Proc. Natl. Acad. Sci. USA
Vol. 78, No. 5, pp. 3108–3112, May 1981
Genetics

Is there a difference among human populations in the rate with which mutation produces electrophoretic variants?

(tribal mutation rates/human electrophoretic variants/human population sampling theory)

JAMES V. NEEL AND EDWARD ROTHMAN

Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Michigan 48109

Contributed by James V. Neel, January 19, 1981

ABSTRACT **Data are summarized that suggest that tropical-zone/tribal/nonindustrialized populations have higher frequencies of certain types of protein variants than temperate-zone/civilized/industrial populations, and it is demonstrated that these differences are not an artifact produced by the contagious type of sampling used with respect to tribal populations. Evidence is reviewed that suggests that a possible explanation of this difference is higher mutation rates in the tribal populations studied.**

The demonstration of substantial differences in the average mutation rate among various human populations would raise a number of important questions concerning the responsible factors. In this paper, we examine recent pertinent data from two kinds of populations—tropical-zone/tribal/nonindustrialized populations and temperate-zone/civilized/industrialized populations—data that do indeed raise the possibility of such differences.

There are two general approaches to this question. The first simply compares a variety of populations with respect to the frequency of genetic variation assumed to be maintained primarily by mutation and, in the absence of other explanations, attributes the higher mutation rate to the population that has the greater amount of such variation. The second attempts, on the basis of the same kind of variation, to generate an actual estimate of mutation, by either direct or indirect formulations. The results of both approaches will be considered.

A type of genetic variation especially useful for such purposes is that shown by surveys of proteins by using electrophoretic techniques. For studies involving our own species, blood is the most readily sampled tissue. The variants shown by such surveys, originally undertaken among more accessible, civilized populations, are usually classified as polymorphisms (allele frequencies ≥ 0.01) and nonpolymorphisms (allele frequencies < 0.01). The distinction between these two categories of variation is arbitrary but operationally convenient. Considerable disagreement exists as to the extent to which the former are, for the most part, traits responding to positive selection or, rather, in general selectively neutral traits, the alleles for which have 'drifted up' to this relatively high frequency. Less disagreement exists concerning the nonpolymorphic traits; they are generally assumed to be predominantly neutral or deleterious traits whose frequency is maintained by mutation pressure.

Some terminological problems arise with respect to traits of the latter type. The early work defining these traits as variants encountered with average frequencies of 1–10 per 1000 examinations of individuals for any specific protein was largely done on civilized populations (1). When, later, work was extended to tribal populations (from whose relatively recent amalgamation the civilized populations have arisen), a different picture was encountered. Within a given tribe, *specific and apparently unique* variants, seemingly restricted to that tribe (or to adjacent tribes), might range from a few examples in a single family to many examples, with allele frequencies of 0.05–0.10 (cf. refs. 2 and 3). We have referred to these latter as "private polymorphisms" (4). This is the picture to be expected of alleles arising relatively recently (in terms of human history) through mutation, most of which are neutral, slightly advantageous, or even slightly deleterious in their phenotypic effects, when gene frequencies drift upward or are subject to positive selection but tribal boundaries constitute an effective barrier to gene flow (5, 6). In most civilized populations, whose ancestors were drawn from 100 or more different tribes, the commingling will, as a consequence, reduce the overall frequency of any specific example of these localized polymorphisms to that of a rare (i.e., nonpolymorphic) variant (although the total number of copies when distributed across all the tribes that commingled remains the same). In this treatment, we shall include in our calculation of the average frequency of nonpolymorphic variants in tribal populations, all variants of *restricted* distribution, on the justification that this in fact what has been done in the study of conglomerate civilized populations. If mutation rates were similar in the tribal populations and in the populations antecedent to the civilized conglomerates (and in the civilized conglomerates since their fusion) and the average variant was close to neutrality in its phenotypic effects, we would anticipate that the average frequency of rare variants per system should be the same but the number of different variants encompassed within that average should be greater within the civilized groups, because the number of tribes that have contributed "their" variants to civilized populations is greater than the number of tribes thus far sampled in this respect in studies of tribal populations.

FREQUENCY OF NONPOLYMORPHIC VARIANTS

Table 1 presents data on the frequency of nonpolymorphic variants of restricted distribution found in three recent summaries of tribal populations and in three summaries of civilized populations. The data are all concerned with proteins of the erythrocytes, this being the most convenient cell type to sample under field conditions. The criteria for inclusion of a protein in the table are (*i*) data are available for at least three of the six groups and (*ii*) the technique used to demonstrate the variants was starch gel electrophoresis. For all of the vacant cells in the Ann Arbor series, data are available from polyacrylamide gel electrophoresis, with results in general consistent with the starch gel data for the other two samples from civilized populations, but it seemed wise to restrict the presentation to the results of a single technique. Variants occurring as well-known genetic polymorphisms that have been excluded from the tabulation are listed in a footnote.

The number of tribes contributing to the Amerindian data is 21; the sample of Australian aborigines has been widely drawn but the number of tribal entities has not been specified; and 55 language groups (which we will equate to tribes) contributed to the New Guinea sample. Only the northern aspect of Australia can be classified as tropical, so that the rubric under which we are grouping these three sets of populations is something of a misnomer in this case. The civilized populations were com-

Table 1. Comparison of the frequency of alleles encoding for rare variants or private polymorphisms in three tribal-type and three civilized populations

	Tribal population				Civilized population			
Trait	Amerindians (2)	Australian aborigines (7)	New Guinea natives (8)	No./1000*	London (1)	Japan (9)	U.S. (10)	No./1000
Acid phosphatase	33/9171	1/4016	0/7421	1.65	0/7887	0/4431	1/1970	0.07
Adenosine deaminase	2/5391	0/1437		0.29	2/4798	0/4841		0.21
Adenylate kinase 1	0/7513	0/3535	0/5551	0	1/6760	0/4767	0/1940	0.07
Carbonic anhydrase 1	0/4129	200/3751	0/1857	20.54	3/10115	1/4823		0.27
Carbonic anhydrase 2	39/3874	166/3751	0/1370	22.79	0/434	2/4828		0.38
Esterase A	77/4125			18.67		3/4425	1/1968	0.63
Esterase D	0/3872	0/1556	1/5453	0.09	0/454	0/4415	0/1969	0
Hemoglobin α	0/8073	0/2692	46/6874	2.61	4/10971	1/4847	0/1738	0.28
Hemoglobin β	0/8073	0/2692	0/6874	0	8/10971	0/4847	0/1738	0.46
Isocitrate dehydrogenase	8/5746	0/1226	0/4908	0.67	4/718	4/4845	6/1927	1.87
Lactate dehydrogenase A	2/6521	2/4180	10/7858	0.75	2/1015	0/4835		0.34
Lactate dehydrogenase B	65/6521	1/4180	2/7858	3.66	0/1015	0/4835		0
Malate dehydrogenase	1/5891	0/2694	107/7856	6.57	1/516	0/4845		0.19
Nucleoside phosphorylase	0/3887			0	2/1542	5/4560	0/1960	0.87
Peptidase A	28/5904	1/3034		3.24	8/8798	5/4839	5/1962	1.15
Peptidase B	17/6018	67/3189	1/5108	5.94	15/7041	6/4847	6/1967	1.95
Phosphoglucomutase 1	6/8907	5/3919	257/7775	13.01	12/10333	25/4814	3/1967	2.34
Phosphoglucomutase 2	41/8560	109/3790	113/7787	13.06	7/10333	5/4844	3/1972	0.87
Phosphogluconate dehydrogenase	0/9176	65/4035	39/7809	4.95	1/4939	6/4841		0.72
Phosphoglucose isomerase	13/5772	1/1569	9/6934	1.61	1/1550	37/4847		5.94
Triosephosphate isomerase	0/3888			0	2/1705	2/4562		0.64
No./1000	2.53	11.19	5.89	5.38	0.72	1.02	1.08	0.89

The following electrophoretic variants that occur as genetic polymorphisms of widespread distribution have been specifically excluded: $CRPL_A$, $CRPL_{Michigan}$, $CRPL_{New\ Haven}$, TF_{D1}, $TF_{D\,Chi}$, HP_2, $ACP1_A$, ADA_2, AK_2, $CA2_2$, ESD_2, $HGBA1_S$, $HGBA1_C$, $PEPA_2$, $PGM1_2$, $PGM1_7$, $PGM1_9$, and $6PGD_C$.

* Includes homozygotes.

posed of "individuals of European origin, the great majority being residents of the United Kingdom" (1), of Japanese residents of Hiroshima and Nagasaki (9), and of newborn infants in Ann Arbor, Michigan, largely of Caucasian origin, but 8% were born to American Blacks and 2% to Mongoloids, and there is no information for 4% of the sample (10). The tribal antecedents of the three civilized "populations" are of course only partially known. Also, for most of the proteins listed, only a single electrophoretic technique was used in the search for variants, so that these numbers are underestimates of the total amount of electrophoretic variation.

We have included in the data on tribals four traits that, at this stage in our knowledge, are polymorphisms in the ethnic group in which they have been sampled. The basis for this is the belief that, in view of their restricted distribution within the ethnic group, they would in this ethnic group fall below the level of a polymorphism as more tribes (for the most part lacking the variant) are studied. Otherwise stated, we believe that such polymorphisms must have existed among the tribes whose consolidation ultimately led to the various civilized populations treated herein. However, this position cannot be rigorously defended. These variants are PGM1 type 3 in 254 New Guinea natives, PGM2 type 3 in 103 Australian aborigines, CA-1 type 9 in 192 Australian aborigines, and CA-2 type 4 in 166 Australian aborigines.

Table 1 indicates that there are considerable differences among the three sets of tribal populations but that, on average, the frequency of individuals that have variants of restricted distribution is some six times greater in the tribal than in the civilized populations. For some of the private polymorphisms, allele frequencies are high enough that some of the trait bearers are homozygotes. For this reason, it would be somewhat more accurate to present the data in allele frequencies (in which case the difference between tribal and civilized populations would be even greater), but we retain the present approach for ease of comparison with the literature. If the four possible exclusions mentioned above are observed, then the average rare variant frequency in tribal populations drops to $\approx 2.9/1000$, still three times greater than in the civilized groups. On the other hand, this action would force us to recognize a type of low-frequency polymorphism in the tribal populations that is substantially less common in the civilized populations.

Before these differences can be accepted as valid, we must consider the possibility that the manner in which tribal populations are sampled creates a bias that falsely inflates the observed relative frequency of restricted variants. A careful examination of this issue is complicated by our inability to specify an underlying joint distribution for the number of copies of the variants at any given time. On the other hand, by using a collection of models for this distribution that describes most possibilities, we can shed some light on the question of bias.

Let us define $G(j)$, $j = 1, 2, \ldots 2N - K + 1$ as the number of alleles with j copies in a population of size $2N$ having K distinct alleles. These random variables must satisfy two constraints:

$$\sum_{i=1}^{2N-K+1} G(j) = K \quad \text{and} \quad \sum_{j=1}^{2N-K+1} jG(j) = 2N.$$

For a sample of size $2n$, we also define the sample analogues of these random variables, labeled $g(j)$, $j = 1, 2, \ldots 2n - k + 1$, that satisfy the corresponding constraints:

$$\sum_{i=1}^{2n-k+1} g(j) = k \quad \text{and} \quad \sum_{j=1}^{2n-k+1} jg(j) = 2n.$$

Our approach is to examine the expectation of two statistics under a variety of models. These statistics are (*i*) k, the number

of different alleles in the sample at a given locus, and (*ii*)

$$r = \sum_{j=1}^{2n\phi} \frac{jg(j)}{2n}, \quad [1]$$

where ϕ is the arbitrarily defined allele frequency below which a variant is scored; in this case 0.01. r is thus the fraction of genes in the sample representing alleles with frequency $\leq \phi$—i.e., the number of variants per 1000 determinations in Table 1.

To evaluate the expectations of both k and r, we note that

$$E[g(j)] = \sum_{i \geq j} E[G(i)]P[i,j;n] \quad [2]$$

where $P[i,j;n]$ is the conditional probability that an allele having i copies in a population of size $2N$ will have j copies in a sample of size $2n$. For a random sample (without replacement),

$$P(i,j;n) = \frac{\binom{i}{j}\binom{2N-1}{2n-j}}{\binom{2N}{2n}}$$

and, if $n \to \infty$, $N \to \infty$ with $n/N \to f$, then $\lim P(i,j;f) = \binom{i}{j} f^j (1-f)^{1-j}$. It is this limiting form that we shall consider for the case of a civilized population. Because for tribal populations the sampling is "contagious" in nature, the procedure for sampling tribal populations can be viewed as a mixture of these binomials. In particular, we take

$$P(i,j;f) = \int_0^1 \binom{i}{j} z^j (1-z)^{i-j}\, dF(z), \quad [3]$$

where $F(z)$ is a distribution with mean

$$f = \int_0^1 z dF(z).$$

By mixing this binomial distribution, we are increasing the probability of an extreme sample, which is precisely what is required here; the probabilities of observing either every copy of a particular allele present in the population or of observing no copies of this allele is greater than the corresponding quantities for a random sample.

Even without specifying a particular form for the $E[G(j)]$, we can draw some conclusions concerning the expectation of k for a given f and K. As

$$E(k|K,f) = K - \sum_{i \geq 1} E[G(i)]\,(1-f)^i$$

is a concave function in f, then (as a consequence of Jensen's inequality), we have for any positive sequence $E[G(i)]$,

$$E_{\text{civilized}}(k|K,f) \geq E_{\text{tribal}}(k|K,f). \quad [4]$$

Thus, from Eq. 4, we would expect to see more different 'alleles' in a random sample from a civilized population than in a stratified random sample from a tribal population. This result is rather unfortunate because the statistic k is, under most models of the neutral hypothesis, a sufficient statistic for the mutation rate and would be the statistic of choice for this discussion. We are thus led to consider the less efficient statistic, r. Inequalities for the expectation of r, however, depend on the specific form of the sequence $E[G(i)]$. Consider Ewen's result (11), which states that, under certain conditions

$$E[G(i)|K,N] = \frac{2N!}{(2N-i)!}\,\frac{1}{i}\,\frac{|S_{2N-i}^{(K-1)}|}{|S_{2N}^{(K)}|},$$

where $S_a^{(b)}$ is the Stirling number of the first kind (cf ref. 12). When $N \to \infty$, $K \approx \theta \ln 2N$, we have

$$E[G(i)] = \frac{\theta}{i},\ i = 1, 2, \ldots, \quad [5]$$

where $\theta = 4N\mu$. In this case, we can show that the distribution of r does not depend on the sampling scheme and so we have,

$$E_{\text{civilized}}(r) = E_{\text{tribal}}(r) = \theta\phi. \quad [6]$$

So, when Ewen's sampling theory for the neutral hypothesis is appropriate and no difference in the mutation rate is assumed, we expect to find no difference in the relative frequency of *restricted* variants. Furthermore, using the fact that the sequences $G(j)$ are asymptotically mutually independent Poisson random variables, it can be shown that the standard error of this estimate is (as $n \to \infty$),

$$\text{SEE} = r\sqrt{\phi}. \quad [7]$$

(The variance of r is $\approx \phi^2\theta$ as $n \to \infty$; hence, r/ϕ has nonzero asymptotic variance.) In this paper, ϕ is taken to be 1% so the SEE is 10% of r. Thus, the standard errors for variant frequencies in Amerindians and Australian and New Guinea natives are 0.3, 1.1, and 0.6, respectively, and the differences from the frequencies observed in samples from civilized populations are significant.

It is possible that other models (e.g., Bose–Einstein sampling) of the joint distribution of the allele frequencies may be appropriate and, in these cases, differences between the expectation of r under random sampling and stratified sampling may be found. Even with Bose–Einstein statistics, however, we have not, under a wide variety of data patterns, been able to account for the magnitude of the quantitative difference found in the data.

This greater frequency of rare variants in the tribal populations, which we take to be a solid fact, can be explained by any or some combination of the following: (*i*) better survival value of these rare variants in tribal than in civilized populations, (*ii*) a population structure (i.e., stratification, level of inbreeding, and such) more favorable to retaining such variants in tribal populations, (*iii*) higher mutation rates in tribal populations. With respect to the first of these possibilities, it is *a priori* difficult to see why such variants, often characterized by reductions or instability in enzyme activity, should have greater survival value in tribal than civilized populations. Although such a possibility cannot be excluded, there are no data that can be brought to bear on the question. With respect to the second of these possibilities, we can examine the effects of population structure in the same manner as above in our investigation of nonrandom (or contagious) sampling, demonstrating that, under Ewen's sampling formulation, in a population with more structure fewer distinct alleles, k, should be observed but the value r is expected to be unchanged. We are thus left with the third of these alternatives.

DIRECT ESTIMATES OF HUMAN MUTATION RATES

The most satisfactory approach to the study of human mutation rates is the direct method—i.e., the examination of children for the presence of genetic traits not present in either parent. For electrophoretic variants, this requires electrophoretic examination of a series of proteins selected only for the ease and clarity with which variants can be recognized. When a variant is encountered in a child, the parents must be examined. Absence of the variant in both parents is presumptive evidence for mutation or a discrepancy between nominal and biological parentage. The steps to place a probability on the latter contingency have recently been enunciated by (unpublished results). In the application of the direct method, attention is focused on

the rare (nonpolymorphic) variants, on the argument that variants corresponding in electrophoretic mobility to the common polymorphisms have such a low probability of representing new mutations that the possibility can be ignored. This argument could lead to overlooking a mutation that mimics in electrophoretic mobility an established polymorphism.

The direct method is for many reasons more easily pursued among the conveniences of civilization; of the four sets of data available, three are based on civilized groups. The results are given in Table 2. No mutations were observed in 433,261 locus tests conducted in three quite different civilized areas, nor were there any among 94,796 locus tests on these Amerindians. From the relationship

$$0.05 = (1 - \mu)^n,$$

where μ = mutation rate per locus per generation and n = number of locus tests, one can determine hypothesized values of μ that can be rejected at a 5% level of significance. The upper limit of acceptable values is 0.7×10^{-5} per locus per generation for civilized populations such as those studied thus far and 3.2×10^{-5} per locus per generation for Amerindians. As no mutations were encountered in either type of population, the higher bound for the Amerindians simply reflects the smaller number of locus tests, but the usefulness to the argument of setting upper limits will be apparent below.

INDIRECT ESTIMATES OF MUTATION RATE

In situations in which data permitting direct estimates of mutation rate can be accumulated only with great difficulty, as exemplified by remote and thinly settled tribal populations, an indirect approach must be pursued. This approach assumes an equilibrium population in which the traits used in the calculations are neutral with respect to selective value. Although the assumption of neutrality for all the traits listed in Table 1 can be challenged, it has an indispensible heuristic value. The implementation of this approach requires estimates of the number of individuals in a single generation of the population (N), the average proportion of alleles per locus surveyed that are variant (K) (polymorphisms of wide distribution are usually excluded from the estimation of this proportion) and, usually, either an estimate of the average survival time, in generations, of a newly introduced neutral mutation that is destined for ultimate loss ($\bar{t}_0$) or the conditional probability of the transmission of j copies of a mutant allele from i copies of this allele (P_{ij}). Three different formulations for manipulating these data are in current use [those of Kimura and Ohta (13), of Nei (14), and of Rothman and Adams (15)]. Because in practice only a portion of any tribe under study is sampled, some variants occurring in limited numbers could be missed. Rothman and Adams have devised an estimator of the average total number of different variants at a given locus in any specific tribe, given data on the proportion of the tribe sampled and the numerical representations of the various alleles already encountered in that tribe. The correction that this estimator permits will be incorporated into all the calculations using the formulas of Kimura and Ohta and Rothman and Adams but is not indicated in the application of Nei's formula. Although Kimura and Ohta apparently would include all variants in the calculation, Neel (16) and Nei have suggested that polymorphisms of wide distribution should be excluded from any calculation, both because of the possibility that they are maintained by selection and to prevent double counting.

Table 2. Results of four efforts to detect by the *direct* method electrophoretic variants due to mutation in human populations

Population	Locus tests	Mutations	Reference
United Kingdom (London)	113,478	0	(1)
Japan* (Hiroshima/Nagasaki)	208,196	0	(9)
United States (Ann Arbor, MI)	111,587	0	(10)
Amerindians (Central/South America)	94,796	0	(10)

* Parents of these children estimated to have received <1 *rem* from the atomic bombs.

To estimate $\bar{t}_0$, Kimura and Ohta (13), on the assumption of an equilibrium population in which the mean number of offspring is two and the variance in offspring number is two, find

$$\bar{t}_0 = 2\frac{N_e}{N}\ln(2N)$$

in which N_e = "effective" population size. Although, in general $N_e < N$, in a calculation this approximate, N_e is usually equated to N. The breeding structure of tribal populations departs rather radically from these assumptions. Li *et al.* (17) find through simulation based on a specific Amerindian tribe, the Yanomama, as well as through the derivation of a progeny distribution for a mutant allele in a Yanomama-type population, a $\bar{t}_0$ for alleles censused in the adult generation of 5.6. The application of Eq. 1 to this same population leads to an estimate of $\bar{t}_0$ of 11.7. We have used the former value in our calculations of mutation rates in Amerindians, but others have used the results of the Kimura–Ohta formulation. Estimates of $\bar{t}_0$ are population specific, so that, although we have some confidence in the figure we have derived for Amerindians, it should not be injudiciously applied to all tribal-type populations.

The most obvious drawbacks to the indirect approach are the assumptions that the population in question has been free of in-migration for a long period of time and that the variants being detected are neutral. In-migration should bias the estimate upward, due to the introduction of additional variants not due to local mutation but is of course effective in this regard only if it introduces a rare variant; rough allowance for the possibility that migration has introduced a variant is made by scoring a specific variant only once (even if in several tribes). Genetic theory and observation suggest that variant departures from neutrality are much more apt to be characterized by loss than by gain in fitness; the assumption of neutrality in these calculations should thus bias the estimates of mutation downward. The two most obvious sources of bias are thus in opposite directions.

Thus far, these approaches have been applied to four different assemblages of tribal populations; the results are shown in Table 3. We specifically have excluded from this table the estimate of Chakraborty and Roychoudhury (20) for the Kadars of Kerala State, India, because the authors of the study on which the estimate is based (21) write "The postulated genetic reconstruction of the ancestral Kadar population suggests that they may have been similar to Melanesian and Australian aboriginal populations, but that this original genetic structure has been modified through incorporating genetic elements not only from black Africans but from surrounding Dravidian populations." This is scarcely the description of even a quasi-equilibrium population. We also exclude the (unusually low) estimate of these authors for five tribes of Andhra Pradesh because the primary electrophoretic data exist only in an inaccessible doctoral thesis and the tribes in question are known to have undergone a 4-to-5-fold population increase since 1891. It is obvious that there are considerable differences in the results given by the various approaches, as applied, and also among groups. The simple average across all populations and all three methods of estimating is 0.95×10^{-5} per locus per generation.

Thompson and Neel (5) have developed a treatment of the expected number of copies of a mutant allele at any number of

Table 3. Estimates of mutation rates in three different groups of tribal populations, based on an indirect approach

Group	Number of tribes	Estimate ($\times 10^{-5}$ per locus) per generation: Kimura–Ohta	Nei	Rothman–Adams	Average	Reference	Comments
Amerindians	12	1.43	1.69	1.71	1.61	(16, 18)	—
Amerindians	2	1.22				(19)	$\hat{t}_0$, calculated from formula $2\ln(2N) = 10.9$
Australian aborigines	?*	0.61	0.28	1.29	0.73	(7)	Current population approximately one-third that at time of first contact, despite recent increase
New Guinea natives	55	0.44	0.24	0.56	0.41	(8)	Tribe equated to language group

* This calculation based on total aboriginal sample.

generations following its origin, from which, given actual data on the numerical representation of each of the alleles of restricted distribution encountered in a survey of a defined tribal population, one can infer the mutation rate most consistent with the data set. The rate inferred from the application of this approach to the Amerindian data summarized by Neel (2) was 0.7×10^{-5} per locus per generation (5). Although it is not possible to assign an error to this estimate, it is consistent with the results of the indirect estimates on tribal populations.

Thus, in keeping with the difference in variant frequencies between the two types of populations, the (indirect) estimates of mutation rates for the tribal-tropical populations have given values that are higher than those excluded by the (direct) estimates on temperate-civilized populations. There is, however, marked variation among the indirect estimates on the tribal groups.

DISCUSSION

The possibility raised by the data given here, of higher mutation rates in certain tropical/tribal/nonindustrialized populations than in certain temperate-civilized/industrial populations, must be considered as tentative. There are, however, aspects of life in the tropics that could provide an explanation for these findings. For instance, the pressure of certain virus diseases, especially those due to the arboviruses, is heavier in the tropics, and at least one of them (yellow fever) is chromoclastic for cells in tissue culture (for review, see refs. 22 and 23). Agents that break chromosomes usually increase mutation rates. In this connection, we have reported evidence of unusual chromosomal damage in two villages of Yanomama Amerindians (24, 25). Contamination of foodstuffs by molds may be more common in the tropics; the mutagenic effect of mold-derived aflatoxin is well known (for review, see refs. 26, 27). Meat is often cooked by direct contact with the fire, in the process acquiring a thick char. Currently, there is growing evidence for the mutagenic potential of the products of protein pyrolysis (28, 29). Alternatively, ethnic differences in mutation rates cannot be excluded.

A test of this suggestion by the direct approach to mutation rates will probably not be feasible for logistic reasons: It will be extremely difficult to build up a series of sufficient size on relatively unacculturated tribal populations. In this connection, it must be recalled that the indirect approach draws on the accumulated mutagenic experience of many generations. Perhaps the best hope for a test lies in the development of better methods for the study of *in vivo* somatic cell mutation rates in the human, which rates presumably relate, with ratios yet to be developed, to the germinal rates reflected in the frequency of electrophoretic variants in populations. It will be of the utmost importance if such studies are undertaken to select subjects who are presumably living under the same conditions in which these higher frequencies of rare variants arose.

This research was supported by U.S. Department of Energy Contract DE-AC02-76EV02828.

1. Harris, H., Hopkinson, D. A. & Robson, E. B. (1974) *Ann. Hum. Genet.* **37**, 237–253.
2. Neel, J. V. (1978) *Am. J. Hum. Genet.* **30**, 465–490.
3. Neel, J. V. (1980) in *Population Structure and Genetic Disorders*, ed. Eriksson, A. (Academic, London), pp. 173–193.
4. Tanis, R., Ferrell, R. E., Neel, J. V. & Morrow, M. (1974) *Ann. Hum. Genet.* **38**, 179–190.
5. Thompson, E. A. & Neel, J. V. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1442–1445.
6. Neel, J. V. & Thompson, E. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1904–1908.
7. Bhatia, K. K., Blake, N. M. & Kirk, R. L. (1979) *Am. J. Hum. Genet.* **31**, 731–740.
8. Bhatia, K. K., Blake, N. M., Serjeantson, S. W. & Kirk, R. L. (1981) *Am. J. Hum. Genet.* **33**, 112–122.
9. Neel, J. V., Otake, M., Fujita, M., Satoh, C., Goriki, K., Nerishi, S., Hamilton, H. B., Kageoka, T. & Asakawa, J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4221–4225.
10. Neel, J. V., Mohrenweiser, H. M. & Meisler, M. H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6037–6041.
11. Ewens, W. J. (1972) *Theoret. Pop. Biol.* **3**, 87–113.
12. Abramowitz, M. & Stegun, I. A. (1966) *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, DC).
13. Kimura, M. & Ohta, T. (1969) *Genetics* **63**, 701–709.
14. Nei, M. (1977) *Am. J. Hum. Genet.* **29**, 225–232.
15. Rothman, E. D. & Adams, J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5094–5098.
16. Neel, J. V. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 3311–3315.
17. Li, F., Neel, J. V. & Rothman, E. (1978) *Am. Nat.* **112**, 83–96.
18. Neel, J. V. & Rothman, E. D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5585–5588.
19. Tchen, P., Séger, J., Bois, E., Grenaud, F., Fribourg-Blanc, A. & Feingold, N. (1978) *Hum. Genet.* **45**, 317–326.
20. Chakraborty, R. & Roychoudhury, A. K. (1978) *Hum. Genet.* **43**, 179–183.
21. Saha, N., Kirk, R. L., Shanbhag, S., Joshi, S. & Bhatia, H. (1974) *Hum. Hered.* **24**, 198–218.
22. Harndin, D. G. (1974) in *Chromosomes and Cancer*, ed. Germany, J. (Wiley, New York), pp. 151–190.
23. Sandberg, A. A. (1980) *The Chromosomes in Human Cancer and Leukemia* (Elsevier, New York).
24. Bloom, A. D., Neel, J. V., Choi, K. W., Iida, S. & Chagnon, N. A. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 920–927.
25. Bloom, A. D., Neel, J. V., Tsuchimoto, T. & Meilinger, K. (1973) *Cytogenet. Cell Genet.* **12**, 175–186.
26. Fishbein, L., Flamm, W. G. & Falk, H. L. (1970) *Chemical Mutagens* (Academic, New York).
27. Flamm, W. G. & Mehlman, M. A., eds. (1978) *Adv. Mod. Toxicol.* **5**, pp. xii and 401.
28. Sugimura, T., Nagao, M., Kawachi, T., Honda, M., Yahagi, T., Seino, Y., Sato, S., Matsukura, N., Matsushima, T., Shirai, A., Sawamura, M. & Matsumoto, H. (1977) in *Origins of Human Cancer*, eds. Hiatt, H. H., Watson, J. D. & Winsten, J. A. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 1561–1577.
29. Sugimura, T. (1978) *Mutat. Res.* **55**, 149–152.