rarely, if ever, be readily apparent. The jury needs to get a sense of whether a particular multilocus pattern is a common or rare event in the world and it gets this sense from estimates based on data bases differentiated solely on the basis of racial lines just as readily as it will ever get it from estimates based on data bases differentiated on the basis of some mythically relevant ethnic lines.

I do not belittle the importance of studying interethnic genetic variation in terms of VNTR genes. I do, however, belittle the effort to bootstrap from a perceived lack of study in this regard to a conclusion which the defense was unable to impress on the court in *United States v. Yee*—i.e., the conclusion that the fact of a match is irrelevant absent meaningful population data and that, since (because of possible substructure) the population data are not meaningful, DNA evidence is irrelevant. DNA evidence is highly relevant as it is currently being presented in our courts, and further studies of ethnic variation will neither diminish nor enhance its relevance to any meaningful extent.

JAMES R. WOOLEY

Organized Crime
Strike Force Division
Northern District of Ohio
Cleveland

Reference

Lander ES (1991) Research on DNA typing catching up with courtroom application. Am J Hum Genet 48:819–823
United States v Yee, 134 FRD 161 (ND Ohio 1991)

0002-9297/91/4904-0024$02.00

Am. J. Hum. Genet. 49:893–895, 1991

Comments on DNA-based Forensic Analysis

To the Editor:

I wish to respond to Eric Lander's (1991) invited editorial recently published in the AJHG *Journal*. I am concerned that this editorial does not contribute to solutions but, instead, confuses the issues.

In my opinion, there are four major scientific issues in application of DNA personal identification technology; they are (1) the scientific principles backing the DNA methods, (2) the criteria which determine match, mismatch, and inconclusive data with regard to RFLPs of forensic sample and suspect, (3) the "significance of the RFLP(s) match," and (4) quality control and assurance of data.

The questions related to issue 1 have been reviewed by the U.S. Congress OTA (1990, pp. 7–8), with the following excerpted conditions: "The Office of Technology Assessment (OTA) finds that forensic uses of DNA tests are both reliable and valid when properly performed and analyzed by skilled personnel." The National Academy of Sciences (NAS) will report its review shortly. The questions related to issue 2 have been studied and reported in the paper by Budowle et al. (1991). Each forensic laboratory is expected to establish, on the basis of laboratory performance of its protocols and staff, its standards for match. I find their decisions on match, mismatch, and inconclusive to be rational. The questions related to issue 3 have been considered in the OTA report, and the following summarizing conclusions have been made: That scientific principles of population genetics can be applied to forensic DNA analysis is not in question, but how best to apply which principles to RFLP analysis is under debate. Disagreement exists as to the extent to which such debate can or should be resolved (OTA 1990). The NAS will report its review shortly. The questions related to issue 4 will be reviewed in the NAS report and have already been published in the Federal Bureau of Investigation's TWGDAM and Association of Crime Laboratory Directors (ASCLAD) quality-assurance policy statements.

Lander argues that "Caucasian," "black," and "Hispanic" are not adequate as genetic classifications of population data bases. I would argue that they are operational genetic classifications which are readily understood in the courts. The possibility of population substructure and significant allelic variation has been proposed by Lewontin as a possible flaw in calculating significance of RFLP(s) match. Lewontin has always chosen outrider alleles to make this point. Lander's editorial implies that we lack significant subpopulation data to make an estimate of match significance. I disagree. I feel that it is possible to make an estimate of significance of RFLP(s) match by using available data bases. Let me illustrate the methods of that estimation.

1. One could ignore population genetics. Using a data base of individual RFLPs, one could argue to the courts that a given matching haplotype had or had

not been observed for X number of individuals. Juries would understand this statement. The significance of RFLP(s) match would be set by the number of individuals studied and would ignore population-genetic principles. It is simple. It would represent a loss if the field of population genetics were to serve the field of forensic science.

2. One could use the population-genetic principles of independent segregation of nonlinked genetic markers, allele frequencies, and population data bases to estimate significance of RFLP(s) match.
3. One could use (*a*) the principles of method 2 and (*b*) *multiple population data bases* to estimate the significance of RFLP(s) match. This would have the benefit of giving a range of values.

I have illustrated these three methods by using haplotypes of three individuals chosen from our Houston data base; from our Caucasian, Afro-American, and Mexican-American data bases a single individual was chosen. Each has a four-probe haplotype with eight RFLPs. *No match for these three haplotypes was found within our data base of 748 individuals.* This would satisfy a nongenetic—i.e., empiric—statement of significance of RFLP(s) match. Table 1 illustrates the significance of RFLP(s) match for the haplotype of these three individuals by using the FBI binning method, as published in the Budowle et al. (1991), article and our Houston data bases. It is clear that, regardless of the data base used to calculate significance of RFLP(s) match, a four-probe eight-allele haplotype is a rare genetic character. I feel that it is reasonable to present such data to a jury for their deliberations. It indicates a range of significance of RFLP(s) match. It is operationally useful.

If we were to develop subpopulation data bases, how would we use them in the U.S. courts? Lander suggests that in the calculations of significance of RFLP(s) match there is a problem with regard to the theoretical significant allele substructure—but he does not suggest a solution. Do we use a Neapolitan or Sicilian data base on a fourth-generation Italian defendant? Are we compelled to use a Columbian data base on a Columbian immigrant defendant? Do we use a selected African data base on a U.S. Afro-American defendant? Do we use a Dublin data base on a third-generation New York Irishman? Where does population data gathering stop, and how is it used?

The continuous-allele DNA RFLP analytic method is very powerful in its identification and exclusion power. It exceeds the power of discrimination of blood group and protein markers—which are now admissible in courts. To exclude genetic principles in estimating DNA-based RFLP matches would be a step backward in the application of population genetics to forensic science. To argue that lack of population substructure data precludes our estimating a significance of RFLP(s) match is absurd, in my opinion. *The admission to a court that the significance of RFLP(s) match can be calculated by a variety of methods and against several data bases is honest.* The courts deserve that this highly accurate genetic forensic method be admissable for both wrongly and correctly charged defendants.

It is reasonable to use genetic principles regarding DNA genetic markers in the courts, to present data and interpretations regarding matching RFLPs, and to estimate the significance of RFLP(s) match against U.S. data base(s) such as "Caucasian," "black," and "Hispanic." Since the DNA methods are so powerful, the courts should be assured of highest-quality data and laboratory quality assurance.

C. Thomas Caskey
Institute for Molecular Genetics and
Howard Hughes Medical Institute
Baylor College of Medicine
Houston

Table 1

Estimate of Significance of RFLP(s) Match

Defendant	Frequency of Match in		
	Caucasian Data Base	Afro-American Data Base	Mexican-American Data Base
Caucasian 381	2.50×10^{-8}	1.67×10^{-8}	1.00×10^{-7}
Afro-American 861	1.67×10^{-10}	2.00×10^{-8}	5.00×10^{-8}
Mexican-American 661	3.33×10^{-9}	3.33×10^{-8}	5.00×10^{-8}

References

Budowle B, Giusti, AM, Waye JS, Baechtel FS, Fourney RM, Adams DE, Presley LA, et al (1991) Fixed-bin analysis for statistical evaluation of continuous distributions of allelic data from VNTR loci, for use in forensic comparisons. Am J Hum Genet 48:841–855

Lander ES (1991) Research on DNA typing catching up with courtroom application. Am J Hum Genet 48:819–823

US Congress Office of Technology Assessment (OTA) (1990) Genetic witness: forensic uses of DNA tests. Publication OTA-BA-438. US Government Printing Office, Washington DC

0002-9297/91/4904-0025$02.00

Am. J. Hum. Genet. 49:895–897, 1991

Statistical Interpertation of DNA Typing Data

To The Editor:

Both the invited editorial by Lander (1991) and similar earlier commentaries on the subject of courtroom applications of DNA typing data have led to numerous arguments that simply defy well-known human population-genetic principles. In such criticisms, the authors employ a logic that may be called "reverse logic," whose mathematical validity is highly questionable. It is true that population substructure leads to genotypic proportions that deviate from Hardy-Weinberg expectations (HWE). Population substructure also produces gametic (as well as nongametic) disequilibria. These are well-known population-genetic principles. But Lander (1989*a*, 1989*b*, 1991) and others (e.g., see Cohen 1990) fail to recognize that there are other factors, particularly relevant to the RFLP analysis of DNA typing, which may produce these end results. Therefore, from the observed deviation from HWE and from an observed linkage disequilibrium, one cannot necessarily infer population substructure. It is unfortunate that in the peer-reviewed journals the above-mentioned authors have been allowed to make this inference without validating whether other associated features of DNA typing data conform to the substructuring hypothesis.

First, one might note that deviations from HWE, in the direction of deficiency of overall proportions of heterozygotes, have been noted in the DNA typing data in binned classification of alleles (Budowle et al. 1991). In contrast, it is demonstrated that, when we consider both the incomplete resolution of similar-sized alleles and measuremental errors of allele sizing, no deviation from HWE is detected (Devlin et al. 1990). One could argue that such tests do not have sufficient statistical power for detection of deviation from HWE. To ameliorate this problem, population data from several law-enforcement agencies have been subjected to nonparametric correlation analyses to check whether alleles of different sizes aggregate in any nonrandom fashion to form DNA types of individuals. Such tests, when properly applied (considering that the paternal and materal alleles cannot be distinguished in individuals in a population data base), result in no deviation from HWE. A correlation measure, originally devised for any general continuous trait with unknown (and possibly complex-shape) distribution (Karlin 1981), has substantially more power for detection of deviation from HWE. It can also be shown that Karlin's (1981) nonparametric correlation measure applies for quasi-continuous traits such as allele sizes at VNTR loci; it is distribution free, and its expectation can be derived even if nonrandom aggregation of alleles within individauls occurs because of population substructuring. These results indicate that, even if populations such as U.S. Caucasians, U.S. blacks, or Hispanics are truly substructured, their consequence on deviations from HWE is only trivial and cannot produce effects as gross as the ones indicated in the fictitious examples given (e.g., see Cohen 1990). Furthermore, even though it is well known that in RFLP analysis by Southern blot protocol the possibility exists that certain alleles of extreme sizes may remain undetected, Lander and others pay no attention to this in explaining the observed heterozygote deficiency. There is a voluminous literature (e.g., see Skibinski et al. 1983; Gart and Namm 1984; and cited references) that deals with such issues. It can be shown that even an extent of 6%–10% overall heterozygote deficiency can be explained if the frequency of such "nondetectable" alleles is 3%–6%. Samples of quite large sizes (e.g., more than 1,500–5,000 individuals/population) would be required for one to observe any single homozygote individual both of whose alleles are nondetectable. Even if this is found, there is no way to distinguish this type from those due to other vagaries of DNA typing (such as DNA degradation, insufficient DNA, etc.). Therefore, covert nondetectability of extreme-size alleles is a much *simpler* explanation of heterozygote deficiency of binned allele data.