Letters

Forensic DNA Typing

The recent article criticizing current methods in forensic DNA typing by R. C. Lewontin and Daniel E. Hartl (20 Dec., p. 1745) is very properly criticized in turn by Ranajit Chakraborty and Kenneth K. Kidd (Perspective, 20 Dec., p. 1735). They point out that Lewontin and Hartl use old and inappropriate blood group data to bolster their contention that allele frequency differences between human groups might affect the calculated probability of a match between two DNA samples.

A point not mentioned by Chakraborty and Kidd is that the variable number of tandem repeat (VNTR) polymorphisms are utterly different from any polymorphisms that have been dealt with heretofore by population genetics theory. So high is the rate of change at these alleles that a very large reason for the relative uniformity of allele frequencies from one human group to another is a kind of mutational churning process. This results from a high rate of unequal crossover and from other very common genetic events. As a result, most of the "alleles" that are grouped by the observer in a size bin are likely to have different evolutionary histories that by chance have led them to occupy the same position of the gel. I have used (*1*) the analogy of the New York and Tokyo stock exchanges to clarify this point for the lay public. Both stock exchanges consist of general groupings of companies—automotive, financial, electronics, and so on—and viewed at this level appear to be similar. But when individual companies within these categories are examined, most will be found to be different, with different characteristics and histories.

At the end of their article, Lewontin and Hartl make some recommendations, under the heading, "What is to be done?" (This, you will remember, is the title of Lenin's famous 1902 pamphlet, in which he made recommendations about the future of Russia. And we all know what came of *that*.) One recommendation is to gather detailed data about allele frequencies in different racial groups. If there were meaningful differences between racial groups, and if there were knowledge of the likely race of the criminal, it would be sensible to do this. But the churning process has homogenized racial groups to a remarkable degree, as shown in figure 1 of Chakraborty and Kidd's paper. Trying to gather meaningful information about the frequencies of thousands, and perhaps hundreds of millions (*2*), of different highly mutable alleles would surely be a waste of effort.

The high mutability of these alleles does pose a real problem, that of mosaicism. It will be essential in the future to match not only the DNA of the accused but the tissue from which the DNA is derived (*3*). Particularly if only one VNTR is used, blood and semen might show very different patterns!

CHRISTOPHER WILLS
Department of Biology, University of California, San Diego, La Jolla, CA 92093

REFERENCES

1. C. Wills, *Exons, Introns and Talking Genes: The Science Behind the Human Genome Project* (Basic Books, New York, 1991).
2. A. J. Jeffreys, A. MacLeod, K. Tamaki, D. L. Neil, D. G. Monckton, *Nature* **354**, 204 (1991).
3. G. Carmody, personal communication.

The disagreement between Lewontin and Hartl and Chakraborty and Kidd over the probability value of obtaining a false match in forensic DNA fingerprinting reminds me of an exchange some years back between a television reporter and the designer of the rocket car that Evel Kneivel would try to jump over the Snake River Canyon. The designer offered the opinion that Kneivel had about an 80% chance of making the jump successfully. "That good?" said the reporter. "Good?" said the designer, "You think that's *good*?" It all depends on your perspective.

Ultimately at fault for the furor is our legal system, which is innocent of quantitative thinking, as well as logic, in its treatment of evidence. Thus, evidence that has been scientifically demonstrated to be highly unreliable, such as the testimony of eyewitness passersby, is routinely admissible, while the relevant reliability studies are not (*1*). Other evidence that is known to be virtually worthless, such as polygraph results [scientific studies by disinterested parties commonly find error rates of 20 to 30%, not too impressive for true-false questions (*2*)], is admissible in certain cases in selected jurisdictions. Still other forms of evidence, such as forensic ballistics-matching and handwriting identification, do not have their empirical foundation critiqued because they were validated years ago in a less critical scientific climate. Nonscientific evidence, by which is meant any evidence gathered by techniques that the lay public can readily understand, has to meet no reliability criteria at all. Thus, testimony from animal trainers who claim their dogs can make olfactory linkages between suspects and evidentiary objects is routinely accepted without any validation (*3*).

In an ideal world, the DNA controversy might compel the legal profession to rethink the fundamental logic concerning how evidence is treated and perhaps to set up unified criteria applicable to both "scientific" and "nonscientific" evidence and between old and new forensic techniques. I would like to see how real (as opposed to DNA) fingerprint data would stand up to the same sorts of inquiry to which DNA fingerprinting has already been subjected. We already know that real fingerprints can yield false positives and, in the United States anyway, there are no standard criteria for what constitutes a match between fingerprints (*4*). In this dream world, the legal profession might even consider what sort of quantitative meaning there might be to such common legal phrases as "beyond a reasonable doubt."

However, in the real world, even if commonly cited figures for the reliability of DNA fingerprints are overestimates by "two or more orders of magnitude," this still makes them of far more probative value than nearly any other form of forensic evidence. In this context, it is the relative, rather than the absolute, reliability that should concern the courts. In the meantime, of course, all scientists will agree that we should work to keep improving on these estimates.

STEVEN N. AUSTAD
Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

REFERENCES

1. E. F. Loftus, *Eyewitness Testimony* (Harvard Univ. Press, Cambridge, MA, 1979).
2. B. Kleinmuntz and J. J. Szucko, *Nature* **308**, 449 (1984).
3. I. L. Brisbin and S. N. Austad, *Anim. Behav.* **42**, 63 (1991).
4. J. E. Starrs, *J. Police Sci. Administr.* **12**, 287 (1984).

We support Daniel E. Koshland, Jr.'s objective approach in handling the article by Lewontin and Hartl about DNA typing by commissioning a rebuttal from Chakraborty and Kidd. Such balanced journalism is absolutely necessary for a scientific issue of such social and legal impact.

The major controversy raised by Lewontin and Hartl stems from their contention that there should be subgrouping of the current DNA population databases into specific ethnic groups. Lewontin and Hartl assume that subdivision of racial population databases is necessary because of two comparisons they performed. Based on a 1954 population database, the first comparison shows a 247-fold difference in the multilocus genotypic frequencies between Italians and Poles with respect to their red blood cell surface antigens. Chakraborty and Kidd,

using logic identical to that of Lewontin and Hartl but a more recent database (1987), demonstrate that only a 0.97-fold difference exists in the multilocus genotypic frequencies of blood group antigens between these ethnic groups. The second comparison presented by Lewontin and Hartl suggests substantial differences between French and Israeli allele frequencies at the D2S44 locus. Chakraborty and Kidd demonstrate that the appropriate interpretation of fixed bin databases (each bin must contain at least five observations) results in much smaller differences between the French and Israeli D2S44 bin frequencies. Therefore, in these cases, by applying current red cell antigen data and correctly interpreting DNA fixed bin data, specific databases for subpopulations of European ethnic groups are not necessary.

We contend that reference groups based on ethnic origins will not provide a reliable source of information with which to address the issue of substructuring because of the heterogeneity that exists in most second- and third-generation American Caucasians and blacks. Because of this contention, as well as the difficulties inherent in the data collection procedure, we have not accumulated our DNA RFLP (restriction fragment length polymorphism) databases according to the ethnic origin of typed North American individuals. Our databases (n = 20,000 Caucasians, 20,000 blacks, and 3,000 Hispanics) have been accumulated on the basis of race. We are now in the process of subdividing these racial databases on the basis of geography to determine whether these binned RFLP frequency distributions differ significantly.

In paternity testing, the choice of the correct reference population database for stating an appropriate paternity index was addressed at the Airlie Conference in 1982 (*1*). Polymorphic genetic systems are used that correspond to red blood cell antigens and enzymes, serum proteins, and leukocyte antigens. Differences exist between subgroups with respect to individual genetic systems; however, an insignificant change between European ethnic groups was observed when several genetic systems were used to generate the cumulative paternity index (*2*). This is in accordance with the finding by Chakraborty and Kidd that several DNA loci result in insignificant differences between major racial groups.

DNA typing is an extremely useful tool for forensics and paternity analysis. Many times it is the means by which a falsely accused individual can be exonerated (31% of our DNA work for parentage analysis results in an exclusion). If one cannot exclude the accused, then we believe that the current racial population databases are sufficient for calculating a valid estimate of the probability of a match between the suspect and the evidence.

ROBERT A. BEVER
MICHAEL DEGUGLIELMO
RICK W. STAUB
CHARLES M. KELLY
R. SCOTT FOSTER
Genetic Design Inc.,
Greensboro, NC 27409

REFERENCES

1. *Inclusion Probabilities in Parentage Testing* (American Association of Blood Banks, Arlington, VA, 1983).
2. K. Hummel and J. Gerchow, *Biomathematical Evidence of Paternity* (Springer, Berlin, 1981), pp. 97–108.

I am appalled at the events surrounding the publication by Lewontin and Hartl that were described in the December 20 News & Comment article by Leslie Roberts (p. 1721). It seems to me inconceivable that scientists would attempt to suppress publication of a paper because they disagreed with its conclusions, a paper which apparently had gone through what one assumes was a normal and stringent review process by independent referees and had been accepted for publication. The vehemence and lack of scientific objectivity that appear to surround this issue indicate that there may be important concerns other than scientific ones. I urge that *Science* obtain from those most closely involved in this debate information about possible economic interests in DNA typing and provide this information to the reader, as other journals have sometimes done (*1*).

I believe the scientific community owes a tremendous debt of gratitude to those who have called to its attention the possible problems and in some cases misuse that have occurred with forensic DNA typing. We are all aware of the costs one can encounter by going against the prevailing dogma, whether it was developed by government or the scientific community. Those who would question it often pay a very high price.

The calls by Eric Lander and others for quality control and proficiency testing are right on the mark. If human lives are going to be taken on the basis of such evidence, and they are, then it is incumbent on all involved to ensure that, in as much as humanly possible, data are correctly obtained and fairly interpreted. For this reason, it would probably be best if forensic cases were analyzed, with the use of standardized protocols, in nonprofit state or national labs that are subject to proficiency testing. As an example, the European countries have apparently recently standardized DNA profile testing (*2*).

Last, I believe that serious questions remain about several aspects of forensic DNA typing, such as band shifting, match criteria, and the definition of allele. Some define allele on the basis of length and assume that all repeats are identical (*3*), whereas A. J. Jeffreys *et al.* (*4*) define alleles on the basis of sequence. Jeffreys *et al.* show clearly that bands may have identical numbers of repeats yet differ in sequence. Thus, as stated by B. Budowle *et al.* (*5*), the currently used methodology permits phenotyping of variable number of tandem repeats profiles but not genotyping, in contrast to the conclusion of Chakraborty and Kidd.

LYNWOOD R. YARBROUGH
Department of Biochemistry and Molecular Biology,
University of Kansas,
School of Medicine,
University of Kansas Medical Center,
Kansas City, KS 66160–7421

REFERENCES

1. J. Brookfield, *Nature* **355**, 207 (1992).
2. *Lancet* **335**, 381 (1990); J. Hodgson, *Bio/Technology* **8**, 798 (1990).
3. B. Devlin, N. Risch, K. Roeder, *Am. J. Hum. Genet.* **48**, 662 (1991).
4. A. J. Jeffreys, A. MacLeod, K. Tamaki, D. L. Neil, D. G. Monckton, *Nature* **354**, 204 (1991); A. J. Jeffreys, R. Neumann, V. Wilson, *Cell* **60**, 473 (1990).
5. B. Budowle *et al.*, *Am. J. Hum. Genet.* **48**, 841 (1991).

Regarding the adequacy (or lack thereof) of DNA typing as it is practiced at present in the forensic community, surely it is not often that an Editor insists on revisions to the galleys of an article accepted after peer review. Even more remarkable (and all credit no doubt due to the Editor) is to commission a rebuttal to the article and to publish it contemporaneously. Save for an uncritical account filtered through a staff reporter (Leslie Roberts), oddly missing has been direct comment, so often heard on other issues, from the Editor who stands at the center (or more accurately to one side) of the controversy. Having first stirred the pot, where was he when it came time to eat the meal, be it cake or crow?

DON W. CLEVELAND
Department of Biological Chemistry,
Johns Hopkins University
School of Medicine,
725 North Wolfe Street,
Baltimore, MD 21205–2185

Response: I am pleased to address the questions raised by readers Yarbrough and Cleveland, particularly because it may help to correct the impression some readers may have received from erroneous reports in the

popular press. *Science* received no communications whatsoever from either the FBI or any other government agency in regard to the article by Lewontin and Hartl, and we would not have tolerated such pressure had it been exerted.

Science was approached by several geneticists at the October 1991 International Congress of Human Genetics who had seen a copy of the article in press. These scientists were concerned that some of Lewontin and Hartl's more theoretical arguments were not adequately supported by data. We had already been advised by peer reviewers that such concerns were correct, but that the article discussed issues of importance to those concerned with the application of DNA fingerprinting and should be published. To represent this situation to our readers, we asked Lewontin and Hartl either to revise their article so that it was more consonant with the data or to have their article published as a Policy Forum, that is, an opinion piece. When Lewontin and Hartl rejected these options, we asked Chakraborty and Kidd to write a Perspective (which was also peer-reviewed and edited) from an alternative viewpoint. The decision to publish Lewontin and Hartl's article and the accompanying Perspective was guided by our desire to present to our readers the best and most up-to-date developments in contemporary science. Our judgments were consistent with our overall philosophy that the pages of *Science* should reflect the most accurate view of highly controversial scientific issues.

It is standard *Science* policy to factor in any information we receive about possible conflicts of interest, on the part of both authors and reviewers, during the peer-review process.—DANIEL E. KOSHLAND, JR.

Response: The preceding letters raise some issues that were not mentioned in our Perspective (*1*) because of space limitations or because they seemed marginally relevant to population genetic principles.

Wills notes that the blood group data used by Lewontin and Hartl (*2*) are old and inappropriate. In addition to our earlier detailed examination (*1*) of this issue, we would like to point out that Lewontin and Hartl's choice of allele frequency data from specific Italian and Polish subgroups to "represent large immigrant populations in the United States" (*2*, p. 1748) seems unrealistic, as immigrants in the United States from Poland and Italy did not come from any single subgroup in these countries. Without knowing the exact demographic and geographic origins of the immigrants, it is appropriate to use only the national averages (in respective countries) of allele frequencies, as was done in (*1*). Furthermore, Lewontin and Hartl's (*2*) quoted set of blood group allele frequencies are not said by Mourant (*3*) to come from a single Italian and a single Polish subgroup. Therefore, Lewontin and Hartl's reported 247-fold difference in blood group profile frequencies (*2*) is an unrealistic example and does not have any relevance to biological differences between the descendants of these groups. Even if a pair of specific subgroups with these frequencies did exist in Mourant (*3*), the relevance of these data to the current U.S. populations of Italian and Polish origins would still have to be demonstrated.

Wills also notes that the "mutational" processes that generate new alleles at variable number of tandem repeat (VNTR) loci are different from those at traditional blood group and protein loci. Size alterations at VNTR loci are probably also "forward-and-backward" in nature; thus, alleles of an identical size would have nonidentical origins [as noted in our Perspective (*1*), but not acknowledged by Yarbrough]. Mosiacism of VNTR genotypes in different tissues of the same individual, if undetected, might lead to false exclusion, not false inclusion. Thus, wrongful incrimination of innocent suspects would not be caused by mosaicism.

Austad discusses the legal settings of the admissibility of DNA typing in court proceedings. DNA typing is only one line of evidence, and its significance must be weighed in the light of the other evidence available. The use of different standards of admissibility of different pieces of evidence is not only beyond the realm of science, it sets double standards for legal proceedings.

Bever *et al.* present another criticism of Lewontin and Hartl's article (*2*). While DNA "matching" criteria are more stringent than binning and rebinning of allele-size data, the large degree of conservativeness in DNA profile frequencies is evident from the tabulations of rebinned allele size data published recently (*4*). This tabulation is used in casework analysis at forensic laboratories that use the FBI Academy's protocol. In the light of this defined rebinning procedure, the Israeli-French differences in D2S44 allele frequencies, reported in (*2*), are irrelevant, and they would not contribute to DNA profile frequency estimates that would be computed from rebinned allele size data. We agree with the comment by Bever *et al.* about the use of major racial group data in U.S. court cases and note that a government-sponsored epidemiologic study, conducted before 1980, of geographically different populations in the United States revealed phenotype frequency differences at blood group and protein loci (*5*). This study explained such differences by the racial makeup of populations, not by geography and economic indicators. Remarkable similarities of phenotype frequencies were shown between samples collected according to a highly structured study design and those collected at blood banks, evidence that blood donor samples provide a reasonably accurate representation of the population frequencies of DNA or other genetic profiles of individuals.

Both Yarbrough and Cleveland raise questions about our Perspective having been invited in an inappropriate way. Cleveland's use of the verb "to commission" is particularly objectionable, because it implies that there might have been an economic interest in our accepting the invitation. Such an implication is entirely baseless. The issue here is the scientific basis of criticisms of the applicability of DNA typing, and to put a different label on this issue is injurious to the scientific value of the debate. Like all other scientific contributions to *Science*, our Perspective (*1*) went through the stringent impartial review process of the journal, and it was accepted on the basis of these reviews and our response to them.

We would answer Yarbrough by stating that protocols for standardization and proficiency testing are in place in the United States as well as in Europe [(*1*), reference 4; (*6*)], and internal proficiency testings in several laboratories have been in operation for more than 3 years. Binned definition of alleles also does not diminish the validity of frequency estimates because, as we mentioned (*1*), the definition of the alleles is technology-based and the similarity of binned VNTR alleles and serologically defined alleles is not our invention; it is a generally accepted view (*7*). Moreover, Risch and Devlin clearly show (*8*) that the crude definition of binned alleles leads to conservative DNA profile frequencies, further diminishing the already miniscule possibility of wrongful incrimination of innocent suspects.

RANAJIT CHAKRABORTY
Center for Demographic and Population Genetics, University of Texas Graduate School of Biomedical Sciences, Houston, TX 77225
KENNETH K. KIDD
Department of Genetics, Yale University School of Medicine, New Haven, CT 06510

REFERENCES AND NOTES

1. R. Chakraborty and K. K. Kidd, *Science* **254**, 1735 (1991).
2. R. C. Lewontin and D. L. Hartl, *ibid.*, p. 1745.
3. A. E. Mourant, *The Distribution of the Human Blood Groups* (Blackwell, Oxford, United Kingdom, 1954).

4. B. Budowle *et al.*, *Crime Lab. Digest* **18**, 9 (1991).
5. B. Cohen *et al.*, *DHEW Publication No. (PHS) 80-1664* (National Center for Health Statistics, Department of Health, Education and Welfare, Hyattsville, MD, 1980), p. 28, paragraph 2.
6. Technical Working Group on DNA Analysis Methods, *Crime Lab. Digest* **18**, 44, 1991; M. A. Farley and J. J. Harrington, *Forensic DNA Technology* (Lewis, Chelsea, MI, 1991), p. 217.
7. J. W. Morris *et al.*, *J. Forensic Sci.* **34**, 1311 (1989).
8. N. Risch and B. Devlin, *Science* **255**, 717 (1992).

Response: The 20 December 1991 issue of *Science* contained our scientific article (*1*) on the forensic applications of DNA typing. In an unprecedented move, the Editor commissioned a rebuttal from Chakraborty and Kidd that was published in the same issue as a Perspective (*2*). In response, we would like to make a few points that are relevant not only to the Perspective but also to the letters published in this issue.

Chakraborty and Kidd (*2*) argue that genetic differentiation among subgroups has a negligible effect on Hardy-Weinberg frequencies in mixtures of different subgroups. This was one of the main points of our article. When one assesses the reliability of chain multiplication in estimating the probabilities of DNA matches, the critical question is whether significant variation in gene frequencies occurs among human groups. As we all agree that statistical tests for deviations from Hardy-Weinberg are virtually useless for detecting variation among subgroups, this argument is a red herring that diverts attention from the real issue.

Chakraborty and Kidd "revise" our quoted blood group frequencies by averaging frequencies over many Polish and Italian subpopulations. But Poles and Italians are real people, not averages. Our particular Polish and Italian subgroups had a real difference in genotype frequency of a factor of 247 [(*1*), table 1]. Chakraborty and Kidd reduce the difference down to a factor of 1.21 [(*2*), table 1], as noted by Bever *et al.* in their letter, but in the final analysis the exercise demonstrates the point we tried to make—averaging gene frequencies over many disparate subgroups obscures real differences between them and results in incorrect estimates. Even within countries, significant genetic variation from population to population exists as a result of past migrations. For example, the Piemontese in northern Italy are related to Germanic peoples through the Lombard invasions, while Sicilians have affinities with Greeks and North Africans. In the averaging done by Chakraborty and Kidd (*2*), the 19th-century Italian Unification, a political event, is not distinguished from the biological reality of Italian subpopulations. Much past history is retained in subpopulations, even in advanced countries with supposedly mobile populations. Southern Spain still retains the gene frequencies of its Arab occupiers after 500 years. In Japan, there is a cline in gene frequencies from Southern Honshu to Northern Hokkaido as a result of repeated mainland invasions that ended in the 13th century. The British Isles still show clear clines of gene frequency corresponding to the displacements of the Celts and Picts by the Jutes and Angles in the first millennium.

Our comparison of the French and Israeli data [(*1*), table 2] was based on bins with ten or more observations in at least one of the subgroups. Chakraborty and Kidd assert (*1*) that no bin should have less than five observations in any of the relevant subgroups. This interesting new rule has the obvious consequence that all allelic types must be present at this level or higher in all ethnic subgroups in the database. Applying this rule consistently would require precisely the kind of data on ethnic subgroups that we believe is necessary. Bever *et al.* enthusiastically support the new rule, but we do not see how it could be applied to the databases of their firm, Genetic Design Inc., which performs paternity tests, because, as Bever *et al.* state, their databases contain no ethnic information beyond race. Furthermore, the databases cited by Bever *et al.* contain numerous trios of mother, child, and putative father, so the individuals in the databases are not statistically independent.

Chakraborty and Kidd also inform us (*2*) that the French and Israeli data are wrong anyway. The French and Israeli data are not *our* data, but were introduced by the FBI in the case of *U.S.* v. *Yee* (*3*) in the apparently mistaken belief that they demonstrated the absence of population differentiation. The FBI's chief DNA typist is quoted as saying that the data are wrong (*2*, reference 35). This troublesome revelation is reminiscent of other FBI data provided in *U.S.* v. *Yee*, in which the DNA types of 16% of 225 individuals were not verified in a second examination. Chakraborty and Kidd and others argue fine points of population genetics but do not discuss the apparent high rate of laboratory irreproducibility.

Chakraborty and Kidd (*2*) give a citation without page numbers [(*2*), reference 47] to a book by F. M. Salzano and S. M. Callegari-Jacques to support their statement that Amazonian Indians form an essentially genetically homogeneous equilibrium population (*4*). On the contrary, this particular book, and Salzano's lifetime of work, is a documentation of genetic differentiation and clines among these very tribes, as illustrated by the statement that "north-south clines have been observed in the prevalences of alleles *Le*, R^1, and R^2; east-west gradients . . . could be demonstrated for ESD^1 and Gc^1 . . . correlations between alleles exist . . . . [reflecting] possible ancient migration routes . . ." (*4*, p. 208). Interested readers should also examine chapter 9 (*4*, pp. 178–204) and particularly figures 8.1 (*4*, p. 167), 9.9 (*4*, p. 198) and 9.10 (*4*, p. 199). Chakraborty and Kidd do not distinguish the lack of differentiation between villages *within* tribes (because of the so-called "fission-fusion" process of village formation) from the often very large differences *between* tribes. The census group "Hispanics," which is the group at issue in DNA typing (*3*), includes groups with different ancestries: South American Indian, Central American Indian, Northern Mexican Indian, African, or none of the above.

Chakraborty and Kidd assert (*2*) that American demography for descendants of Caucasian immigrants is close to a "melting pot," but they cite no data. Chakraborty contradicts himself on this point in his recent paper (*5*). He states, "Older industrial American cities such as Pittsburgh attracted European immigrants because they offered economic opportunities. These immigrants maintained their cultural separateness, *still reflected in the urban neighborhoods of Pittsburgh*" [(*5*), p. 152, emphasis added]. The same paper also points out that "African slaves originated from a variety of geographic locales, and their disposition in the New World was not random" [(*5*), p. 152].

Chakraborty and Kidd state (*2*) that the data in the reference populations "are not collected from homogeneous endogamous subgroups . . . but these databases are congruent with the operational definitions of reference populations" [(*2*) p. 1736]. In other words, in the Alice-in-Wonderland world of DNA typing, a reference population is what you say it is, nothing more, nothing less. However, up here in the real world, there are valid reference populations that give the right answers and invalid mixtures that give the wrong answers, and the treatment by Chakraborty and Kidd of the Polish and Italian data demonstrates this difference convincingly.

Chakraborty and Kidd assert (*2*) that the use of separate reference databases for Caucasians, blacks, and Hispanics is tantamount to the method of ethnic ceilings that we described (*1*). We discussed the method of ethnic ceilings explicitly with regard to ethnic subgroups *within* the major races; moreover, the ceilings are defined locus by locus. Their treatment of the Polish and Italian data demonstrates the errors that arise when heterogeneity within human groups is not considered. Probabilities differing by two or more orders of magnitude are commonplace when different databases are used (even the crude FBI racial databases), as demonstrated by the scatter of points around the straight

line in figure 1 of (*2*). Chakraborty and Kidd refer to more than 2000 U.S. court cases that have employed DNA evidence [(*2*), reference 4]. This is misleading, as the majority are paternity cases in which the DNA types of mother, child, and putative father are usually available for comparison. Chakraborty and Kidd compare binning the variable number of tandom repeat (VNTR) sequences with the grouping of A_1 and A_2 blood type alleles. This comparison is inappropriate because the problem with binning VNTRs is not that alleles are grouped but that sometimes they are assigned to the wrong bin. Their statement that the "worst case" scenario is an equal mixture of Poles and Italians is incorrect—it is actually the best case for their argument. Their statement [(*2*), p. 1737] that the "arithmetic and underlying principles are identical" for linkage equilibrium and Hardy-Weinberg equilibrium is simply incorrect. In a quote from our article, they also did not include an important qualifier present in our original text [(*2*), reference 6].

Wills states in his letter that our arguments are based on "old" blood group data. Why Wills would disregard reliable blood group data is unclear, because even Chakraborty and Kidd concede their relevance. Wills also cites a high mutation rate among some VNTRs as the basis of a "mutational churning process" that is "a very large reason for the relative uniformity of allele frequencies from one human group to another. . . ." This hypothesis "explains" what has not yet been shown to exist, as relative uniformity of allele frequencies is precisely the point in dispute. We see no evidence of VNTR "churning" in French and Israelis or in the South American Indian tribes studied by Kidd *et al.* (*6*) [(*1*), p. 1749]. Wills also makes much of the observation that one of our subheadings (*1*) is identical to the title of a 1902 pamphlet by V. I. Lenin. This coincidence has no more relevance to DNA typing than the fact that Wills' letter makes favorable reference to Japanese automobile companies.

Austad concedes the validity of our arguments, but points out that DNA typing is certainly better than polygraph results, so why all the fuss? His argument seems to be that new sources of scientific evidence should be held to a standard of reliability no greater than methods currently in use. We would make a fundamental distinction between the intrinsic limitations of a technology and limitations imposed by the use of false assumptions, particularly when simple alternatives are available. For example, although polygraph examinations have a high intrinsic error rate, we suppose that Austad would object to a test in which an electrical short in the machine produced additional erratic readings. Our view is that erroneous assumptions about genetic uniformity among ethnic groups are no more necessary to DNA typing than electrical shorts are necessary to polygraph machines.

We would finally like to emphasize that this dispute is not about the use of DNA evidence in the courtroom. DNA typing is a very powerful procedure. We regard it as "possibly the most powerful innovation in forensics since the development of fingerprinting in the last part of the 19th century" [(*1*), p. 1746]. All we ask is a basic degree of candidness in reporting the statistical significance of a match. With databases as large as $x = 10{,}000$, why not use $1/x$ as a conservative estimate of the probability [(*1*), p. 1749]? After all, 0.0001 is already a pretty small number. Why invoke unsupported assumptions in order to obtain a still smaller probability that is exaggerated and unreliable? Perhaps it is because the organizations whose interests are served by numerical exaggeration have also been in charge of choosing the statistical procedures.

RICHARD C. LEWONTIN
Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138
DANIEL L. HARTL
Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110

REFERENCES

1. R. C. Lewontin and D. E. Hartl, *Science* **254**, 1745 (1991).
2. R. Chakraborty and K. K. Kidd, *ibid.*, p. 1735.
3. *U.S.* v. *Yee*, 129 Federal Rules Decisions 629 (Northern District of Ohio, 1990).
4. F. M. Salzano and S. M. Callegari-Jacques, *South American Indians: A Case Study in Evolution* (Oxford Univ. Press, New York, 1988).
5. R. Chakraborty *et al.*, *Am. J. Hum. Genet.* **50**, 145 (1992).
6. J. R. Kidd *et al.*, *Human Biol.* **63**, 775 (1991).

Erratum: In the abstract and in the text (line 35 in the middle column of page 185) of the report "Electrical resistivity and stoichiometry of K_xC_{60} films" by G. P. Kochanski *et al.* (10 Jan., p. 184), the minimum resistivity was given incorrectly as 2.2 microohm-cm. The correct value is 2.2 milliohm-cm.

Erratum: In the News & Comment article "Is homosexuality biological?" by Marcia Barinaga (30 Aug., p. 956), it was suggested incorrectly that the suprachiasmatic nucleus is not part of the hypothalamus.

Erratum: The Table of Contents for the issue of 31 January 1992 (p. 508) incorrectly listed a letter by J. Bello as appearing in the Letters section beginning on page 514. The letter appeared in the issue of 14 February on page 784.

Erratum: In figure 1 (p. 1509) of the Research Article "Radar images of Mars" by D. O. Muhleman *et al.* (27 Sept., p. 1508), the Mars longitude of the sub-Earth point was mislabeled in each of the six snapshot radar images of Mars. None of the labels should have contained a decimal point. The values of λ in the labels should have been 78, 92, 104, 120, 133, and 147, respectively.