ENFIELD, F. D. 1977. Selection experiments in *Tribolium* designed to look at gene action issues, pp. 177–190. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr. (eds.), Proc. Int. Conf. Quant. Genetics Iowa State Univ. Press, Ames.

JONES, L. P., R. FRANKHAM, AND J. S. F. BARKER. 1968. The effects of population size and selection intensity in selection for a quantitative character in *Drosophila*. II. Long-term response to selection. Genet. Res. 12:249–266.

KIMURA, M. 1957. Some problems of stochastic processes in genetics. Ann. Math. Stat. 28:882–901.

KIMURA, M. 1968. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. Genet. Res. 11:247–269.

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

LANDE, R. 1977. Statistical tests for natural selection on quantitative characters. Evolution 31:442–444.

LEVIN, D. A. 1979. The nature of plant species. Science 204:381–384.

LEVIN, D. A., AND H. W. KERSTER. 1974. Gene flow in seed plants. Evol. Biol. 7:139–220.

NEI, M. 1975. Molecular Population Genetics and Evolution. North-Holland, Amsterdam, Neth.

ROBERTSON, A. 1962. Selection for heterozygotes in small populations. Genetics 47:1291–1300.

Corresponding Editor: M. Slatkin

Evolution, 40(3), 1986, pp. 643–645

DEFINITION AND ESTIMATION OF FIXATION INDICES

MASATOSHI NEI
Center for Demographic and Population Genetics, University of Texas at Houston, P. O. Box 20334, Houston, TX 77225

Received June 28, 1985. Accepted January 16, 1986

Recently, Weir and Cockerham (1984) presented a simplified version of Cockerham's (1969, 1973) method for estimating Wright's (1943, 1951) fixation indices (*F*-statistics). They concluded that their method is more precise than the others currently in use. In this study, however, they did not consider the difference in assumptions between the different methods. In the following, I would like to clarify this point, particularly the difference between Cockerham and Weir's method and mine.

Cockerham's (1969) and Weir and Cockerham's (1984) method is based on the following assumptions. 1) Conceptually, there are infinitely many populations which are derived from the same ancestral population at the same time, and *s* populations are sampled from this ensemble population. All fixation indices are defined in terms of the ensemble population, and the fixation indices are estimated from genotype frequency data from the *s* populations sampled. 2) All populations are statistically independent, and no mutation, no migration, and no selection are assumed (Weir and Cockerham, 1984). However, certain types of migration and selection can be incorporated as long as the statistical independence is maintained (Cockerham, 1969, 1973). 3) The same population size (*N*) is maintained for all populations and for all generations (Weir and Cockerham, 1984).

The above assumptions are essentially the same as those of Wright's (1951) inbreeding theory, and under these assumptions Weir and Cockerham's statistical method seems to be correct. However, the problem lies with the validity of the assumptions. Obviously, most natural populations do not satisfy any of these assumptions. Population sizes usually vary greatly, and they are almost never constant over time. This is so even if we consider a relatively short evolutionary time. Migration between populations is also a serious problem as Cockerham (1973) himself realized; the migration rate usually varies greatly from population pair to population pair. The number of populations within a species also varies with time, and at a given time some populations are more closely related to each other than to others. In addition to these factors, mutation and selection are also expected to disturb Weir and Cockerham's formulation.

What should we do then? Can we do nothing with natural populations? My solution to this problem is to abandon the above assumptions and to reformulate fixation indices in a different way (Nei, 1977). I have done this by redefining all fixation indices in terms of existing populations for which genotype frequencies are studied. In this case, no assumption is required about the various factors that caused the differentiation of the populations. We are simply interested in the extent of genetic differentiation of the populations studied. Obviously, there is no need to consider replicate populations in this approach. However, since only a limited number of individuals are sampled from each population, it is necessary to estimate fixation indices (parameters), as in the case of estimation of gene frequencies or genetic distances (Nei and Chesser, 1983).

One might think that the fixation indices defined above are less useful than Weir and Cockerham's indices. This is not true. Certainly, my fixation indices refer only to the populations studied, but they can be used for comparing the extent of genetic differentiation among different sets of populations. For example, the extent of genetic differentiation of European popula-

tions is much smaller than that of the South American Indian populations. If such a difference is established, we can then look into factors that caused the difference using some other type of information. It is usually dangerous to extrapolate estimates of fixation indices obtained from one set of populations to another, because their evolutionary histories are usually different. Another use of fixation indices is to identify an allele or an allelic pair at a multiallelic locus that shows an unusual genotypic distribution among populations by computing fixation indices for all alleles or for all allelic pairs separately. This may lead to detection of certain types of selection (Nei, 1965, 1977; Wright, 1978).

Nevertheless, there is one deficiency in my approach. In F-statistics analysis, the most important parameter is F_{ST}, and this is defined as D_{ST}/H_T, where D_{ST} is a measure of interpopulational gene differences and H_T is the heterozygosity expected under Hardy-Weinberg equilibrium for the total population (Nei, 1973, 1977). D_{ST} is dependent on the number of populations (s), as noted by Nei (1973, 1975). Let p_{ik} be the frequency of the kth allele in the ith population. D_{ST} is then given by

$$D_{ST} = \left(\sum_i \sum_j D_{ij} \right) / s^2, \qquad (1)$$

where $D_{ij} = \sum_k (p_{ik} - p_{jk})^2/2$. Obviously, $D_{ii} = 0$, so that a more appropriate measure of interpopulational gene differences would be

$$D_{ST}' = \sum_{i \neq j} D_{ij} / [s(s-1)]. \qquad (2)$$

(This was denoted by $\bar{D}_m$ in my 1973 paper.) This is equal to the average pairwise (minimum) genetic distance, and I have suggested that D_{ST}' be used as the absolute measure of gene differentiation. D_{ST} and D_{ST}' are related by $D_{ST}' = sD_{ST}/(s-1)$.

Despite the above dependence of D_{ST} on s, I have used the definition $F_{ST} = D_{ST}/H_T$ in my 1977 paper, because 1) this is equivalent to Wright's (1951, 1965) original definition of F_{ST}, 2) the effect of s is minor unless s is very small, and 3) mathematical simplicity is maintained.

Of course, if necessary, one can redefine F_{ST} and F_{IT}, using D_{ST}' rather than D_{ST}, and estimate them. (F_{IS} is unaffected.) In this case, the total heterozygosity should also be redefined as $H_T' = H_S + D_{ST}'$ rather than $H_S + D_{ST}$, where H_S is the average heterozygosity expected under Hardy-Weinberg equilibrium within populations (see Nei, 1973, 1977). One can then redefine the fixation indices as $F_{IS} = H_O/H_S$, $F_{ST}' = D_{ST}'/H_T'$, and $F_{IT}' = H_O/H_T'$, where H_O is the average observed heterozygosity within populations. These indices can be estimated by a slight modification of the method of Nei and Chesser (1983). We have shown how to obtain unbiased estimates ($\hat{H}_O$, $\hat{H}_S$, and $\hat{H}_T$) of H_O, H_S, and H_T. Therefore, unbiased estimates of D_{ST}' and H_T' are given by $\hat{D}_{ST}' = s(\hat{H}_T - \hat{H}_S)/(s-1)$ and $\hat{H}_T' = \hat{H}_S + \hat{D}_{ST}'$, respectively. Thus, we have

$$\hat{F}_{IS} = 1 - \hat{H}_O/\hat{H}_S, \qquad (3a)$$

$$\hat{F}_{IT}' = 1 - \hat{H}_O/\hat{H}_T', \qquad (3b)$$

$$\hat{F}_{ST}' = 1 - \hat{H}_S/\hat{H}_T'. \qquad (3c)$$

These above equations are closer to Weir and Cockerham's estimators than the previous ones are, but there are still some differences because of the different assumptions made.

Conducting a computer simulation with $s = 3$ and $\tilde{n}$ (harmonic mean of n) = 19 *under their assumptions,* Weir and Cockerham (1984) concluded that "the traditional estimator $\hat{\theta}_T$ is the worst in terms of bias." In their computation of $\hat{\theta}_T$, however, they did not make the traditional sample size correction ($\frac{1}{2}\tilde{n}$) (Wright, 1943; Nei and Imaizumi, 1966; Workman and Niswander, 1970). If we make this correction, all of their $\hat{\theta}_T$ values decrease by 0.026 and become very close to the values obtained by their method. It should also be noted that in order to obtain a reliable estimate of their parameter θ, both s and n must be quite large, and in this case their method and ours give essentially the same result.

Weir and Cockerham criticized Nei and Chakravarti's (1977) contention that F_{ST} be defined as 0 when the same allele is fixed in all s ($<\infty$) populations. In my view, Nei and Chakravarti's definition is quite appropriate, since there is no other allele in the entire population. Note that Nei and Chakravarti's study is concerned with the evolutionary change of population F_{ST} for a finite number of populations and has nothing to do with the estimation of F_{ST}. Note also that we are interested in the degree of gene differentiation among populations rather than in the coefficient of inbreeding or coancestry within populations in which Weir and Cockerham (1984 p. 1358) are primarily interested.

Finally, I should like to clarify the thrust of my theory of gene diversity analysis (Nei, 1973). This theory was developed primarily for the decomposition of the total average heterozygosity (H_T) into the within-population (H_S) and between-population (D_{ST}) components, and thus it was intended to be applied to a large number of loci, including both polymorphic and monomorphic loci (Nei, 1975, p. 149). I called the ratio (G_{ST}) of D_{ST} to H_T the coefficient of gene differentiation. Therefore, G_{ST} is conceptually different from F_{ST}, though when a single polymorphic locus is considered they become identical with each other (Nei, 1973). Obviously, G_{ST} is dependent on s because of the way it is defined. This dependence can again be removed if we redefine G_{ST} by using D_{ST}' instead of D_{ST} as in the case of F_{ST}. In this case, however, H_T' is no longer the total heterozygosity. For this reason, I did not recommend this correction. Instead, I suggested that D_{ST}' be used as a measure of the gene differentiation among populations. However, there is nothing wrong in using this redefined quantity (G_{ST}') as long as the definition is clearly understood. In my view, users should decide which method is to be used, depending on their purpose and circumstances.

Acknowledgments

I thank Peter Smouse, Pekka Pamilo, and Clay Stephens for their comments. This work was supported by research grants from the National Institutes of Health and the National Science Foundation.

Literature Cited

Cockerham, C. C. 1969. Variance of gene frequencies. Evolution 23:72–84.

———. 1973. Analyses of gene frequencies. Genetics 74:679–700.
NEI, M. 1965. Variation and covariation of gene frequencies in subdivided populations. Evolution 19: 256–258.
———. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. USA 70:3321–3323.
———. 1975. Molecular Population Genetics and Evolution. North-Holland, Amsterdam, Neth.
———. 1977. *F*-statistics and analysis of gene diversity in subdivided populations. Ann. Hum. Genet. 41:225–233.
NEI, M., AND A. CHAKRAVARTI. 1977. Drift variances of F_{ST} and G_{ST} statistics obtained from a finite number of isolated populations. Theoret. Popul. Biol. 11:307–325.
NEI, M., AND R. K. CHESSER. 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Genet. 47:253–259.
NEI, M., AND Y. IMAIZUMI. 1966. Genetic structure of human populations. I. Local differentiation of blood group gene frequencies in Japan. Heredity 21:9–35.
WEIR, B. S., AND C. C. COCKERHAM. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370.
WORKMAN, P. L., AND J. D. NISWANDER. 1970. Population studies on southwestern Indian tribes. II. Local genetic differentiation in the Papago. Amer. J. Hum. Genet. 22:24–29.
WRIGHT, S. 1943. Isolation by distance. Genetics 28: 114–138.
———. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.
———. 1965. The interpretation of population structure by *F*-statistics with special regard to systems of mating. Evolution 19:395–420.
———. 1978. Evolution and the Genetics of Populations, Vol. 4. Variability Within and Among Natural Populations. Univ. Chicago Press, Chicago, IL.

Corresponding Editor: M. Slatkin

Evolution, 40(3), 1986, pp. 645–649

EVOLUTIONARY IMPLICATIONS OF CHROMOSOMAL POLYMORPHISMS IN *PEROMYSCUS BOYLII* FROM SOUTHWESTERN MEXICO

KATHLEEN M. DAVIS[1], STEVEN A. SMITH,[2] AND IRA F. GREENBAUM[1]
[1]*Department of Biology, Texas A&M University, College Station, TX 77843*
[2]*Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77843*

Received July 29, 1985. Accepted December 30, 1985

Chromosomal evolution in the deer mouse genus, *Peromyscus* (Cricetidae), has been well documented to involve alteration in the number of autosomal arms (fundamental number; FN) within a constant diploid number of 48. Chromosomal banding data from representatives of more than half of the approximately 60 species of *Peromyscus* indicate that karyotypic differentiation among species encompasses a range of FN = 52 to 92, and involves differences in the presence of heterochromatic short arms and pericentric inversions. Intraspecific chromosomal polymorphisms documented for relatively few species of *Peromyscus* have also been shown to involve heterochromatin and/or pericentric inversion variation (Greenbaum and Baker, 1978; Baker et al., 1983; Pengilly et al., 1983; Greenbaum and Reed, 1984). Although heterochromatin differences might not be expected to be involved with the dynamics of speciation, the incorporation of euchromatic structural rearrangements, such as pericentric inversions, should provide a ready mechanism for generating reproductive barriers among populations (for reviews see White, 1973, 1978). Since McClintock's (1931, 1933) initial observations of chromosome pairing at meiosis in inversion heterozygotes of *Zea mays*, such heterozygosity has been commonly assumed to result in a significant reduction of reproductive fitness. In this report, we present data which document the occurrence of three autosomal pericentric inversion polymorphisms which appear to be in equilibrium within populations of *P. boylii* from southwestern Mexico. These data are interpreted to indicate that, in some cases, structural rearrangements of chromosomes do not result in the generally assumed reduction of heterozygote fitness.

MATERIALS AND METHODS

Specimens of *P. boylii* (N = 78, see Table 1) were live-trapped from natural populations from the Sierra Madre del Sur in southwestern Mexico (Fig. 1). Karyotypes were obtained from bone marrow using the procedure described by Lee and Elder (1980), as modified by Baker et al. (1982). G- and C-band preparations were produced using modifications of the techniques described by Seabright (1971) and Sumner (1972), respectively. G-banded chromosomes were identified according to the standardized karyotype for *Peromyscus* (Committee, 1977), and C-banded chromosomes were correspondingly arranged. The chromosomal polymorphisms were identified from G- and C-banded karyotypes from six individuals. As each of the three