

## GENETIC DIFFERENTIATION IN THE COLONISING LIZARD *ANOLIS GRAHAMI*

ARAVINDA CHAKRAVARTI

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

Received 7.vi.76.

### SUMMARY

Taylor and Gorman's (1975) gene frequency data from the Jamaican and Bermudan populations of *Anolis grahami* are reanalysed by using  $F_{ST}$  statistics. Statistical drawbacks of the original analysis are indicated. The present analysis shows that the data are consistent with the theory of random genetic drift of selectively neutral mutations.

In a study of the distribution of Wright's (1943, 1951)  $F_{ST}$  for enzyme loci in the lizard *Anolis grahami*, Taylor and Gorman (1975) used Lewontin and Krakauer's (1973) test of selective neutrality of polymorphic genes and concluded that not all the alleles examined were selectively neutral. This conclusion, however, does not seem to be warranted, since their computation involved some errors.

Taylor and Gorman computed  $F_{ST}$  for each polymorphic allele by using gene frequency data from the Jamaican and Bermudan populations. They used the following formula.

$$F_{ST} = \frac{(x_2 - p)^2}{p(1 - p)} \quad (1)$$

where  $p$  is the gene frequency in the Jamaican population and  $x_2$  is that in the Bermudan population. An assumption underlying the use of this formula is that the gene frequency in the Jamaican population remained constant for the entire evolutionary time involved, since this population was very large. It is known that the Bermudan population was established in 1905 (68 generations before 1973) from 26 males and 45 females introduced from the Jamaican population.

They then computed the observed cumulative frequency distribution of  $6F_{ST}/\bar{F}_{ST}$ , where  $\bar{F}_{ST}$  is the average of  $F_{ST}$  over the seven alleles used. This distribution significantly deviated from the  $\chi^2$  distribution with six degrees of freedom. This led them to conclude that their data are not consistent with the neutral mutation hypothesis. It is, however, noted that the number of degrees of freedom to be used for  $F_{ST}/\bar{F}_{ST}$  is not six but one, since the number of populations ( $s$ ) is two. They erroneously used the number of alleles minus one as the number of degrees of freedom for this quantity. If we make this correction and recompute the observed cumulative distribution of  $(s-1)F_{ST}/\bar{F}_{ST}$ , then it agrees well with the  $\chi^2$  distribution with one degree of freedom. In this case the theoretical variance of  $(s-1)F_{ST}/\bar{F}_{ST}$  is  $2(s-1) = 2$ , while the observed variance is 1.63. Therefore, the null hypothesis of equality of the observed and theoretical variances cannot be rejected ( $F_{\infty}^{(1)} = 0.81$ ;  $P > 0.75$ ).

Besides the above error there seems to be another problem in Taylor and Gorman's data analysis. Since their sample sizes were small, they estimated the population  $F_{ST}$  by

$$F'_{ST} = F_{ST} - \left( \frac{1}{2M_1} + \frac{1}{2M_2} \right) \quad (2)$$

where  $M_1 = 38$  and  $M_2 = 43$  are the sample sizes from the Jamaican and Bermudan populations, respectively. This correction decreases the mean  $F_{ST}$  by the amount of  $(1/2M_1 + 1/2M_2)$ , whereas the variance of  $F_{ST}$  remains the same. Therefore, the ratio of the observed to the theoretical variances of  $(s-1) F_{ST}/\bar{F}_{ST}$  is necessarily inflated. In my opinion, this correction is unnecessary. Since we are interested in the effect of genetic drift, we can regard the sample size as the population size in the following generation. This is equivalent to assuming a bottleneck for the next generation.

It should also be noted that the theoretical variance of  $(s-1) F_{ST}/\bar{F}_{ST}$  can be larger than that of a  $\chi^2_{s-1}$  distribution even under pure genetic drift when the initial gene frequency is close to 0 or 1, as shown by Nei and Chakravarti (1977). Therefore, we must be cautious in using this type of test even in isolated populations.

Taylor and Gorman (1975) estimated the effective population size of the Bermudan population by using Wright's formula:

$$E(F_{ST}) = 1 - \left( 1 - \frac{1}{2N_2} \right)^t \quad (3)$$

where  $E(F_{ST})$  is the expectation of  $F_{ST}$  at time  $t$  and  $N_2$  is the effective size of the Bermudan population. This formula, however, is not appropriate when the number of subpopulations is extremely small. In this case a simple way to relate  $E(F_{ST})$  to evolutionary time is as follows:

Following Nei (1975), we define  $E(F_{ST})$  as:

$$E(F_{ST}) = \frac{E(s_x^2)}{E[\bar{x}(1-\bar{x})]} \quad (4)$$

where  $\bar{x}$  and  $s_x^2$  denote the mean and variance of gene frequency ( $x$ ). In the case of two subpopulations,

$$\begin{aligned} \bar{x} &= (x_1 + x_2)/2 \\ s_x^2 &= (x_1 - x_2)^2/4 \end{aligned}$$

where  $x_1$  and  $x_2$  are the gene frequencies in populations 1 and 2, respectively. Assuming that only random genetic drift is the cause for gene differentiation, it can be shown (see Crow and Kimura, 1970) that:

$$E(s_x^2) = \frac{p(1-p)}{4} \left\{ 2 - \left( 1 - \frac{1}{2N_1} \right)^t - \left( 1 - \frac{1}{2N_2} \right)^t \right\} \quad (5)$$

$$E[\bar{x}(1-\bar{x})] = \frac{p(1-p)}{4} \left\{ 2 + \left( 1 - \frac{1}{2N_1} \right)^t + \left( 1 - \frac{1}{2N_2} \right)^t \right\}. \quad (6)$$

If we assume that the Jamaican population ( $N_1$ ) has remained very large throughout the 68 generations (*i.e.*  $N_1 = \infty$ ), we have

$$E(F_{ST}) = \frac{E(s_x^2)}{E[\bar{x}(1-\bar{x})]} \approx \frac{1 - e^{-t/2N_2}}{3 + e^{-t/2N_2}} \quad (7)$$

The exponential approximation in the above formula is generally valid unless the effective population size is smaller than 10.

In applying formula (7) we have to use gene frequency data for all alleles from both polymorphic and invariant loci (see table 1 of Taylor and Gorman, 1975). Using 33 such alleles, we obtain 0.0213 as the estimate of  $E(F_{ST})$ . If we take  $t = 68$ , the effective population size of the Bermudan population is estimated to be 391—a value considerably smaller than Taylor and Gorman's estimate of 641. On the other hand, if we take  $t = 69$ , we obtain  $N_2 = 397$  which is still substantially smaller than 641.

*Acknowledgements.*—I thank Drs M. Nei, R. Chakraborty, and P. Fuerst for discussions and suggestions on improving the paper. This study was supported by research grants from the U.S. National Institutes of Health and the U.S. National Science Foundation.

#### REFERENCES

- CROW, J. F., AND KIMURA, M. 1970. *An Introduction to Population Genetics Theory*. Harper and Row, New York.
- LEWONTIN, R. C., AND KRAKAUER, J. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, *74*, 175-195.
- NEI, M. 1975. *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam and New York.
- NEI, M., AND CHAKRAVARTI, A. 1977. Drift variances of  $F_{ST}$  and  $G_{ST}$  statistics obtained from a finite number of isolated populations. Submitted for publication.
- TAYLOR, C. E., AND GORMAN, G. C. 1975. Population genetics of a "colonising" lizard: Natural selection for allozyme morphs in *Anolis grahami*. *Heredity*, *35*, 241-247.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, *28*, 114-138.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, *15*, 323-354.