The forensic debut of the NRC's DNA report: population structure, ceiling frequencies and the need for numbers

D.H. Kaye
College of Law, Arizona State University, Box 877906, Tempe AZ 85287-7906, USA

Received 16 July 1994 Accepted 10 August 1994

Key words: forensic DNA, ceiling frequencies, NRC report, courts

Abstract

This paper reviews judicial opinions that have discussed the April 1992 recommendations of a committee of the U.S. National Research Council concerning the statistics of forensic DNA profiles obtained with single-locus VNTR probes. It observes that a few courts have held ‘ceiling frequencies’ (as opposed to less ‘conservative’ estimates) admissible, but that the implications of the scientific criticisms of the ceiling procedures have yet to be addressed adequately in court opinions. It urges courts to distinguish between policy judgments and scientific assessments in both the NRC report and the scientific literature, and to defer less to the former than to the latter.

Introduction

In deciding whether the results of DNA tests are admissible in criminal cases, courts in the United States have relied on the testimony of scientists about the state of the science, on the scientific literature, and on other judicial decisions and opinions. These inputs – and the resulting caselaw – have not been static. Since 1988, a fusillade of objections to the admission of DNA tests has been raised in court. These include questions about the possible effects of contaminants on forensic samples, the use of radioactive isotopes in exposing X-ray film, the addition of ethidium bromide, ad hoc corrections for band shifting, the records of laboratories on proficiency tests, the criteria and procedures for deciding whether two VNTR fragments ‘match’, the size and sources of databases used to assess the significance of matching bands, and the method of calculating the frequency of matching DNA patterns within a reference population. Three phases in the judicial responses to DNA evidence can be discerned: an initial period of enthusiastic acceptance of the evidence, a damping counterreaction, and the current period of mixed outcomes and opinions expressing grave concern over certain aspects of DNA evidence. (For reviews of these legal developments, see Kaye, 1993; Thompson, 1993).

In April 1992, a committee of the National Research Council (NRC, 1992) entered this free fire zone. This paper examines the courts’ use of the NRC report in their opinions. I confine the discussion to one issue – the statistics of DNA ‘matches’ with single-locus VNTR probes – for the committee’s views on this topic have been the most controversial. I do not purport to show that the report has caused the courts to reach any particular results; it is possible, and often probable, that the outcomes of cases in which the report is cited would have been the same had no report been written. Nor do I assess the contribution the report has made to clarifying the issues, resolving doubts about other matters, and prompting reforms in forensic work. I merely seek to describe how one major part of the report has been treated in opinions and to offer some suggestions to avoid its being misunderstood or misused.

Proving the existence of a controversy

In some jurisdictions, the lack of general acceptance within the scientific community precludes admission of scientific evidence. In others, it weighs against admissibility but is not invariably fatal (see, for example, McCormick, 1992, § 203). Pretrial hearings lasting
months and generating thousands of pages of testimony have addressed the question of scientific consensus, and courts frequently cite the NRC report to support a finding that a major scientific controversy is raging over the proper method for ascertaining the frequency of a 'match'.

Prior to the NRC report, the standard practice was to estimate multi-locus frequencies by multiplying single-locus frequencies, which in turn, were obtained by multiplying the two 'allele' frequencies at each locus. The allele frequencies were derived from histograms of fragment weights within a database. Under the protocol developed by the Federal Bureau of Investigation, the histograms involved very large class intervals to ensure that the single-locus frequencies would not be underestimated. The similarly sized fragments falling into each large 'bin', which are analogous to an allele, may be called a 'binelle'; the set of binelles, which is analogous to a genotype, may be called a 'binotype' (compare Devlin, Risch & Roeder 1992).

The computations, in other words, presupposed Hardy-Weinberg equilibrium and linkage equilibrium: if the estimated binelle frequencies of a heterozygote at the jth locus are \(p_j\) and \(q_j\), then the estimated binotype frequency \(P\) for a match at \(n\) loci is

\[ P = 2^n \prod_{j=1}^{n} p_j q_j \]

In court, the most potent criticism of this simple calculation has been the possibility of population structure - the presence of subgroups, with varying binelle frequencies, that tend to mate among themselves. One exposition of this criticism often cited in judicial opinions is Lewontin and Hartl (1991). The critics argue that, until direct studies of subpopulations are completed, there is no way to be certain that the departures from equilibrium do not make the standard frequency estimates too small (or for that matter, too large) by several orders of magnitude. Other population geneticists and statisticians maintain that direct studies of subpopulations are unnecessary and unlikely to be productive (see, for example, Devlin, Risch & Roeder, 1994; Roeder, 1994).

This debate is not easy for the courts to penetrate. In People v. Pizarro, 12 Cal. Rptr. 436, 456 (Ct. App. 1992), for instance, a California court of appeals quoted at length from early scientific publications, and lamented:

The difficulty is, where does this place us? It places us in the middle of the conflict as to whether or not the basic theory of population genetics involving broad racial and ethnic groups as opposed to the argument of substructure has any general acceptance in the relevant scientific community - a conflict which we cannot resolve on the present record.

The NRC report can only have reinforced the perception of dissension. If there was any doubt left in early 1992 that prominent scientists were divided over the adequacy of the forensic computations, the NRC committee put it to rest. Starting with People v. Barney, 10 Cal. Rptr. 2d 731, 741 (Ct. App. 1992), court after court has noted the committee's report of 'considerable dispute' and a 'substantial controversy'. A New Mexico appellate court was especially impressed:

The report discusses the debate over the need for subpopulation databases, and concludes that they indeed are necessary. This report is indicative of the absence of general acceptance. There is not just one author trying to make a point, but rather a group of people that has reached a consensus in rejecting one aspect of the current methods of forensic use of DNA evidence.


Indeed, a few judges have reacted to the NRC report like sharks scenting blood. Despite the report's endorsement of the principles underlying VNTR studies, these judges perceive vulnerability without pinpointing its location or extent. In State v. Futch, 860 P.2d 264 (Or. Ct. App. 1993), prosecution and defense experts presented binotype frequencies ranging from \(10^{-18}\) to \(10^{-5}\), and a majority of the appellate court concluded that such disagreements were matters for the jury to sort out. But a dissenting judge, citing only the first page of the report's summary, insisted that:

The National Research Council, an equally august body [as the Office of Technology Assessment], reports that important questions have been raised about the reliability and validity of forensic RFLP. . . . [T]he technique has not yet achieved general acceptance.