. The more DNA in a sample, the darker the bands on the autorad will appear. Whenever the bands are of the same intensity, they fade equally as the amount of DNA is reduced. If a probe produces bands of quite different intensities, however, normally dark bands may appear faint and normally faint bands may disappear entirely as the amount of DNA decreases. This points again to reliability as a situational notion. Genetic researchers never face this problem because they draw fresh blood samples having a quantity of DNA that can be readily calculated. Forensic analysts work with

hence some probes and some alleles may be inappropriate for use in DNA fingerprinting. Another problem in the *Castro* case is that in analyzing the autorad of probe DXYS14, Lifecodes ignored three bands, discounting them as contaminants "of non-human origin." Defense experts disagreed with this conclusion.[32] The court, sensitive to subjective discounting of exculpatory evidence, suggested a preference for synthetic probes over bacterial probes.[33] This suggestion implies that certain probes are inappropriate for DNA fingerprinting.

It is commendable that the court identified some of the problems and ambiguities with the probes used by Lifecodes. This searching analysis of the evidence is critical for such complex testimony, and highlights the need for a pool of experts to challenge the evidence in an adversarial arena.

As the court made clear, flaws are dependent upon the probe and band used. To be challenged or accepted effectively, the particular probes used and alleles compared must be familiar to the experts who are testifying in the case. Unfortunately, there are hundreds of probes, thousands of RFLPs, and tens of thousands of alleles that may be used in "DNA identification testing," and more are being discovered every day. By June 1989, over 3000 different RFLPs had been identified. Approximately 100 of these loci may have as many as 50 to 100 alleles, many of which are quite similar and hence difficult to differentiate.[34] Of these many thousands of possible RFLPs, each scientist uses only a small number.[35] This points to the clear danger in a sweeping acceptance of "DNA forensic identification tests." A more prudent approach would have been to limit approval to the particular restriction enzymes, probes, and bands used for identification in this particular case.

The treatment of DNA evidence in *Castro* exemplifies a weakness not so much in the court, but in the *Frye* test itself. Although requiring

---

degraded, dried, and contaminated samples of unknown concentration and therefore may not know the quantity of DNA.

32. Lander, *supra* note 12, at 502.

33. 545 N.Y.S.2d at 994. Probes obtained from genetically engineered bacteria may hybridize with bacterial contaminants in a sample due to bacterial contamination in the probes, whereas probes synthetically manufactured will not. This, too, points to situational differences in reliability. Genetic diagnosticians work in a sterile environment with sterile equipment, drawing sterile samples. In many forensic situations, bacteria cannot be avoided since samples must be taken as found. In the present case, DNA was extracted from dried blood on a watch.

34. Lander, *supra* note 12, at 501.

35. *See id.* If the defense experts had not been familiar with the allele pattern of probe D2S44, they could not have effectively challenged the decision of Lifecodes to ignore non-matching bands. *Id.* at 503. It is difficult to imagine what "general acceptance" means in a field where most scientists are unfamiliar with most of the information available to them.

"general acceptance," the *Frye* decision gives no guidance about what exactly should be accepted or how deep the inquiry should be. The result in *Castro* was a sweeping acceptance in principle of a handful of procedures vaguely described as "DNA forensic identification tests." The court's broad holding neglected much of the detail in the test. Scientific evidence should be examined in greater detail and accepted in specific, not general terms. Furthermore, the results of the court's admissibility analysis should be reported more precisely. To analyze evidence in more detail, a general framework for review is needed.

V. SCIENTIFIC IDENTIFICATION EVIDENCE

Courts need to concentrate their efforts on standardizing and clarifying the use of novel scientific evidence. They should determine what the essential elements of each scientific identification test are, and describe each element in detail for use in later cases. I propose a four-step approach to such complex scientific identification evidence:

(A) Identify the actual characteristic that is claimed to match in two samples. This characteristic should be generally accepted as measurable to the precision claimed.
(B) Evaluate the specific methods used to produce the physical evidence, if any, and determine whether they are generally accepted as capable of measuring the characteristic to the precision claimed. These methods should be included in the record either explicitly or by reference.
(C) Assess the rule used to declare a match between two samples and determine the degree of certainty of the match. The matching rule should be similarly stated explicitly.
(D) Analyze the probabilistic basis for determining the significance of a match. If the samples are used for inculpatory purposes, the court should determine whether the probabilities are intuitive to the jury, or intuitive to the experts, or whether they must be derived from population statistics gathered scientifically. If the statistical assessment is necessary, the court should determine the general acceptance of the database, and include that database either directly or by reference in the record.

These four steps are elaborated below.

A. The Characteristic

1. Definition of the Characteristic

The court should state precisely which characteristics are identified in both the suspect sample and the perpetrator sample. For example, characteristics should not be understood as simply "fingerprints" or "DNA" since they are not the actual subject matter of the comparisons between the specimens. In fingerprinting, the "points of identity," such as loops, whorls, deltas, ridge endings, bifurcations, and dots, are the actual characteristics that are measured and compared.[36] Similarly, in DNA testing, the characteristics are the existence, size, and intensity of a band produced using a particular probe and specific restriction enzyme. The importance of clearly defining the relevant characteristics was amply illustrated by *Castro*. The practical utility of the decision would have been much greater if the court had defined the test it accepted as a process measuring and comparing particular bands, produced by specific probes and restriction enzymes, by a specific matching rule.

2. Mutability of the Characteristic

The identification of a common characteristic in two samples will not constitute a match where the characteristic can change unpredictably.[37] If a scientific test or procedure depends on external conditions, there is no legitimate basis for relying on the tests unless the external conditions are known and are determined not to affect the validity of the results. Therefore, once the characteristics to be measured are identified, the court should ask whether each characteristic is immutable *in the particular situation*. Unfortunately, samples gathered in forensic situations often may have undergone significant change.

---

36. *See generally* P. GIANNELLI & E. IMWINKELRIED, *supra* note 4, §§ 16–6, 16–7. Fingerprint evidence rarely compares two full sets of prints that can be superimposed for a perfect match. Instead, they identify particular characteristics and points of similarity in two samples for comparative purposes. For example, a measurement or comparison of the distances between lines would be misleading since the finger's skin might be stretched when making the print. Therefore, testimony that the prints did not match because the absolute distances varied would be based upon an inadequate characteristic for comparison. This is a trivial case currently, since there are a wealth of fingerprint experts, international standards, and a well established system of comparison. In the case of a new technology such as DNA fingerprinting, legitimate bases of comparison are ill-defined.

37. A lack of understanding of the characteristic measured can trivialize the discourse in a *Frye* hearing. DNA and fingerprints may not change, but the characteristics actually measured and compared may indeed change. For this reason, fingerprint experts consider relative orientation but not absolute distances as a relevant characteristic. *Id.* at § 16–7(A).

B. Precision, Accuracy, and Method of the Characteristic's Measurement

Although scientists may accept that a characteristic can be measured, the precision and accuracy of a measurement will vary depending on how the measurement was made. Many scientists may use autorads of a person's DNA to diagnose genetic diseases; however, it is unclear whether they make their measurements with the same degree of accuracy as should be required of forensic researchers seeking inculpatory evidence. If scientific evidence depends on a highly precise measurement, the ability to make a measurement with the necessary precision must be generally accepted.[38]

C. The Matching Rule

Whether the evidence takes the form of fingerprints, bitemarks, blood typing, or DNA fingerprinting, test results from two samples are compared to see if the two samples match.[39] A match requires the identification of similar characteristics in both samples and an absence of dissimilar ones. For each characteristic compared there must be an explicit "matching rule." In some instances such as ABO blood typing, the matching rule is simple: Do the blood cells agglutinate?[40] The case of DNA fingerprinting is more problematic. A band may appear both in the suspect's and perpetrator's sample. If these bands differ significantly

---

38. A genetic researcher testifying for the defense in *Castro* noted the difficulty with modern measurement methods in reliably differentiating between alleles:

> Because hypervariable RFLP loci often involve 50–100 alleles yielding restriction fragments of very similar lengths, reliably recognizing a match is technically demanding. At one commonly used locus, for example, most alleles lie within a mere 2 per cent of the length of the gel.

Lander, *supra* note 12, at 501.

39. Matching rules are not always objective. For example, toolmark evidence is often very subjective. The only way to challenge a subjective expert judgment is to challenge the experience or credentials of the expert herself. Such evidence suggests its own problems, which relate in large part to the weight of the evidence rather than its admissibility. In any event, a discussion of this type of evidence is beyond the scope of this Recent Development.

40. However, even with ABO blood typing the definition of agglutination (clumping) of the blood cells should be challenged. How clumped must the cells be before we say they are agglutinated? Is there a range of sizes where agglutination and hence blood type is questionable? Here at least we can rely on medical science for the answer. Since ABO blood typing is used widely, and since the effect of mis-typing is apparent (agglutination of transfused blood cells within a recipient's body and possible death), medical researchers have developed reliable rules to guide the legal community.

in their molecular weights, the samples do not match and hence come from different people. Certainly if they are of identical weight, they can be said to match. But at what point along the broad spectrum of measurement should the experts decide whether the samples match or not? In other words, how big is the "grey area" in which it cannot be said with confidence that the samples are either the same or different? Accordingly, it is crucial that the matching rule should also be generally accepted.

D. Population Statistics

For a scientific identification test to have incriminating power it must not only identify a matching characteristic, but it must also be supported by some rational inference that the match could not have occurred randomly. An assessment should be made of the probability of coincidental occurrence in the relevant suspect population.[41] Without information on the random occurrence of the characteristic, it is impossible to know the significance of a match. Perhaps everyone has the characteristic, in which case even if a match was declared, it would have no probative value. The absence of any generally accepted probabilistic information should not be considered as going to the weight of the evidence, since without some probability basis, the evidence has no incriminating weight whatsoever, and it should be inadmissible. For scientific identification evidence to have *exonerating* power, no probabilities are necessary. If two samples exhibit non-matching characteristics, they come from two different sources.

Of course, some identification evidence has probative force even in the absence of purely statistical data. For example, a blood test may indicate that a perpetrator was a male. The jury can easily calculate the odds of a coincidental match (a random occurrence) as approximately

41. Determining the relevant population is difficult. After identifying matching bands in the *Castro* case, Lifecodes calculated the possibility of those bands appearing randomly, using statistics for the occurrence of those bands in the Hispanic population. Lander, *supra* note 12, at 502. The unspoken assumption is that the blood on the watch was of Hispanic origin, since the neighborhood was Hispanic. But this assumption may not have been correct. If the alleles identified in the watch's bloodstain were found in all Caucasian and African-American New Yorkers, the alleles would be quite common for New York as a whole even if they were uncommon in the Hispanic sub-population. A large percentage of all New Yorkers would have that pattern of alleles and the chance of finding that pattern at random in the New York metropolitan area would be very high. The evidence of a match would have lost its inculpatory power. While this discussion is only hypothetical, it demonstrates the possible danger in the selection of the relevant population.

fifty percent in the general population.[42] Alternatively, a court may allow the intuitive probabilities of an expert to be admitted.[43] Finally, the court may allow the expert to present numeric, statistical evidence to give the match weight. It is important for the court to address probability inferences explicitly, since without probability information, the evidence is not probative.

Secret databases of population statistics used to evaluate DNA fingerprinting tests present a very real danger. The statistical databases used by Lifecodes and Cellmark for the allele occurrence frequencies were generated independently by each company. They create and analyze autorads in-house to identify alleles and compile the statistics. They consider the databases proprietary. At trial, the companies routinely refuse access to the autorads used to produce their statistics, preventing the defense from challenging them.[44] Certainly, researchers from these

---

42. Bitemark evidence comparing the defendant's teeth and bitemarks found on the victim was admitted in People v. Marx, 54 Cal. App. 3d 100, 112 (1975). The court questioned comparisons made by experts in matching the samples, but admitted them anyway because the jury could evaluate the evidence themselves. *Id.* at 112. Rather than a failure of the *Frye* test as Professor Starrs implies, this is a case where the judge believed the jury could supply its own matching rule and its own intuitive population statistics. Starrs, Frye v. United States *Restructured and Revitalized: A Proposal to Amend Federal Rule of Evidence 702*, 115 F.R.D. 79, 94 (1987).

43. In Delaware v. Pennell, 1989 LEXIS 520 (Del. Super. Ct.), the court of first instance admitted testimony on the DNA evidence, but disallowed testimony on the statistics. At trial, the statistics were inadvertently mentioned and the case was appealed. The appellate court did not reverse or remand. This points to the particular problem discussed here: A mere match has no probative value. A non-polymorphic probe can identify an allele all humans have in common. Allowing the jury to supply their own "intuitional" statistics makes no sense since a jury cannot have developed rational intuitions concerning such novel evidence. This problem is exacerbated by the media, which often touts DNA evidence as providing positive identification.

44. How the statistics can be admitted when scientists and courts are denied access is puzzling.

> In response to a discovery motion, Cellmark disclosed to the defense its "DNA fingerprinting" protocol, laboratory notes from the testing in this case, the autoradiographs produced during RFLP analysis and statistical frequency tables. The defense request for more specific information regarding its methodology and population data base was denied by Cellmark. Arguably, trade secrets may be at stake for the commercial laboratories. Protective measures could be pursued, however, before denial of discovery is appropriate.

State v. Schwartz, 447 N.W.2d 422, 427 (Minn. 1989).

One commentator notes:

> Because the . . . private laboratories conceal much of their process from the public view with the shield of "trade secrets," their techniques cannot gain general acceptance in the scientific community. The scientists from these labs have not publicly revealed much of their protocols through publishing. On the individual case level, some defendants have not been allowed access to labora-

companies have published many peer-reviewed articles that describe the summary statistics derived from these autorads, but the articles of necessity do not include the autorads themselves.

The *Castro* court cast doubt on Lifecodes' ability to interpret autorads reliably. Since their probability statistics were compiled from proprietary autorads interpreted in-house, presumably by the same methods that were suspect in *Castro*,[45] all the statistical factors compiled from them and used to calculate probabilities in future cases may be products of similar misinterpretations.

CONCLUSION

In a *Frye* hearing on scientific identification evidence, the court will uncover many problems if it looks beyond the surface of the tests. For a test to be generally accepted, each specific element of the test must be accepted. The court should identify the traits to be compared, the means of measuring them, the rule used to identify a match, and the statistical studies (or accepted intuitional theories) used to show the significance of the match. The court should ask whether *all* elements are generally accepted, for if any of these links are missing, the chain of inference is broken and the evidence is useless.

Some commentators have recommended that a national "science court" be formed to approve new technologies before they are admitted in courts of law. In the distant future this may occur, but at present courts must rely on themselves to provide accurate analysis before accepting scientific evidence. Consequently, all tests should be described explicitly with reference to specific characteristics, matching rules, procedures, and statistical databases.[46]

The secrecy in DNA testimony that results from the use of proprietary databases is alien to our notions of justice. Such secrecy prevents a *Frye* decision from having any precedential value because subsequent courts will have inadequate access to the specific details involved in the previous decision. Thus, each new *Frye* proceeding will have to start from scratch. Although the *Castro* court clearly anticipated that future

---

> tory documents. If during trial opposing experts have been allowed to review the protocols, the experts must sign agreements not to disclose them. Therefore, independent validation has not been possible. The private labs have restricted access to their probes. . . . Clearly, scientific acceptance cannot occur with proprietary rights restricting access to the technology.

See Note, *supra* note 11, at 502 (footnotes omitted).

45. For a discussion of inculpatory interpretation, see 545 N.Y.S.2d at 996–98.

46. *See generally* Note, *supra* note 11.

hearings would only engage in a Prong III analysis,[47] the failure of the court to describe specifically an accepted methodology will require future courts to engage in detailed analysis. The process undertaken by the court should be recounted in the opinion and the test defined thereby.

Moreover, an explicit analysis by the trial judge conducting a *Frye* proceeding would provide an appellate court with significantly better information on which to base its review of the case. Poor or inappropriate tests may or may not suffice to overturn a decision in a particular case depending on the sufficiency of other evidence. However, lives and liberty will often depend on DNA evidence, and the ability to review that evidence is critical.

If the elements of the test were noted in the judgment, the probability of deliberate misrepresentations by scientists would be reduced because they would be more easily detected. Given that the resources of criminal defense attorneys are slim, the ability to get to the heart of future claims by experts about specific tests would increase the probability of mendacious expert testimony being challenged. If claims of accuracy or reliability were referred to explicitly in the opinion, their enhanced profile might reduce the tendency of scientific promoters to make such claims.

If a test is refined and changed after its initial acceptance, there is a danger that the changes will not be reviewed in later applications. Courts should take care not to be swayed by a company's assertions that its test is still a "DNA forensic identification test." By providing a complete definition of the procedure accepted, the court's attention will be focused on the technology of the test and not its title. Careful definition of the test will highlight the portions that have changed, thus increasing the efficacy of a *Frye* review. If past courts have accepted a theory of RFLP identification and the ability of a particular process to identify a match according to a known rule, the addition of another measurable characteristic (such as another allele) would not necessitate a review of the entire process, but would spotlight the particular trait such as the new allele that needs to be evaluated *de novo*. The court could limit its review to the methods used to identify this particular allele, the population statistics associated with it, and any dangers, such as degradation of the allele, that are unique to it.

Science has progressed dramatically since *Frye* and scientific evidence has become more complex. Statistics have assumed a position of great importance. To deal with these developments, a more methodical approach to reviewing scientific evidence in *Frye* hearings is necessary. By defining tests more accurately and addressing more directly the

47. *See supra* notes 25–27 and accompanying text.

"thing from which the deduction is made," courts will assure they are engaging in a meaningful exercise.