over the surface of the animal in reversal is identical to that of the animal's surface over the substratum in righting (Fig. 1*b*). A reflexive basis for the reversal of the covering reaction supports the suggestion of a reflexive basis for the covering reaction itself. The functional consequences of covering as proposed by other workers could still be valid, but would not be the casual explanations of the phenomenon.

This work was done at the Galeta Marine Laboratory of the Tropical Research Station of the Smithsonian Institution. I thank Dr G. Hendler for his assistance.

JOHN M. LAWRENCE

Department of Biology,
The University of South Florida,
Tampa, Florida 33620

Received April 1; accepted May 25, 1976.

1 Gamble, J. C., in *Symp. Underwater Assoc. Malta*, 1965 (edit. by Lythgoe, J. N., and Woods, J. D.), 47–50 (1966).
2 Dix, T. G., *N. Z. Jl mar. Freshwat. Res.*, **24**, 187–194 (1970).
3 Lees, D. C., and Carter, G. A., *Ecology*, **53**, 1127–1133 (1972).
4 Lewis, J. B., *Can. J. Zool.*, **36**, 607–621 (1958).
5 Sharp, D. T., and Gray, I. E., *Ecology*, **43**, 309–313 (1962).
6 Millott, N., *J. exp. Biol.*, **33**, 508–523 (1956).
7 Glynn, P., *Mar Biol*, **1**, 226–243 (1968).
8 Orton, J. H., *J. mar. biol. Ass. U.K.*, **16**, 289–296 (1929).
9 Dayton, P. K., Robilliard, G. A., and Paine, R. T., in *Antarctic Ecology*, **1** (edit. by Holgate, M. W.), 244–258 (Academic, New York, 1968).
10 Pequignat, E., *Nature*, **210**, 397–399 (1966).
11 Pequignat, E., and Tiffon, Y., *C. hebd. Séanc. Acad. Sci. Paris*, **264D**, 3014–3015 (1967).
12 Pequignat, E., *Mar. Biol.*, **12**, 28–41 (1972).
13 Yanagisawa, Y., *Publ. Seto mar. Biol. Lab.*, **19**, 431–435 (1972).
14 Dambach, M., and Hentschel, G., *Mar. Biol.*, **6**, 135–141 (1970).

Testing the neutral mutation hypothesis by distribution of single locus heterozygosity

ONE of the most controversial issues in population genetics at present is whether the widespread protein variation in natural populations is maintained by some form of balancing selection[1–3] or merely represents the drifting polymorphism of neutral or nearly neutral mutations[4–6]. By the neutral mutation hypothesis the level of genetic variability in an equilibrium population is determined by mutation rate, v, and effective population size, N_e. If we use the infinite allele model, in which new mutations are assumed to be always different from the pre-existing alleles in the population, the average heterozygosity per locus is given by

$$H = M/(1+M) \quad (1)$$

where $M = 4N_e v$ (ref. 7).

Whether the observed heterozygosity agrees with the above theoretical value or not is, in practice, however, very difficult to examine, since the exact values of N_e and v in a population are hardly obtainable. Recently, it was shown[8,9] that the variance of single locus heterozygosities is given by

$$V(h) = \frac{2M}{(1+M)^2(2+M)(3+M)} \quad (2)$$

From this the neutral mutation hypothesis can be tested without knowing the values of N_e and v separately. By estimating M from the observed value of average heterozygosity, we can compute the theoretical variance by using equation (2) and then compare it with the observed variance[10]. Preliminary results of the application of this method to a number of organisms have suggested that the agreement between the theoretical and observed variances is rather good[11,12], but the amount of data used was too small to make any general conclusion. We have accumulated all published data which are amenable to such a test. In this report we show that the observed variances of heterozygosity in most organisms so far studied can be reasonably explained by the neutral mutation hypothesis. The pattern of the distribution of heterozygosity also agrees with that expected from this hypothesis.

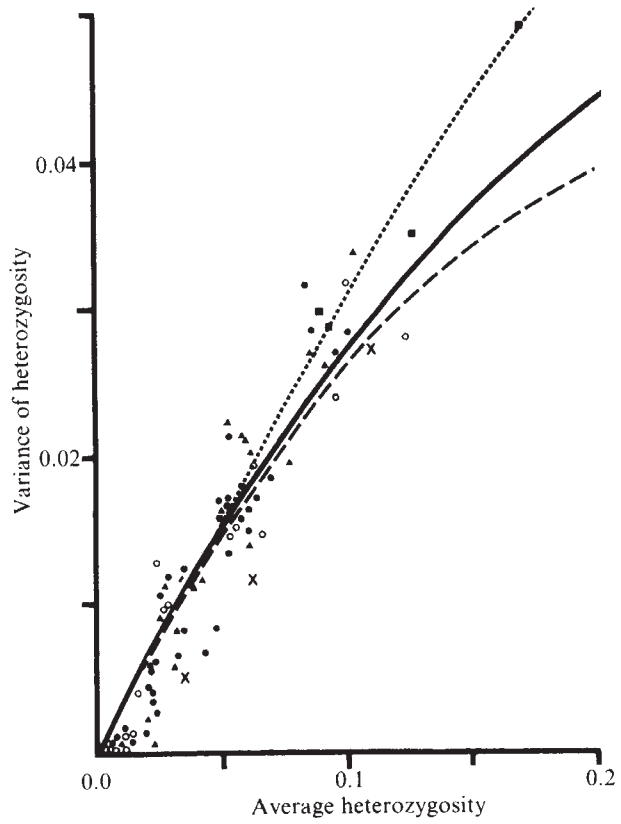

Fig. 1 Relationship between the mean and variance of heterozygosity for vertebrate species: ●, mammals (33 species); ×, birds (2 species, 1 subspecies); ○, fish (18 species, 1 subspecies); ▲, lizards (21 species); ■, amphibians (3 species, 1 subspecies). The solid and dashed lines are the theoretical relationships obtained from the infinite allele and stepwise mutation models, respectively. The dotted line is the theoretical relationship for the infinite allele model with varying mutation rate.

Surveying 52 publications, we collected gene frequency data obtained by electrophoresis for 123 species and 8 subspecies, in which 15 or more loci were examined for at least 30 genomes. It is known that for estimating average heterozygosity a large number of loci should be used, whereas the sample size per locus can be rather small[10]. For each of the species and subspecies used, the mean, $\hat{H}$, and variance, $\hat{V}(h)$, of single locus heterozygosity was computed with a correction for sampling variance[10]. When a species or subspecies comprised a number of subpopulations, the average values of $\hat{H}$ and $\hat{V}(h)$ for these were used. The relationship between $\hat{H}$ and $\hat{V}(h)$ was compared with the theoretical one (solid lines in Figs 1 and 2) obtained from equations (1) and (2). It can be argued that the stepwise mutation model[13] is more appropriate for electrophoretic data than the infinite allele model. Therefore, we have also computed the theoretical curve for this model (broken line) using Ohta and Kimura's formula[13] for H and Moran's formula[11] for $V(h)$. Actually, as Figs 1 and 2 show, there is not much difference between the two models unless H is very large. Since there seems to be a significant difference in average heterozygosity between vertebrates and invertebrates[15], the two groups have been considered separately (Figs 1 and 2).

It is clear that the observed variances for vertebrates are generally in good agreement with the theoretical values, in spite of the fact that the number of loci used is relatively small and data on diverse organisms are included. A close examination of Fig. 1, however, indicates that when $\hat{H}$ is smaller than 0.05

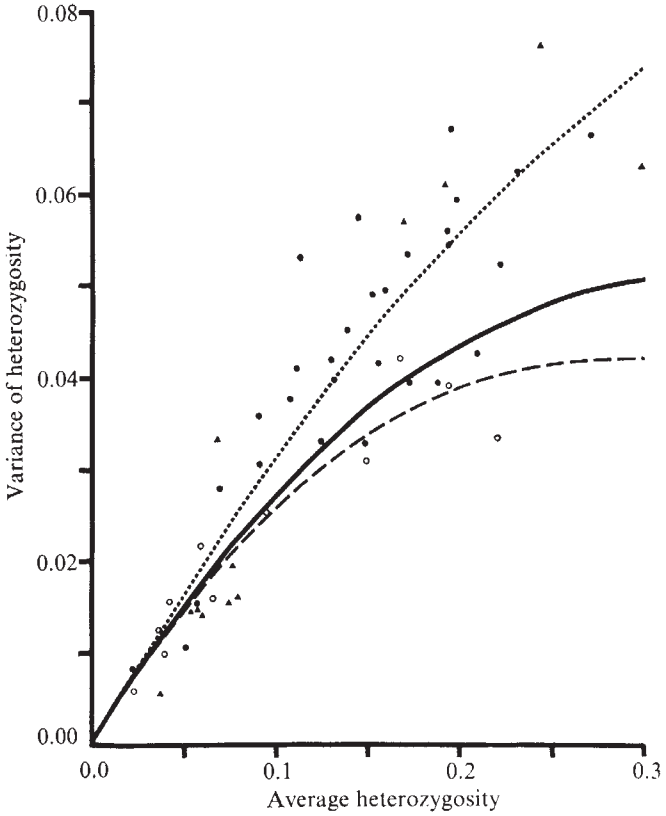

Fig. 2 Relationship between the mean and variance of heterozygosity for invertebrate species: ●, *Drosophila* (24 species, 5 subspecies); ▲, other insects (12 species); ○, other invertebrates (10 species). The solid and dashed lines are the theoretical relationships obtained from the infinite allele and stepwise mutation models, respectively. The dotted line is the theoretical relationship for the infinite allele model with varying mutation rate.

the observed variance tends to be smaller than the theoretical curve, whereas for $\hat{H} > 0.1$ it tends to be larger. The agreement between the theoretical and observed variances in invertebrates is also generally good, though the tendency for the observed variance to deviate upward is pronounced because of higher heterozygosities in this group of organisms.

The tendency for the observed variance to deviate upward for large $\hat{H}$ does not contradict the neutral mutation hypothesis and is expected to occur for the following reason. The formula for $V(h)$ is based on the assumption of identical mutation rates for all loci in both theoretical models used here. In practice, this assumption is surely incorrect, and when mutation rate varies among loci, the variance of heterozygosity is expected to inflate. To get an idea about the amount of inflation, we computed the expected variance of heterozygosity from the assumption that mutation rate varies among loci according to the gamma distribution with the coefficient of variation equal to one. The gamma distribution with this value of coefficient of variation was suggested by the examination of the rates of amino acid substitution for 19 polypeptides in evolution (Nei, Chakraborty and Fuerst, unpublished). From the assumption of neutral mutations the rate of amino acid substitution per polypeptide is equal to the mutation rate per locus, neglecting synonymous mutations[4]. At any rate, the results obtained are given by the dotted lines in Figs 1 and 2. It is clear that the agreement between the theoretical and observed variances is now very good.

The tendency for the observed variance to deviate downward for a small value of $\hat{H}$ is apparently caused by the relatively small number of loci studied (about 20) in this group of organisms. When the theoretical value of H is relatively small, the distribution of single locus heterozygosity, h, is extremely skewed towards the class of $h = 0$ (Fig. 3). Thus, whenever the sample estimate of H is small, the observed value of $V(h)$ tends to be smaller than its expected value. Our computer simulation has shown that the difference between the observed and theoretical variances for $\hat{H} \leqslant 0.05$ is of the order of magnitude that is expected to occur when the number of loci is about 20. The details of this study will be published elsewhere. Thus, the data presented in Figs 1 and 2 are consistent with the neutral mutation hypothesis.

In Fig. 3 the frequency distributions of single locus heterozygosity for three species groups in which a large number of loci have been studied are presented together with one theoretical distribution obtained by computer simulation[16]. This theoretical distribution is based on 2,000 replicate evaluations of h with the expected heterozygosity of 0.09, which is close to the observed average heterozygosities for the three species groups. All empirical distributions are L-shaped and very similar to the theoretical one. In both empirical and theoretical distributions there is a little peak around $h = 0.5$. Existence of this peak has been theoretically predicted[8,17]. In addition to those in Fig. 3, we examined the distribution of heterozygosity for all other species. The pattern of the distribution was essentially the same as that expected from the neutral mutation hypothesis for most species.

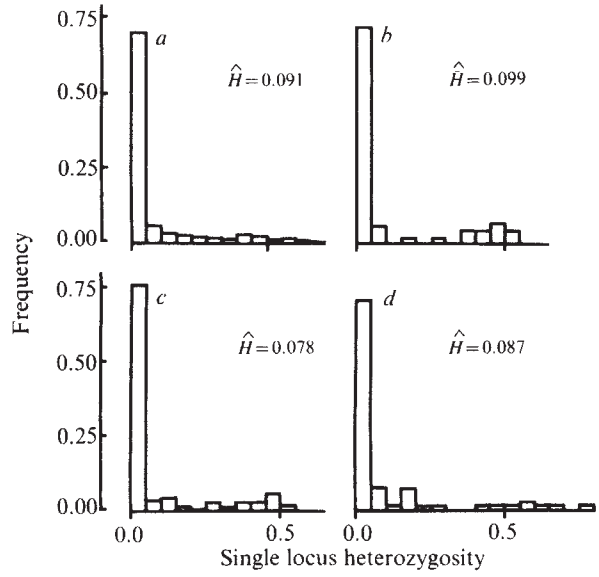

Fig. 3 Distributions of single locus heterozygosities for *b*, man (Caucasian)[11]; *c*, the house mouse[20,21] (*Mus musculus*) and *d*, the *Drosophila mulleri* species group[22]. The theoretical distribution (*a*) was obtained by computer simulation with the infinite allele model. $\hat{H}$ stands for the average heterozygosity. The distribution for the stepwise mutation model is almost identical with that for the infinite allele model when $\hat{H} = 0.09$.

We are aware that our observations about the variation and distribution of heterozygosity can be explained by some elaborate combination of different types of selection. Such an explanation is, however, strained, since we then have to assume that selection is adjusted so as to have a good agreement between the observed and theoretical variances in diverse organisms. It seems to us that the simplest and most natural explanation is the neutral mutation hypothesis. Note that this hypothesis does not exclude the existence of a small proportion of overdominant or selectively advantageous mutations. Furthermore, our analysis will not exclude the possible importance of slightly deleterious mutations[18,19], of which the contribution to heterozygosity is relatively small.

We thank Yoshio Tateno for some numerical computations involved in this study. We are also grateful to Professor P. A.

P. Moran for his unpublished formula. This work was supported by grants from the NIH.

MASATOSHI NEI
PAUL A. FUERST
RANAJIT CHAKRABORTY

Center for Demographic and Population Genetics, The University of Texas at Houston, 1100 Holcombe Boulevard, Room 1109, Houston, Texas 77030

Received January 21; accepted June 9, 1976.

1 Sved, J. A., Reed, T. E., and Bodmer, W. F., *Genetics*, **55**, 469–481 (1967).
2 King, J. L., *Genetics*, **55**, 483–492 (1967).
3 Milkman, R. D., *Genetics*, **55**, 493–495 (1967).
4 Kimura, M., *Nature*, **217**, 624–626 (1968).
5 King, J. L., and Jukes, T. H., *Science*, **164**, 788–798 (1969).
6 Kimura, M., and Ohta, T., *Nature*, **229**, 467–469 (1971).
7 Kimura, M., and Crow, J. F., *Genetics*, **49**, 725–738 (1964).
8 Stewart, F., *Theor. Populat. Biol.* (in the press).
9 Li, W.-H., and Nei, M., *Genet. Res.*, **25**, 229–248 (1975).
10 Nei, M., and Roychoudhury, A. K., *Genetics*, **76**, 379–390 (1974).
11 Nei, M., and Roychoudhury, A. K., *Am. J. hum. Genet.*, **26**, 421–443 (1974).
12 Nei, M., *Molecular Population Genetics and Evolution* (North-Holland, Amsterdam, 1975).
13 Ohta, T., and Kimura, M., *Genet. Res.*, **22**, 201–204 (1973).
14 Moran, P. A. P., *Theor. Populat. Biol.* (in the press).
15 Selander, R. K., and Kaufman, D. W., *Proc. natn Acad. Sci. U.S.A.*, **70**, 1875–1877 (1973).
16 Nei, M., and Tateno, Y., *Proc. natn Acad. Sci. U.S.A.*, **72**, 2758–2760 (1975).
17 Ewens, W. J., and Gillespie, J. H., *Theor. Populat. Biol.*, **6**, 35–57 (1974).
18 Ohta, T., *Proc. natn Acad. Sci. U.S.A.*, **72**, 3194–3196 (1975).
19 Latter, B. D. H., *Nature*, **257**, 590–592 (1975).
20 Selander, R. K., Hunt, W. G., and Yang, S. Y., *Evolution*, **23**, 379–390 (1969).
21 Selander, R. K., and Yang, S. Y., *Genetics*, **63**, 653–667 (1969).
22 Zouros, E., *Evolution*, **27**, 601–621 (1973).

Y chromosome visibility in quinacrine-stained human spermatozoa

THE fluorescent dye, quinacrine, binds strongly to the Y chromosome and only to a lesser extent to the other chromosomes in human metaphase cell preparations[1]. A bright fluorescent spot (F body), which is presumed to be the Y chromosome, is clearly visible in stained interphase cells from various tissues from the human male[2-5] and in mature spermatozoa[6,7]. Spermatozoa bearing one F body (about 45% of the total) are taken to be Y spermatozoa, and the staining method is now being used widely to detect physical differences, for example, in DNA content and head size[7,8] and motility[9] between X and Y spermatozoa, and for testing the efficiency of procedures designed to separate them[10-13]. Spermatozoa bearing two F bodies (about 1.3% of the total) have been interpreted as YY spermatozoa[7,8], and the relatively high number of such spermatozoa has led to speculation concerning the seemingly high non-disjunction rates during spermatogenesis[8,14,15]. We here examine the factors which affect Y chromosome visibility in sperm heads, and conclude that the F body count depends critically on observer performance, that the presence of YY spermatozoa cannot be demonstrated decisively using the quinacrine stain, and that measured physical differences between spermatozoa with and without an F body may partly reflect a dependence of the Y chromosome detection efficiency on the size of the sperm head.

When a brightly fluorescent chromosome inside a nucleus is examined under the ultraviolet microscope it appears as a bright spot against a less bright background which is produced by fluorescence from the other nuclear contents and from any unbound dye which may be present. The chromosome is detected by the observer provided that the spot intensity is more than 1 or 2% above that of the background[16], the condition for detection then being

$$(I_0 \exp(-k\bar{x}) - I_B)/I_B \gtrsim 0.01\text{–}0.02 \qquad (1)$$

where $\bar{x}$ is the mean depth (weighted and averaged over all directions which lie within the acceptance cone of the objective lens) at which the chromosome is buried in the head, I_0 is the intensity of the spot which would be produced if the chromosome were located at $\bar{x} = 0$, I_B is the background intensity and k is an optical attenuation coefficient.

Almost 50% of spermatozoa are expected to carry a Y chromosome, and since only about 45% exhibit an F body it seems that about a tenth of all Y chromosomes are not detected by the staining procedure. There are two possible explanations. The first is that stain uptake among Y chromosomes from cells of any one individual may be very variable, the missing chromosomes being those which absorb so little stain as to be indiscernible above the background fluorescence. The second is that the optical attenuation coefficient is sufficiently great that chromosomes lying deep in the head are not detected[6]. This coefficient has two components. The refractive index of the DNA within the head is high (probably about 1.5 (ref. 17)) compared with that of the aqueous mounting medium (1.33) which surrounds and permeates the sperm head. If the fluorescent Y chromosome is located at the bottom of a uniform sphere of refractive index 1.5 immersed in a medium of refractive index 1.33, the average refractive deviation of light leaving the chromosome and passing through the sphere is found using Snell's law to be about 15°, sufficient to reduce the intensity of a chromosome of diameter 500 nm to about 25% of its undiminished value. If the other chromosomes each refract the light from the Y chromosome in this way the total deviation will be considerably greater than 15° across the optically inhomogeneous sperm head. In addition, the dimensions of the chromosomes are comparable with the wavelength of the fluorescent light, and diffractive scattering will occur within the sperm head, increasing background fluorescence at the expense of spot intensity. Optical attenuation may therefore be the dominant mechanism, and some support for this hypothesis is afforded by the observation that about 70% of F bodies appear at or very near the region where the background fluorescence becomes less dense[15]. We have estimated the intensities of F bodies by comparing microscope images of quinacrine-stained spermatozoa with matching images projected on to a screen in the same conditions (green light; dark-adapted vision) and measuring the screen intensities with a photometer. F body intensities varied from about 10% (the brightest) to about 1% (that is, the detection threshold) above background. Assuming that the brightest spots correspond to Y chromosomes at $\bar{x} = 0$ and that the Y chromosome is randomly distributed within the head, then $k \sim 50\ \mathrm{mm}^{-1}$ would give the observed proportion of spermatozoa carrying an F body.

Other chromosomes, although fluorescing less brightly than the Y chromosome, may nevertheless be detected if they are located near the surface of the sperm head. If P_y is the probability of detecting a Y chromosome in a Y spermatozoon, and P_c is the probability of detecting any other chromosome (or other artefact such as adhering bacteria or vacuoles within the head[15]) then, assuming that X and Y spermatozoa occur in equal numbers (and, initially, that no YY spermatozoa are present) the percentages of spermatozoa exhibiting one and two F bodies, P_1 and P_2, are given by

$$P_1 = 50P_y(1-P_c) + 100(1-\tfrac{1}{2}P_y)P_c$$

and

$$P_2 = 50P_yP_c + 100(P_c)^2$$

Solving for $P_1 = 45\%$ and $P_2 = 1.3\%$ yields $P_y \sim 0.89$ and $P_c \sim 0.0275$, this being the only solution for which both P_y and P_c lie within the range 0–1. These figures may be compared with those from buccal mucosa cells and lymphocytes where cells of known karyotype are examined[2,3]. A fluorescent spot indistinguishable from a normal Y body[2] occurs in about 5.5% of cells from 46,XX females who carry no Y chromosome, and cells from 46,XY males, carrying one Y chromosome, have one F body in about 67% of cases[2,3]. Thus $P_y \sim 0.65$ and $P_c \sim 0.055$, and the Y chromosome is clearly more difficult to detect in these cells than in spermatozoa.