A MIXED MODEL OF MUTATION FOR ELECTROPHORETIC IDENTITY OF PROTEINS WITHIN AND BETWEEN POPULATIONS[1]

WEN-HSIUNG LI

Center for Demographic and Population Genetics, University of Texas Health Science Center at Houston, Houston, Texas 77030

Manuscript received November 18, 1975
Revised copy received February 2, 1976

ABSTRACT

A model which is a mixture of the model of infinite alleles and the OHTA-KIMURA model of stepwise mutation has been proposed for the study of eletcrophoretic variants in natural populations. Mutations which alter the mobility of a protein are divided into two classes: stepwise mutations and nonstepwise mutations. It is assumed that stepwise mutations follow the OHTA-KIMURA model while nonstepwise mutations follow the infinite allele model. It is then shown that even if the proportion of nonstepwise mutations is only 5%, with the other 95% stepwise mutations, the effective number of alleles given by the present model is considerably larger than that given by the OHTA-KIMURA model of stepwise mutation. The result has also been applied to study NEI's genetic distance.

FOUR years before WATSON and CRICK (1953) proposed the double-helix model of the structure of DNA, WRIGHT (1949) had already used the so-called infinite-allele model to study the number of selectively neutral alleles that can be maintained in a population. Fifteen years later, KIMURA and CROW (1964) independently proposed the same model and gave an extensive mathematical analysis of the model. This model seems to be appropriate if allelic variants are identified at the nucleotide or amino acid level. However, more than a quarter of a century later, our experimental analyses of the genetic variability of natural populations are still at a much cruder level of identifying variants, mostly by electrophoresis. This situation has prompted OHTA and KIMURA (1973) to propose a model of stepwise mutation for the study of electrophoretic variants in natural populations. NEI and CHAKRABORTY (1973), prompted by KING (1973), have also proposed a similar model for the study of genetic distance between populations. Recently, the mathematical property of this model has been studied fairly extensively (e.g., EWENS and GILLESPIE 1974; OHTA and KIMURA 1974; WEHRHAHN 1975; BROWN, MARSHALL and ALBRECHT 1975; AVERY 1975; KIMURA and OHTA 1975; LI 1976a). However, as pointed out by JOHNSON (1974), this model may not be very realistic for some enzymes. OHTA and KIMURA assume that electrophoretic states can be denoted by integers on a line and an allele "can mutate only to one of the two adjacent states. One positive

[1] This study was supported by U.S. National Institutes of Health Research Grant 5-SO5RR-07148 to the Graduate School of Biomedical Sciences, University of Texas Health Science Center at Houston, Houston, Texas 77030.

and one negative change cancel each other, leading the allele back to the original state." Theoretically, positive and negative alterations in charge very often do not cancel evenly, due simply to differing levels of amino acid side-group ionization (JOHNSON 1974), although in practice the extent to which these differences can be detected is another problem. For instance, a fraction of a single charge difference seems to distinguish the mutants *A*58 and *A*46 of the *A* protein of the tryptophan synthetase of *E. coli* (HENNING and YANOFSKY 1963). Furthermore, there are cases where amino acid substitutions which are themselves electrophoretically neutral do result in mobility changes (WILTSHIRE *et al.* 1972; HUNT and DAYHOFF 1974). For example, Hemoglobin Denmark Hill $\alpha 95$ (G_2) Pro-Ala migrates more slowly towards the anode than Hb A on cellulose acetate electrophoresis at pH 8.6, although the substitution per se is neutral. It is therefore important to take the effect of fractional changes in mobility into account. In this study, I shall attempt to incorporate this effect into OHTA and KIMURA's model. I shall use the method of probability generating function (pgf), which has proved to be very powerful for this type of study (WEHRHAHN 1975; WATTERSON 1975; LI 1976a,b).

Electrophoretic Identity of Proteins

The purpose of this section is to study the probability of electrophoretic identity of two randomly chosen proteins. My model is a mixture of the model of infinite alleles and the model of stepwise mutation. I consider only mutations which result in mobility changes. These mutations are divided into two classes: stepwise mutations and nonstepwise mutations. In the first class, a mutation changes the mobility of a protein either one unit step in the positive direction or one step in the negative direction; one positive and one negative step mutations cancel each other completely. That is, mutations of this class follow OHTA and KIMURA's model of stepwise mutation. In the second class, a mutation changes the mobility of a protein a fraction of one unit step, either in the positive or in the negative direction; each mutation is unique and detectable. Mutations of this class follow the model of infinite alleles. It is noted that two proteins can be identical in state without being identical by descent with respect to the first class of mutations, but they can be identical only by descent with respect to the second class of mutations. Whether these assumptions are realistic or not will be discussed later.

I assume that the numbers of negative and positive step mutations occurring to a protein in each generation follow Poisson laws with mean u_{-1} and u_1 respectively, i.e., the probability generating function (pgf) for the distribution of the mobility changes in one generation is given by $\exp(u_{-1}x^{-1} + u_1x - u_{-1} - u_1)$. It can be shown that if two proteins are identical in state now, the distribution for their mobility differences with respect to the class of stepwise mutations in the next generation follow the distribution $\exp[(u_{-1} + u_1)x^{-1} + (u_{-1} + u)x - 2u_{-1} - 2u_1] = \exp(ux^{-1} + ux - 2u)$, where $u = u_{-1} + u_1$. (For a more detailed derivation, readers may refer to WEHRHAHN 1975). Similarly, I assume that in each generation the number of nonstepwise mutations occurring to a protein fol-

lows a Poisson law with mean v, i.e., the pgf for the distribution of this number is $e^{v(y-1)}$. (This is of course only a first approximation, since the relationship between mobility change and amino acid substitution is a very complicated matter (c.f. JOHNSON 1974).) The number of new nonstepwise mutations occurring to two proteins in one generation is obviously given by the distribution $e^{2v(y-1)}$. (Note that in regard to the effect of mutations of the second class on the probability of identity of two proteins I am concerned only with whether mutations have occurred to them or not, but not with the direction of mobility changes. Therefore, I do not consider the "direction" of mutation in this case.) I assume that stepwise mutations and nonstepwise mutations occur independently so that if two proteins are identical now, in the next generation the "differences in mobility" between them with respect to both classes of mutations is given by the distribution

$$\begin{aligned} g(x,y) &= e^{(ux^{-1}+ux-2u)} \cdot e^{2v(y-1)} \\ &= e^{-2U+ux^{-1}+ux+2vy} \end{aligned} \tag{1}$$

where $U = u + v$ is the total mean mutation rate.

Now consider a randomly mating population of constant effective size N. Let $P(i,k;t)$ be the probability that in generation t two randomly chosen proteins differ by i steps with respect to stepwise changes and by k nonstepwise changes, and let

$$G(x,y;t) = \sum_{i=-\infty}^{\infty} \sum_{k=0}^{\infty} P(i,k;t) x^i y^k$$

be its pgf. Note that in generation t "the differences in mobility" between two randomly chosen proteins follow the distribution $g(x,y)$ if the two genes which produce these two proteins were derived from replication of a gene in generation $t-1$, but follow the distribution $G(x,y;t-1)g(x,y)$ if they were derived from two genes in generation $t-1$. Thus

$$G(x,y;t) = \frac{g(x,y)}{2N} + cG(x,y;t-1)g(x,y) , \tag{2}$$

where $c = 1 - 1/2N$. The solution of (2) is given by

$$\begin{aligned} G(x,y;t) &= \frac{g(x,y)[1 - \{cg(x,y)\}^t]}{2N[1 - cg(x,y)]} + \{cg(x,y)\}^t G(x,y;0) \\ &\approx \frac{g(x,y)[1 - e^{a(x,y)t}]}{2N[1 - e^{a(x,y)}]} + G(x,y;0)e^{a(x,y)t} , \end{aligned} \tag{3}$$

where $a(x,y) = -\lambda + ux^{-1} + ux + 2vy$ and $\lambda = 1/(2N) + 2U$. In the following, I shall consider only the probability of identity, i.e., homozygosity $J(t)$. As mentioned earlier, for two proteins to be identical in mobility, there should be no difference between them with respect to the second class of mutations, therefore, $y = 0$. On the other hand, the expansion of (3) in infinite series in terms of x

will be done around $x = 1$. It follows that $g(x,0) = 1$ and $e^{a(x,0)} - 1 = a(x,0)$ approximately. Thus,

$$G(x,0;t)) = \frac{-1 + e^{a(x,0)t}}{2Na(x,0)} + G(x,0;0)e^{a(x,0)t}, \tag{4}$$

approximately. Note that at the steady state,

$$G(x,0;\infty) = \frac{-1}{2Na(x,0)} = \sum_{i=-\infty}^{\infty} \frac{x_1^{|i|}}{\sqrt{d^2 - \alpha^2}} x^i, \tag{5}$$

that is,

$$P(i,0;\infty = \frac{x_1^{|i|}}{\sqrt{d^2 - \alpha^2}},$$

where $d = 1 + \theta = 1 + \alpha + \beta$, $\theta = 4NU$, $\alpha = 4Nu$, $\beta = 4Nv$ and $x_1 = (d - \sqrt{d^2 - \alpha^2})/\alpha$. Therefore, the homozygosity at the steady state is

$$J(\infty) \equiv P(0,0;\infty) = \frac{1}{\sqrt{(1+\theta)^2 - \alpha^2}}. \tag{6}$$

It is also easy to see that

$$\frac{e^{a(x,0)t}}{2Na(x,0)} = -G(x,0;\infty)e^{a(x,0)t} = e^{-\lambda t} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I_j(2ut)P(j-i,0;\infty)x^i, \tag{7}$$

$$G(x,0;0)e^{a(x,0)t} = e^{-\lambda t} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I_j(2ut)P(j-i,0;\infty)x^i; \tag{8}$$

where $I_j(x)$ is a modified Bessel function of the first kind and defined as $I_j(2ut) = \sum_{s=0}^{\infty} (ut)^{2s+j}/s!(s+j)!$ if $j \geq 0$ and $I_j(x) = I_{-j}(x)$ if $j < 0$. From (4), (6), (7) and (8), it follows that

$$J(t) = J(\infty) + e^{\lambda - t} \sum_{j=-\infty}^{\infty} I_j(2ut)[P(j,0;0) - P(j,0;\infty)]. \tag{9}$$

When $v = 0$, (9) reduces to (11) of LI (1976a). The recursion relationship $I_{j+1}(x) = -2jI_j(x)/x + I_{j-1}(x)$ makes numerical computations of (9) fairly easy, though caution should be taken against rounding errors. When ut is small, $I_j(2ut) \approx (ut)^j/j!$, $j \geq 0$, while when ut is large $e^{-\lambda t}I_j(2ut)$ is small. Therefore, (9) may be approximately given by

$$J(t) = J(\infty) + e^{-\lambda t}[J(0) - J(\infty)]. \tag{10}$$

The result of LI (1976a) indicates that this approximation formula holds rather well.

The heterozygosity at generation t is given by $1 - J(t)$ while the effective number of alleles is given by the reciprocal of $J(t)$. In particular, the effective number of alleles at the steady state is given by

$$n_e = \sqrt{(1+\theta)^2 - \alpha^2} \tag{11}$$

$$= \sqrt{(1+\beta)^2 + 2(1+\beta)\alpha}. \tag{11'}$$

Note that if only stepwise mutations are possible, then $v = 0$, $\beta = 0$, $u = U$, $\alpha = \theta$ and (11') reduces to

$$n_e = \sqrt{1 + 2\theta}$$

which is identical with formula (9) of OHTA and KIMURA (1973). On the other hand, if all mutations are nonstepwise, then $u = 0$, $\alpha = 0$, $v = U$, $\beta = \theta$, and (11) reduces to

$$n_e = 1 + \theta$$

which is identical with the result of KIMURA and CROW (1964). Therefore, the present model includes the infinite-allele model and the stepwise model as special cases. Figure 1 shows the effective number of alleles maintained in an equilibrium population for various values of NU, where $U = u + v$ is the total mutation rate. The upper limit of this number is given by the line for $v = U$, which corresponds to the model of infinite alleles, while the lower limit is given by the line for

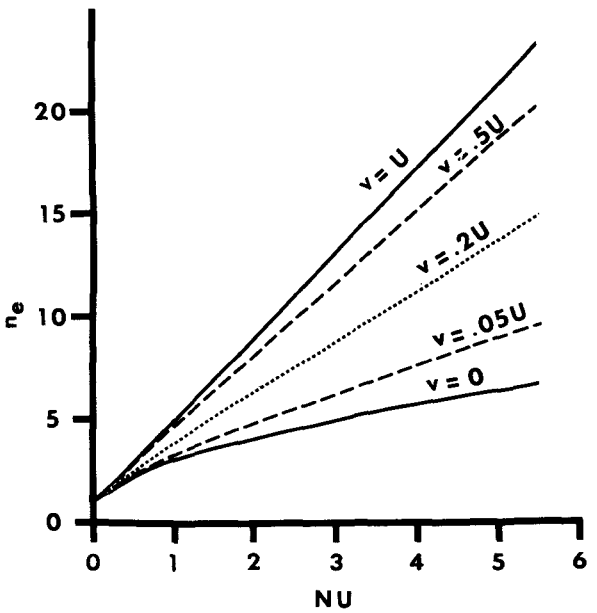

FIGURE 1.—Relationship between the effective number of alleles (n_e) and NU under the infinite-allele model ($v = U$), the present model, and the OHTA-KIMURA model ($v = 0$). $U = u + v$ is the total rate of mutations involving mobility changes, u the rate of stepwise mutations, and v the rate of nonstepwise mutations.

$v = 0$, which corresponds to the O-K (OHTA-KIMURA) model of stepwise mutation. The line for $v = .5U$ is the case where 50% of the mutations are nonstepwise and the other 50% are stepwise. Note that this line is very close to the line for $v = U$, i.e., the model of infinite alleles. The line for $v = .2U$ is the case where 20% of the mutations are nonstepwise. This line is seen to be closer to the line for $v = U$ rather than to the line for $v = 0$. Even the line for $v = .05U$, that is only 5 percent of the mutations are nonstepwise, is quite different from the line for $v = 0$, the O-K model of stepwise mutation.

Nei's Genetic Distance

NEI and CHAKRABORTY (1973) and LI (1976a) have studied the genetic distance under the model of stepwise mutation. In the following, I shall study the genetic distance under the present model. Let a_j and b_j be the frequencies of the jth allele A_j in populations 1 and 2, respectively. NEI's (1972) genetic distance is defined as

$$D = -\log_e(J_{12}/\sqrt{J_1 J_2}), \tag{12}$$

where J_1, J_2 and J_{12} are the average of Σa_j^2, Σb_j^2, and $\Sigma a_j b_j$ over all loci, respectively (or the expectations at a locus).

Now suppose that at $t = 0$ a population splits into two populations and thereafter no migration occurs between them (the effect of migration will be considered in another study). Since each population evolves independently, J_1 and J_2 can be computed from the formulas given in the previous section while J_{12} can be computed as follows. Let $W(x,y) = \sum_{i=-\infty}^{\infty} \sum_{k=0}^{\infty} Q(i,k) x^i y^k$ be the pgf for the distribution of "mobility differences" for the ancestral population at the time of divergence and $E(i,k;t)$ be the probability that at generation t two randomly chosen proteins, one from each population, differ by i steps with respect to the class of stepwise changes and by k nonstepwise changes. The pgf of $E(i,k;t)$'s is

$$\begin{aligned} B(x,y;t) &\equiv \sum_{i=-\infty}^{\infty} \sum_{k=0}^{\infty} E(i,k;t) x^i y^k \\ &= W(x,y) e^{(-2u-2v+ux^{-1}+ux+2vy)t}. \end{aligned} \tag{13}$$

It is not difficult to see that

$$J_{12}(t) \equiv E(0,0;t) = e^{-2(u+v)t} \sum_{i=-\infty}^{\infty} I_i(2ut) Q(i,0). \tag{14}$$

The simplest case for consideration is that where the sizes of the ancestral and the two descendant populations are more or less the same and the ancestral population was at the steady state at the time of separation. In this case it may be assumed that

$$J_1 = J_2 = P(0,0;\infty) = \frac{1}{\sqrt{d^2 - \alpha^2}}$$

and

$$Q(i,0) = \frac{x_1^{|i|}}{\sqrt{d^2 - \alpha^2}}$$

it then follows that

$$J_{12}(t) = e^{-2(u+v)t} \sum_{i=-\infty}^{\infty} I_i(2ut) \frac{x_1^{|i|}}{\sqrt{d^2 - \alpha^2}}, \tag{15}$$

$$D(t) = 2vt + 2ut - \log_e \sum_{i=-\infty}^{\infty} x_1^{|i|} I_i(2ut). \tag{16}$$

When $v = 0$, these formulas reduce to those of Li (1976a). Table 1 shows the D values for the infinite allele model, the present model, and the O-K model. Under the above assumptions $D = 2Ut$ in the case of the model of infinite alleles (Nei and Feldman 1972). This value represents the expected number of amino acid substitutions which occurred in either population, each substitution resulting in mobility change. Note that if $4NU = 0.1$, the increase in detectability due to nonstepwise mutations is small if $v \leq 0.2U$ and $2Ut$ is 0.2 or less, although it becomes appreciable when $2Ut$ is close to 2. On the other hand, if $4NU = 1$, the effect of nonstepwise mutations becomes stronger. The use of electrophoretic data for estimating the genetic distance between populations is not recommended when $2Ut$ becomes larger than two (Nei and Chakraborty 1973; Li 1976a). The present result seems to confirm this point. The result of Nei and Chakraborty (1973) is not compared with the present result because they studied the quantity $D = -\log_e J_{12}$ with the initial condition $J_{12}(0) = 1$.

DISCUSSION

In the above formulation, it has been assumed that in the class of stepwise mutations only one-step mutations can occur. If one takes two-step mutations into consideration, the homozygosity at the steady state is given by

$$J(\infty) = \frac{2}{(\alpha_1 + 2\alpha_2 W_1)\sqrt{W_1^2 - 4}} - \frac{2}{(\alpha_1 + 2\alpha_2 W_2)\sqrt{W_2^2 - 4}}, \tag{17}$$

TABLE 1

Genetic distance under the infinite-allele model, the present model, and the Ohta-Kimura *model*

Generation		10^3	10^5	10^6	10^7	10^8
$U = 10^{-7}$, $\theta = .1$	Inf.: $2Ut$	.000200	.0200	.200	2.00	20.0
	$v = .2U$, $u = .8U$	.000194	.0194	.188	1.40	6.22
	$v = .1U$, $u = .9U$	.000193	.0192	.185	1.26	4.28
	$v = .05U$, $u = .95U$	.000192	.0191	.183	1.19	3.30
	O-K ($v = 0$)	.000191	.0190	.181	1.11	2.32
$U = 10^{-6}$, $\theta = 1$	Inf.: $2Ut$	.00200	.200	2.00	20.0	200.0
	$v = .2U$, $u = .8U$	.00166	.161	1.19	5.89	43.0
	$v = .1U$, $u = .9U$	.00157	.150	1.01	3.89	23.0
	$v = .05U$, $u = .95U$	.00152	.144	.92	2.89	13.0
	O-K ($v = 0$)	.00146	.137	.82	1.88	3.02

Inf. and O-K refer to the model of infinite alleles and Ohta and Kimura's model of stepwise mutation. $U = u + v$ is the total rate of mutations involving mobility changes, u the rate of stepwise mutations, v the rate of nonstepwise mutations. $\theta = 4NU$.

where

$$W_1, W_2 = \frac{-\alpha_1 \pm \sqrt{\alpha_1^2 + 8\alpha_2(1 + \theta + \alpha_2)}}{2\alpha_2},$$

$\theta = 4NU$, $\alpha_1 = 4Nu_1$, $\alpha_2 = 4Nu_2$, and $U = u_1 + u_2 + v$, in which u_1 and u_2 are the rates of one-step and two-step mutations, respectively, and v is the rate of nonstepwise mutations. It can be shown that formula (17) is practically the same as formula (6), since the amount of two-step mutations seems to be somewhat less than 10 percent of the one-step mutations (NEI and CHAKRABORTY 1973). For example, if $\theta = 20$, $\alpha = 16$, and $\beta = 4$, then formula (11) implies $n_e = 13.6$ while (17) gives $n_e = 14.2$ if $\theta = 20$, $\alpha_1 = 14.4$, $\alpha_2 = 1.6$, and $\beta = 4$.

The present results indicate that the effective number of alleles is highly dependent on the proportion of nonstepwise mutations (Figure 1). This proportion seems to vary from protein to protein and depends on the buffer pH value (JOHNSON 1974). It has also been noted that, even for a given protein, the value of pK_a for a given amino acid side group may vary widely, depending upon its location in the protein (HENNING and YANOFSKY 1963; JOHNSON 1974). As HENNING and YANOFSKY argued, it would be possible to observe different mobilities with changes involving the same amino acid at different positions in the protein. Furthermore, different geometries can exhibit different mobilities in a gel and change in conformation may alter the mobility of the protein molecule by exposure of charged side chains that previously did not contribute to the charge of the protein (HENNING and YANOFSKY 1963; JOHNSON 1974). For example, as mentioned earlier, electrophoretically "neutral" substitutions may sometimes result in mobility changes (WILTSHIRE *et al.* 1972; HUNT and DAYHOFF 1974). Therefore, it seems reasonable to conclude that at least an appreciable proportion of mutations will result in fractional mobility changes, although it is almost impossible to determine the exact amount.

It may be criticized that the assumption that nonstepwise mutations never cancel completely and are unique and detectable is not very realistic. The net effect of this assumption will tend to increase the effective number of alleles and genetic distance. The most realistic approach would be to consider the distribution of mobility changes and study the mobility changes in terms of the smallest unit that one can read. However, this approach is obviously very complicated. The simplest way to compensate for this inflation is to increase the proportion of stepwise mutations. But, at all events, the great differences between results of the present model and OHTA and KIMURA's model should be worth noting.

All the above arguments are based on theoretical considerations. DR. W. FITCH has noted (personal communication) that in practice which model should be used depends on the care in which the data are obtained. If one reads the mobility differences very crudely, perhaps the model of stepwise mutation may be realistic enough. On the other hand, if one reads very carefully, the effect of fractional changes should be considered. However, it should be noted that even if the amount of detected fractional changes is only five percent of the total mutations

involving mobility changes, the effective number of alleles may be considerably larger than that given by OHTA and KIMURA's model, as indicated by Figure 1. In view of the present result, it should be clear that a small increase in detectability by combining a technique such as heat denaturation treatment with electrophoresis (BERNSTEIN, THROCKMORTON and HUBBY 1973) may reveal a considerably larger amount of genetic variability.

I am greatly indebted to DRS. M. NEI, W. FITCH, J. F. CROW and L. H. THROCKMORTON for discussions and valuable suggestions.

LITERATURE CITED

AVERY, P. J., 1975 Extension to the model of an infinite number of selectively neutral alleles in a finite population. Genet. Res. **25**: 145–153.

BERNSTEIN, S. C., L. H. THROCKMORTON and J. L. HUBBY, 1973 Still more genetic variability in natural populations. Proc. Nat. Acad. Sci. **70**: 3928–3931.

BROWN, A. D. H., D. H. MARSHALL and L. ALBRECHT, 1975 Profile of electrophoretic alleles in natural populations. Genet. Res. **25**: 137–143.

EWENS, W. J. and J. H. GILLESPIE, 1974 Some simulation results for the neutral allele model, with interpretations. Theoret. Pop. Biol. **6**: 35–57.

HENNING, U. and C. YANOFSKY, 1963 An electrophoretic study of mutationally altered A proteins of the tryptophan synthetase of *Escherichia coli*. J. Mol. Biol. **6**: 16–21.

HUNT, L. T. and M. O. DAYHOFF, 1974 Table of abnormal human globins with references. In: *Hemoglobin: Comparative Molecular Biology Models for the Study of Disease*. Edited by H. KITCHEN and S. BOYER. Ann. N.Y. Acad. Sci., Vol. **241**.

JOHNSON, G. B., 1974 On the estimation of effective number of alleles from electrophoretic data. Genetics **78**: 771–776.

KIMURA, M. and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49**: 725–738.

KIMURA, M. and T. OHTA, 1975 Distribution of allelic frequencies in a finite population under stepwise production of neutral alleles. Proc. Nat. Acad. Sci. **72**: 2761–2764.

KING, J. L., 1973 The probability of electrophoretic identity of proteins as a function of amino acid divergence. J. Mol. Evol. **2**: 317–322.

LI, W-H., 1976a Electrophoretic identity of proteins in a finite population and genetic distances between taxa. (Submitted for publication). ——, 1976b Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. (Submitted for publication).

NEI, M., 1972 Genetic distance between populations. Am. Nat. **106**: 283–292.

NEI, M. and R. CHAKRABORTY, 1973 Genetic distance and electrophoretic identity of proteins between taxa. J. Mol. Evol. **2**: 323–328.

NEI, M. and M. W. FELDMAN, 1972 Identity of genes by descent within and between populations under mutation and migration pressures. Theoret. Pop. Biol. **3**: 460–465.

OTHA, T. and M. KIMURA, 1973 A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet. Res. **22**: 201–204. ——, 1974 Simulation studies on electrophoretically detectable genetic variability in a finite population. Genetics **76**: 615–624.

WATSON, J. D. and F. C. CRICK, 1953 Genetical implication of the structure of deoxyribose nucleic acid. Nature **171**: 964.

WATTERSON, G. A., 1975 On the number of segregating sites in genetical models without recombination. Theoret. Pop. Biol. **7**: 256–276.

WEHRHAHN, C. F., 1975 The evolution of selectively similar electrophoretically detectable alleles in finite natural populations. Genetics **80**: 375–394.

WILTSHIRE, B. G., K. G. A. CLARK, P. A. LORKIN and H. LEHMAN, 1972 Haemoglobin Denmark Hill $\alpha 95$ (G_2) Pro-Ala, a variant with unusual electrophoretic and oxygen-binding properties. Biochim. Biophys. Acta **278**: 459–464.

WRIGHT, S., 1949 Genetics of populations. Encyclopaedia Britannica, 14 ed. **10**: 111, 111A-D, 112.

Corresponding editor: W. J. EWENS