*Proc. Natl. Acad. Sci. USA*
Vol. 81, pp. 6073–6077, October 1984
Evolution

# Group selection for a polygenic behavioral trait: Estimating the degree of population subdivision

**(altruism/kin selection/population structure/*F* statistics/polygenic inheritance)**

JAMES F. CROW* AND KENICHI AOKI†

*Laboratory of Genetics, University of Wisconsin, Madison, WI 53706; and †National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan

*Contributed by James F. Crow, June 14, 1984*

ABSTRACT **For assessing the degree of population subdivision, and therefore the extent to which group selection might favor an altruistic trait, an appropriate measure is Nei's $G_{ST}$, defined by $(F_0 - \bar{F})/(1 - \bar{F})$. $F_0$ is the probability that two alleles drawn from the same group are identical in state and $\bar{F}$ is the probability for two alleles drawn at random from the entire population. These probabilities can be assessed from molecular polymorphisms. $G_{ST}$ has a number of properties that make it useful for empirical studies. When the mutation rate is small relative to the migration rate and the reciprocal of the group size, $G_{ST}$ depends mainly on the absolute number of migrants per generation, moves rapidly to near equilibrium, and is independent of the number of alleles. The relative homogenizing effect of migration in the island and stepping-stone models is not as different as might be expected; one immigrant chosen randomly from the rest of the population is only one to two times as effective as one from a neighboring group, appreciably exceeding 2 only when there are 1000 or more groups. The use of molecular data to estimate the degree of population subdivision may permit testable predictions of the extent of altruistic behavior.**

In an earlier article (1) we noted that an additively polygenic trait that is deleterious to the individual but beneficial to the group will increase in frequency if

$$-B_w/B_b < 2F/(1 - F). \qquad [1]$$

$B_b$ and $B_w$ are the between- and within-group genetic regressions of fitness on the value of the trait; in the case of interest $B_b > 0$ and $B_w < 0$. $F$ is Wright's fixation index, $F_{ST}$ (2). If the population is at equilibrium between migration and random drift

$$(1 - F)/2F \approx 2M_e, \qquad [2]$$

where $M_e$ is the effective number of migrants exchanged each generation between a group and rest of the population (1). If the sex ratio of the group or of the immigrants is 1, $M_e$ is the absolute number of migrants, $M$. If the sex ratio of immigrants deviates from equality in the same direction as that of the group, $M_e < M$; if they deviate in opposite directions, $M_e > M$. If $c$ (= $-B_w$) is the expected cost in reduced fitness per unit change in the value of the character and $b$ (= $B_b - B_w$) is the increase in group benefit per unit change in the average value of the character, then Eq. **1** becomes

$$c/b < 2F/(1 + F) = r. \qquad [3]$$

This is the familiar Hamilton cost/benefit inequality (3) and $r$ is Wright's coefficient of relationship, a measure of shared genes. Very similar conclusions have been reached by Leigh (4), who employed an entirely different procedure. We assumed that group selection occurs by differential increase and decrease in group sizes. Eq. **1** is only slightly modified if the mechanism of group selection is extinction balanced by fission of larger groups (5). If the selection occurs over long time periods, both processes are involved. The evolution of reciprocal altruism by kin or group selection can be analyzed by using the same approach (6).

The more the population is subdivided the greater is the ratio of the between- to the within-group variance and the larger $-B_w/B_b$ (or $c/b$) can be and still meet the conditions for the altruistic trait to increase. Eq. **1** tells us that the present extent of population subdivision can be determined by measuring $F$. It could also be determined from Eq. **2** by observing the number of migrants and assuming that this pattern of migration has existed long enough for the population to be at equilibrium.

The purpose of this note is to discuss more fully two points: First, what is the appropriate definition of $F$ and how might it be measured? In particular, how can molecular data be used to assess the degree of population subdivision? Second, if migration data are available, what are the consequences when the migrants come from nearby groups rather than at random from the rest of the population, as was assumed in the derivation of Eq. **2**?

## Estimation of *F*

Wright's $F_{ST}$ (2) measures the genetic differentiation of a group relative to the population as a whole. This could be measured from pedigrees if there were sufficient information, but this is unrealistic in most, if not all, situations. A more practical possibility is to use data on protein or DNA polymorphism to estimate allelic identity.

We are assuming that the molecular markers are near enough to selective neutrality that they can be used to assess the degree of group differentiation brought about by migration and random drift. Likewise, we are assuming that the altruistic trait, which need not be behavioral, is polygenic; the number of loci determining the trait is assumed to be large enough that selection on any one locus is very weak and selection for the trait does not change the genetic variance. Under these circumstances molecular data can be used to assess the relative opportunity for group vs. individual selection.

Let $F_0$ be the probability that two homologous genes drawn at random from a group are the same allele (identical in state) and let $\bar{F}$ be the probability for genes drawn at random from the entire population. Then, following Nei (7, 8) we define $G_{ST}$ as

$$G_{ST} = (F_0 - \bar{F})/(1 - \bar{F}). \qquad [4]$$

$G_{ST}$ and $F_{ST}$ are the same when there are only two alleles (ref. 8, p. 151). $G_{ST}$ is preferable for our purposes since it

---

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

lends itself to specific models of mutation and migration.

Recurrence relations for these quantities are readily derived if we assume that all groups are the same size and that the mutation rates from any allele to any other are the same. Let $m$ be the proportion of genes in a group that are exchanged each generation for genes chosen at random from the rest of the population. $N$ is the number of individuals in a group and $n$ is the number of groups. The mutation rate from an allele to a particular allele is $u/(K-1)$, where $K$ is the number of alleles. Then the total rate of mutation from one allele to all others is $u$. $F_1$ is the probability that two alleles drawn from two different groups are the same.

Letting primes stand for the next generation, we have

$$F_0' = v[a\{c + (1-c)F_0\} + (1-a)F_1] + w[a(1-c)(1-F_0) + (1-a)(1-F_1)] \quad [5a]$$

$$F_1' = v[b\{c + (1-c)F_0\} + (1-b)F_1] + w[b(1-c)(1-F_0) + (1-b)(1-F_1)] \quad [5b]$$

$$\bar{F} = [F_0 + (n-1)F_1]/n, \quad [5c]$$

where

$$v = (1-u)^2 + u^2/(K-1);$$

$$w = \frac{2u(1-u)}{K-1} + \frac{u^2(K-2)}{(K-1)^2} = (1-v)/(K-1)$$

$$a = (1-m)^2 + m^2/(n-1);$$

$$b = (1-a)/(n-1); \; c = 1/2N.$$

The last term in the expression for $v$ is the probability of both alleles mutating to the same allele. Likewise, in the expression for $w$, the first term is the probability that one of two chosen alleles mutates to the other; the second term is the probability of both mutating to a third type. The derivations of $a$ and $b$ are analogous.

We can obtain the exact equilibrium values by dropping the primes and solving the resulting equations. Our interest is in the case where $u << m, 1/N << 1$. Under these conditions (i.e., neglecting terms of order $u^2$, $um$, $u/N$, $m^2$, and $m/N$) the approximate solutions of interest are

$$F_0 \approx \frac{m + uA(n-1) + 4NmnuB}{m + 4NmnuA + (n-1)uA} \quad [6a]$$

$$\bar{F} \approx \frac{nm + 4Nmn^2uB + (n-1)uA + (n-1)^2uB}{n[m + 4NmnuA + (n-1)uA]}, \quad [6b]$$

where

$$A = K/(K-1); \qquad B = 1/(K-1).$$

When $A = 1$ and $B = 0$ we have the infinite allele model. In this case, Eqs. 6 reduce to Nagylaki's (9) equations 70 and 71. We have derived our equations on the assumption of haploid rather than diploid migrants, but Nagylaki (9) has shown that under many circumstances the haploid equations provide a good approximation.

Substituting Eqs. 6 into Eq. 4 yields

$$G_{ST} \approx \frac{1}{4Nm\alpha + 1} \quad [7]$$

$$\alpha = [n/(n-1)]^2$$

with the pleasing result that this approximation is independent of the mutation rate, $u$, and the number of alleles, $K$. Actually, Eq. 7 is correct even if terms of order $m^2$ and $m/N$ are retained in the approximation.

The infinite allele model has been studied extensively. Eq. 7 for this model was given by Takahata (10). An exact solution was given earlier (ref. 8, equation 5.108); Nei's formula appears different, but only because his $m$ is defined assuming that immigrants come from the entire population, not just from the other groups. A complete solution for the infinite allele island model was found by Li (11).

When $n$ is large Eq. 7 reduces to the familiar form first given by Wright (12)

$$F_{ST} \approx \frac{1}{4Nm+1} = \frac{1}{4M+1}. \quad [8]$$

### Useful properties of $G_{ST}$

From Eqs. 6 we see that the equilibrium values of $F_0$ and $\bar{F}$ depend on all five parameters. On the other hand, Eq. 7 tells us that (to the order of approximation appropriate to the assumption $u << m, 1/N << 1$) the equilibrium value of $G_{ST}$ is independent of the mutation rate and the number of alleles. Furthermore, the dependence on the number of groups, $n$, is very weak unless $n$ is very small. These properties greatly enhance the utility of $G_{ST}$ for our purpose; we need not be concerned with the mutation rate, number of possible alleles, or the total population size. These are quantities that are usually unknown.

If there is any practical utility in assuming that $G_{ST}$ is near its equilibrium value, the equilibrium must be attained rapidly. $F_0$ and $\bar{F}$ do not have this property. Consider first a panmictic population of size $Nn$, and let $X$ stand for the difference between $F$ and its equilibrium value. Then $X' = X(1-c)(v-w)$. Thus, the time to go half way to equilibrium is approximately $(\ln 2)/[2uK/(K-1) + 1/2Nn]$. Unequal mutation rates among the alleles reduce the rate of approach (13). For the finite island model, the maximum eigenvalues are approximately $1 - 2uK/(K-1) - 2m/(n-1)$ if $m << 1/N << 1$ and $1 - 2uK/(K-1) - 1/2Nn$ if $1/N << m << 1$; so the process can be very slow.

On the other hand, $G_{ST}$ goes rapidly almost all of the way to equilibrium, even though $F_0$ and $\bar{F}$ may be changing very slowly and the entire structure approaches equilibrium at rates ultimately determined by the maximum eigenvalue (14). We can get a rough idea of the early change of $G_{ST}$ by using Eqs. 4 and 5 and assuming $u$ is very small and $n$ large. Letting $Y$ be the difference between $G_{ST}$ and its equilibrium value, $Y' \approx Y(1-2m)(1-1/2N)$. The time for $G_{ST}$ to go half way to equilibrium is approximately $(\ln 2)/(2m + 1/2N)$, regardless of allele number. That the value quickly goes almost to equilibrium has been verified by numerical calculations. From numerical examples it also appears that the time is quite insensitive to $n$, although it is less when $n$ becomes very small. So a second useful property of $G_{ST}$ is that it rapidly moves most of the way to equilibrium, at a rate mainly determined by $2m + 1/2N$, even though the final approach is slow.

A third useful property of $G_{ST}$ is that the two parameters $N$ and $m$ enter the equation only as their product, $M$. The equilibrium structure does not depend on the group size, but, except for weak dependence on $n$, only on the absolute number of migrants exchanged per generation, $M$. Thus, groups can contract and expand without changing $G_{ST}$, provided $M$ remains constant. Such change in group size is required, of course, if there is to be any opportunity for group selection by differential growth rates (1).

## Other migration patterns

The island model is often unrealistic. It is more likely that migrants come from nearby groups rather than being randomly chosen. An extreme alternative is the two-dimensional stepping-stone model (15, 16). The groups are regarded as being arranged in a rectangular lattice, with migration occurring only between adjacent colonies. Exact formulae for a finite population have not been found. However, if the opposite edges of the rectangle are brought together, the resulting figure is an abstract torus whose properties are known (17–19). Except for edge effects this should mimic a rectangular plane.

Explicit formulae are not illuminating so we resorted to numerical calculations, which showed that the equilibrium value of $G_{ST}$ is nearly independent of the mutation rate as in the island model. Table 1 shows some representative values. In contrast to the island model, as might be expected, $G_{ST}$ is dependent on the total population size. The dependence is weak, however; in the example as the number of groups changes from 9 to 40,000 the value of $G_{ST}$ changes only by a factor of 2.6. Fig. 1 shows that $G_{ST}$ is roughly proportional to log $n$.

The rapid rate of approach to near-equilibrium of $G_{ST}$ was also studied numerically for the stepping-stone model. The results are not shown here but were very similar to those of the island model except when the total population is very large. We believe that these desirable properties of $G_{ST}$ apply to a wide variety of migration patterns, since the island and stepping-stone models represent opposite extremes and most populations probably lie somewhere between them.

Fig. 2 illustrates another property of stepping-stone models. If the area is narrow the situation approaches that of a linear habitat, and it is well known that this leads to much greater group divergence than an area model. The values in Fig. 2 are thought to be typical of many primate populations. If the length of the habitat is much greater than its width, $G_{ST}$ can increase substantially.

## Application to primate data

$G_{ST}$ can be measured from protein and DNA polymorphism and we can expect increasingly extensive and accurate data in the near future. Aoki and Nozawa (20) have summarized the relevant data from the Japanese monkey (*Macaca fuscata*). From polymorphic variation at 12 blood protein loci in a sample of about 1500 individuals distributed over 33 troops they estimated an average $G_{ST}$ of 0.0935, corresponding to a coefficient of relationship, *r*, of 0.167. The values of *r* from five other primate species lie mostly in the range 0.10–0.15. Aoki and Nozawa suggest that these values may also be appropriate for Pleistocene man.

## Effect of nonrandom migration on the effectiveness of migration

Equation 7 relates the equilibrium value of $G_{ST}$ to the absolute migration rate between groups using the island model. With the stepping-stone model a migrant has less effect than in the island model because of the correlation in allele frequencies in adjacent groups. In our earlier paper (1) we suggested that a migrant from an adjacent group rather than from the entire population has its effectiveness in reducing group differences reduced by a factor of $-\pi/\ln 2u$, where $u$ is the mutation rate. If $u = 10^{-6}$ the factor is 4.2; if $10^{-5}$, it is 3.4. This formula, which comes from Malécot (21), assumes that the total population is infinite. A more accurate expression was given by Kimura and Weiss (22). We now look at finite populations and will see that reduction is typically considerably less than when the population is infinite.

Table 1. Examples showing that, while the equilibrium values of $F_0$ and $\bar{F}$ depend on the mutation rate, $G_{ST}$ is nearly independent of it

| Number of groups | Group size | Number of migrants | Mutation rate | Equilibrium value $F_0$ | $\bar{F}$ | $G_{ST}$ |
|---|---|---|---|---|---|---|
| 25 | 20 | 2.0 | $10^{-7}$ | 1.000 | 1.000 | 0.120 |
| | | | $10^{-6}$ | 0.998 | 0.998 | 0.120 |
| | | | $10^{-5}$ | 0.980 | 0.980 | 0.120 |
| | | | $10^{-4}$ | 0.837 | 0.815 | 0.120 |
| | | | $10^{-3}$ | 0.388 | 0.306 | 0.118 |
| 225 | 30 | 0.3 | $10^{-7}$ | 0.997 | 0.993 | 0.613 |
| | | | $10^{-6}$ | 0.975 | 0.935 | 0.613 |
| | | | $10^{-5}$ | 0.841 | 0.590 | 0.612 |
| | | | $10^{-4}$ | 0.649 | 0.130 | 0.597 |
| | | | $10^{-3}$ | 0.526 | 0.017 | 0.517 |
| 9 | 20 | 1.0 | $10^{-5}$ | 0.993 | 0.991 | 0.172 |
| 25 | | | | 0.981 | 0.974 | 0.223 |
| 100 | | | | 0.928 | 0.899 | 0.283 |
| 400 | | | | 0.784 | 0.675 | 0.337 |
| 2,500 | | | | 0.534 | 0.233 | 0.393 |
| 10,000 | | | | 0.463 | 0.067 | 0.425 |
| 40,000 | | | | 0.453 | 0.017 | 0.443 |

The lower part shows the weak dependence of $G_{ST}$ on the number of groups. The numbers were calculated from a stepping-stone model for an abstract torroidal surface.

We compared the two-dimensional stepping-stone model, using the abstract torus formulation, with the island model for a large number of parameter values. We have computed the number of migrants from the four adjacent groups that is equivalent to one migrant chosen at random from the rest of the population. As expected, this number ranges from near one when the number of groups is small to the Malécot limit as the number of groups increases. Some representative values are shown in Fig. 3. Unless the number of groups is enor-

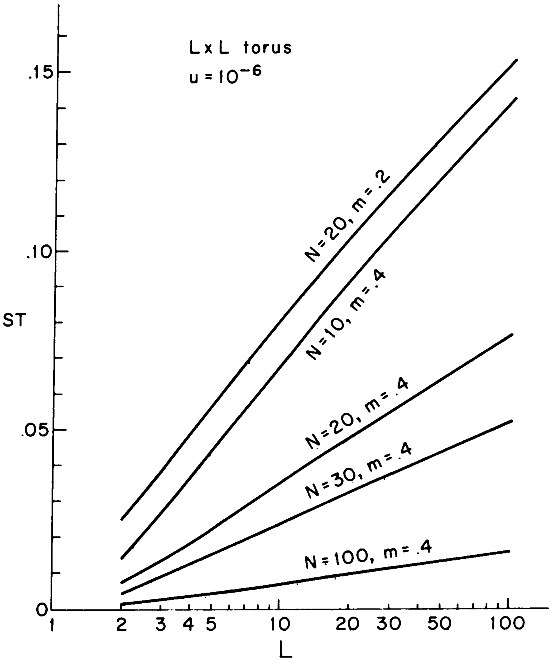

FIG. 1. Equilibrium value of $G_{ST}$ as a function of number of groups in the stepping-stone model. $G_{ST}$ increases approximately as the logarithm of the number of groups. The parameters are thought to be reasonable for many primates. $L \times L$ (=$n$) is the total number of groups in the population; $N$ is the number of individuals per group; $m$ is the migration rate.

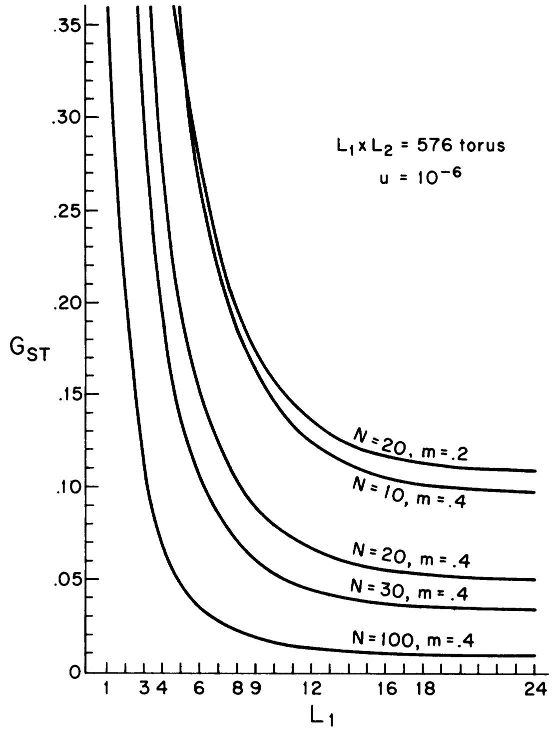

FIG. 2. Effect of habitat shape on the equilibrium value of $G_{ST}$. As the habitat becomes long and narrow $G_{ST}$ increases. $L_1 \times L_2$ is the total number of groups in the rectangular torus.

mous these are in the range 1–2. This is also reflected in calculations of the time for $G_{ST}$ to go half way to equilibrium. The time is greater with the stepping-stone model, but only slightly so.

The exact correspondence between the torus model and a rectangular plane is not known, but according to Maruyama (18), in a continuous model a torus-like figure of size $L \times L$ gives approximately the same value of $F$ as a plane of size $2L \times 2L$. We conclude that, except for narrow habitats, migrants from the population at random are typically between one and two times as effective as if they came from neighboring groups.

## Discussion

Our general approach is to use molecular markers, which are selected very weakly at most, as neutral indicators of population structure. $G_{ST}$ gives us an appropriate description of the relevant aspect of the structure. By using Eq. 3 we can state the maximum value of cost/benefit of a quantitative trait if that trait is to increase in average value or frequency in the population. $G_{ST}$ describes the present structure of the population; it does not tell us how it got that way. If this value has been roughly stable in the past, we could expect that traits with $c/b$ up to this value would have increased in the population, assuming of course that such traits exist and have heritability greater than zero.

Of more interest than the maximum $c/b$ of a selected trait is the average $c/b$ ratio of traits that are incorporated into the population. This question has been investigated by Engels (23), who showed that if there is a continuum of $c/b$ values available, evolution of altruistic behavior tends to maximize $rb - c$. If selection is weak, the average value at equilibrium is $r/2$, one-half the upper limit given by the Hamilton inequality. If we take Aoki and Nozawa's estimate of $r = 1/6$ for Japanese monkeys, we would expect that if this structure is stable traits with $c/b < 1/6$ would increase and that the average value at equilibrium would be 1/12. (Of course we would expect greater altruism of post-reproductives toward pre-reproductives than vice versa, but that is another subject.)

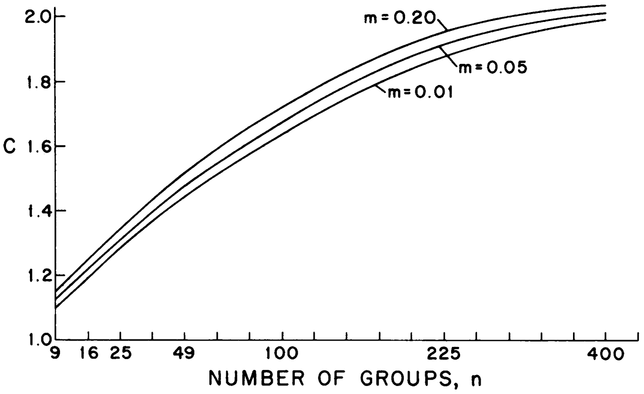

FIG. 3. Migration conversion factor ($C$) as a function of the number of groups, $n$. The conversion factor is the number of immigrants from immediate-neighbor groups that is equivalent to one immigrant chosen randomly from the rest of the population in determining the equilibrium value of $G_{ST}$. The group size is 20 and the mutation rate is $10^{-6}$; $m$ is the migration rate.

The effect of migration in homogenizing the population is more similar between the island and stepping-stone models than one might at first expect. Only with a very large number of groups does the effect of a random migrant exceed that of a migrant from a neighboring group by a factor larger than 2. Actual populations are probably somewhere between these two extreme models. Comparison of $G_{ST}$, which measures the present genetic structure that is the result of the migration pattern in the past, with $M$, which measures the present pattern of migration, can provide a possible test of the stability of the structure.

The relations between local gene identity, $F_0$, and global identity, $\bar{F}$, in this article employ a different definition of $m$ from that originally used by Wright. In Wright's formulation $m$ is the rate at which individuals are exchanged between a group and a pool to which the entire population contributes. Here we assume, as some others have (e.g., refs. 9 and 10), the immigrants have come from a pool drawn from all *other* groups. This makes the model more comparable to the stepping-stone model; as $n$ decreases the island and stepping-stone models become more and more similar. Of course, when $n$ is large the definition of $m$ in the island model is immaterial.

Fig. 2 suggests that the shape of the habitat can be important. The degree of divergence of subpopulations, and hence the opportunity for intergroup selection, is substantially increased if migrants come from nearby groups and the habitat is long and narrow.

This analysis gives no information as to whether the requisite heritable altruistic characters exist and whether there has been sufficient time for such traits to evolve. We have chosen to emphasize polygenic inheritance because this seems to us to be much more likely than monogenic inheritance where complex traits, especially behavioral traits, are involved; hence, our concentration on regressions and variances. Molecular data are brought in because they can reveal the details of population structure.

Much discussion of altruism has involved close relatives. We have tried here to model a system in which the trait, whether behavioral or not, affects the other members of the group without regard to their relationship. When an individual migrates to another group, its descendants behave the same way as if their ancestors had been in this group. In other words, the altruism is dispensed toward whatever group the individual grows up in regardless of where its genes ultimately come from. To us, this seems like a reasonable assumption, but we invite the scrutiny of those more familiar with animal behavior.

In summary, we suggest that molecular measurements can be used to estimate $G_{ST}$ and thus to assess the degree of population subdivision relevant to the evolution of altruistic behavior. This could permit testable predictions—for example, by observations of behavior in otherwise similar populations with differing $G_{ST}$ values or comparison of behavior in populations with areal vs. linear habitats.

A number of people have contributed in a variety of ways to our understanding of this subject. They include Carter Denniston, Bill Engels, Motoo Kimura, Takeo Maruyama, Thomas Nagylaki, Masatoshi Nei, Naoyuki Takahata, and Yoshio Tateno. This is paper no. 2729 from the Laboratory of Genetics, University of Wisconsin, and contribution no. 1564 from the National Institute of Genetics. This work was supported in part by a Grant-in-Aid for Special Project Research on Biological Aspects of Optimal Strategy and Social Structure from the Japan Ministry of Education, Science, and Culture.

1. Crow, J. F. & Aoki, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2628–2631.
2. Wright, S. (1951) *Ann. Eugen.* **15**, 323–354.
3. Hamilton, W. D. (1964) *J. Theor. Biol.* **7**, 1–52.
4. Leigh, E. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2985–2989.
5. Aoki, K. (1982) *Jpn. J. Genet.* **57**, 297–300.
6. Aoki, K. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4065–4068.
7. Nei, M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 3321–3323.
8. Nei, M. (1975) *Molecular Population Genetics and Evolution* (North Holland, Amsterdam).
9. Nagylaki, T. (1983) *Theor. Popul. Biol.* **24**, 268–294.
10. Takahata, N. (1983) *Genetics* **104**, 497–512.
11. Li, W. H. (1976) *Am. Nat.* **110**, 841–847.
12. Wright, S. (1943) *Genetics* **28**, 114–138.
13. Cockerham, C. C. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 530–534.
14. Nei, M., Chakravarti, A. & Tateno, Y. (1977) *Theor. Popul. Biol.* **11**, 291–306.
15. Malécot, G. (1950) *C. R. Acad. Sci. (Paris)* **222**, 841–843.
16. Kimura, M. (1953) *Annu. Rep. Natl. Inst. Genet. (Japan)* **3**, 62–63.
17. Maruyama, T. (1970) *Theor. Popul. Biol.* **1**, 273–306.
18. Maruyama, T. (1971) *Ann. Hum. Genet.* **35**, 179–196.
19. Maruyama, T. (1972) *Genetics* **70**, 639–651.
20. Aoki, K. & Nozawa, K. (1984) *Primates* **25**, 171–184.
21. Malécot, G. (1948) *Les Mathematiques d l'Heredite* (Maisson, Paris).
22. Kimura, M. & Weiss, G. H. (1964) *Genetics* **49**, 561–576.
23. Engels, W. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 515–518.