A TEST OF SPATIAL AUTOCORRELATION ANALYSIS USING AN ISOLATION-BY-DISTANCE MODEL[1]

ROBERT R. SOKAL AND DANIEL E. WARTENBERG

Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York 11794

Manuscript received February 25, 1983
Revised copy accepted May 23, 1983

ABSTRACT

Using the isolation-by-distance model as an example, we have examined several assumptions of spatial autocorrelation analysis applied to gene frequency surfaces. Gene frequency surfaces generated by a simulation of Wright's isolation-by-distance model were shown to exhibit spatial autocorrelation, except in the panmictic case. Identical stochastic generating processes result in surfaces with characteristics that are functions of the process parameters, such as parental vagility and neighborhood size. Differences in these parameters are detectable as differences in spatial autocorrelations after only a few generations of the simulations. Separate realizations of processes with identical parameters yield similar spatial correlograms. We have examined the inferences about population structure that could have been made from these observations if they had been real, rather than simulated, populations. From such inferences, we could have drawn conclusions about the presence of selection, migration and drift in given natural systems.

EVOLUTIONISTS believe that they have a satisfactory understanding of the nature of the processes leading to population differentiation at the infraspecific level. They consider mutation, selection, migration and drift, as well as factors related to the organization of the genetic material, sufficient to explain the phenomena of population differentiation and speciation. However, the relative importance of these factors in these microevolutionary processes continues to be the subject of controversy.

Patterns of spatial variation of morphometric and other variables, and especially of electrophoretic variants, must reflect the workings of these processes, and biologists have looked to such patterns to evaluate the relative roles played by the several microevolutionary factors. However, it is difficult to separate the effects of the several putative factors in spatial variation patterns. Some recent attempts in this direction (CAUGANT, JONES and SELANDER 1982; JONES, SELANDER and SCHNELL 1980; SOKAL and RISKA 1981; SOKAL 1983; SOKAL, BIRD and RISKA 1980; SOKAL 1979a; SOKAL and FRIEDLAENDER 1982; SOKAL and MENOZZI 1982; SOKAL and ODEN 1978a,b; and SOKAL and WAR-

[1] Contribution no. 462 in Ecology and Evolution from the State University of New York at Stony Brook.

TENBERG 1981) have been based on spatial autocorrelation analysis and depend on several assumptions. These are (1) spatial variation patterns (spatial response surfaces) can be summarized and characterized by a "signature" obtained through spatial autocorrelation analysis, spectral decomposition and related techniques. (2) Similar deterministic forces result in similar spatial response surfaces. (3) Stochastic processes with the same parameters yield independent and differing spatial surfaces, but these surfaces will have similar signatures suggesting similar generating processes. (4) Changes in these parameters will be reflected by changes in the signatures. (5) All variables studied, especially electrophoretic variants, are unlikely to track the same environmental factor across geographic space, and the ones that do are unlikely to respond in the same way to the same environmental factor. Linkage may result in similar patterns for a few variables. However, commonality of spatial patterns among all biological variables would, therefore, make a selectional interpretation of the observed spatial variation implausible and would suggest a diffusion process (mass migration) instead.

To test these assumptions, we employ a simple population biological process—that of random differentiation with local dispersal but in the absence of selection. A well-known example of such a process in the evolutionary literature is the isolation-by-distance model (WRIGHT 1969). We shall, in this paper, distinguish between *dispersal* and *migration*, using these terms in the ecological sense (RICKLEFFS 1973). The former relates to movement of organisms away from their place of birth, the latter to mass migrations of populations, typically in one direction. In population genetics, dispersal is commonly termed migration, whereas the second process, relevant to some models of population structure, is considered less frequently. In the isolation-by-distance model, individuals mate at random within a defined neighborhood but are constrained from mating with other, more distant members of their species by limitations on their dispersal.

The isolation-by-distance model and the related stepping stone model have been studied by various authors (KIMURA and WEISS 1964; MORTON, MIKI and YEE 1968; MALÉCOT 1969; ROHLF and SCHNELL 1971; MARUYAMA 1977). These models have been related to a number of cases in human and other populations (NEI and IMAIZUMI 1966; CHAKRABORTY 1976; CHAKRABORTY, CHAKRAVARTI and MALHOTRA 1977; WRIGHT 1978). These authors used a variety of statistical approaches to study the behavior of these models, among them genetic correlations, gene identities and the F-statistics. The relationships of these statistics to spatial autocorrelation coefficients have not yet been established. They will not be investigated here since the major purpose of this paper is to test the assumptions of our approach using spatial autocorrelation analysis (SOKAL and ODEN 1978a,b; SOKAL 1979a; SOKAL and WARTENBERG 1981). In this approach, a spatial correlogram (explained in the next section) serves as the signature of a spatial surface.

In this paper we address the first, third and fourth of the assumptions. Specifically, we shall answer the following questions: (1) Does isolation-by-distance result in spatial autocorrelation? (2) Do identical generating processes

with the same parameters yield similar spatial correlograms? (3) Do changes in the parameters of the generating processes (dispersal or vagility in this instance) result in changes of the spatial correlograms? (4) If the results obtained here had been based on empirical data, could they have led to the rejection of one or more plausible mechanisms, such as selection or migration? As a by-product of these analyses, other interesting observations about the isolation-by-distance process emerge, which we also report.

MATERIALS AND METHODS

To generate gene frequency surfaces under an isolation-by-distance model, we employed the Monte Carlo simulation program developed by ROHLF and SCHNELL (1971), who examined the characteristics of spatial surfaces of gene frequencies representing isolation-by-distance models with varying parameters. This program simulates a population of 10,000 individuals arranged on a 100 × 100 lattice, one individual on each grid point. The individuals are considered monoecious diploids capable of self-fertilization. Initially, this grid is settled at random by individuals representing the genotypes for a one-locus two-allele system sampled from a population of gene frequency 0.5 at Hardy-Weinberg equilibrium. Once per generation, each individual is replaced by one offspring from parents chosen at random from within a defined neighborhood centered on that individual. We carried out five sets of experiments. In sets 1 and 2, one of the parents was taken at random from a neighborhood (WRIGHT 1969, p. 295) of size $N = 9$ individuals (which is a 3 × 3 sublattice centering on the individual to be replaced); the other parent was made sessile by designating the individual being replaced as the second parent of the offspring to be placed at the same location. The mean neighborhood size from which parents come at random for sets 1 and 2 is, therefore, 5.0. In set 3, both parents were vagile and were taken at random from a neighborhood of size $N = 9$. In set 4, both parents were again vagile, but with a greater neighborhood size, $N = 25$ (a 5 × 5 sublattice). Finally, in set 5, both parents were vagile with a neighborhood size of $N = 10{,}000$, effectively resulting in panmixia over the entire lattice. The program computes the results of each "mating" as stochastic realizations of the Mendelian expectations for the given cross.

For sets 1 to 4, we carried out five independent runs, each lasting 200 generations. For set 5, the panmictic case, we computed only a single run of 200 generations. As expected, we found no spatial structure in this set, and additional replications seemed unnecessary. In set 2, we initialized the five runs with independent and random settlement of genotypes for each run, but we kept the same pattern of choice of location of the parents of the 200 subsequent generations in all runs. We did this to determine whether a common structure of the spatial mating pattern would make the results of the separate runs more homogeneous than those of set 1, in which the separate runs represent independent random locations of the parents. The design of the experiments can be summarized in the following table:

Set	N	One parent sessile	No. of runs
1	9 (5.0)	Yes	5
2	9 (5.0)	Yes	5[a]
3	9	No	5
4	25	No	5
5	10,000	No	1

[a] Run 1 of set 2 is the same as run 1 of set 1.

We summarized the results of each run in every fifth generation by dividing the 100 × 100 lattice into 400 sublattices of size 5 × 5 and computing the gene frequency for each of these quadrats containing 25 individuals. Additionally, we computed, for every fifth generation, variances of gene frequencies among quadrats and Wright's F-statistic of individuals with respect to the entire lattice (WRIGHT 1969, p. 294). This quantity, computed as $F = (4DR\text{-}H^2)/(4DR\text{-}H^2 + 2H)$,

where D, H and R are the observed proportions of AA, Aa and aa, respectively, corresponds to F_{IT}.

The independence of the gene frequency surfaces of the 20 separate runs is attested to by correlations among the 20 surfaces at generation 200. We computed these correlations by considering the gene frequencies of each surface as a vector of 400 elements and calculating its product-moment correlation with other similar vectors representing the other surfaces. None of the correlations is high (the maximum is 0.262), and average correlations within sets and between sets are very low (see Table 1). No pattern of relationships differentiating correlations within and between sets emerges. Also, we can find no evidence from this table (or from an examination of the replicated correlograms representing runs in each set, discussed later) that the common spatial mating pattern in the runs of set 2 makes the surfaces of that set more homogeneous than those of other sets. Note that the average correlation of the surfaces within set 2 is not consistently higher than that within the other sets. For purposes of what follows, we shall, therefore, consider set 2 as a replication of set 1.

For each run, we graphed gene frequency surfaces represented by 20 × 20 grids every fifth generation starting at generation 0 and terminating at generation 200 and examined them visually for structure. To summarize the very large number of maps (820 in all) we employed the method of spatial autocorrelation (CLIFF and ORD 1981; SOKAL and ODEN 1978a,b). The set of localities required for the computation of spatial autocorrelation is the 400 quadrats of the 20 × 20 grid. The details of the computation of spatial autocorrelation are furnished by CLIFF and ORD (1981) and, in simplified form, by SOKAL and ODEN (1978a).

We computed both Moran's I, a product-moment coefficient, and Geary's c, a distance-type coefficient. Moran's coefficient is computed as

$$I = n \sum_{ij} w_{ij} z_i z_j / W \sum_{i=1}^{n} z_i^2,$$

where n is the number of localities in the study; $\sum_{ij}$ indicates summation over all i from 1 to n and over all j from 1 to n, $i \neq j$; w_{ij} is the weight given to an edge between localities i and j (w_{ij} need not equal w_{ji}); $z_i = Y - \bar{Y}$, where Y_i is the value of variable Y for locality i, and $\bar{Y}$ is the mean of Y for all localities; and $W = \sum_{ij} w_{ij}$, the sum of the matrix of weights (except for the diagonal entries, if any). For large sample sizes this coefficient ranges from −1 to +1. The formula for Geary's coefficient is

$$c = (n - 1) \sum_{ij} w_{ij} (Y_i - Y_j)^2 / 2W \sum_{i=1}^{n} z_i^2.$$

All terms in this formula have already been explained. Geary's c will range from zero, for perfect positive autocorrelation, to an unbounded positive value for negative autocorrelation, the expected value in the absence of autocorrelation being one. Since the results obtained with these two coefficients are quite similar, we report here only on the results with Moran's I.

We calculated Euclidean distances between the centers of these quadrats and grouped the distances between all pairs of quadrats in a frequency distribution with eight classes with unequal class intervals. The first five classes had an interval of 1.0 and are represented by the class marks: 1, 2, 3, 4, 5; the succeeding three classes ranged from 5.5 to 10.5, from 10.5 to 20.5 and from 20.5 to 30.5 and are represented by class marks 8.0, 15.5 and 25.5, all in quadrat units. Then we calculated the autocorrelation coefficients for each of the distance classes.

The graph of spatial autocorrelation coefficients against distance classes is known as a spatial correlogram. Most of the spatial autocorrelation coefficients (138 of a possible 152) are statistically significant when tested against the null hypothesis that sample I does not differ from its expected value which is $-1/(n - 1)$, using the test procedure given by CLIFF and ORD (1981). In this formula, n is the number of points for which the autocorrelation is being computed. The statistical significance of most coefficients, even when these are as low as 0.01, is due to the unusually large number of grid points ($n = 400$) in this study. In the analyses that follow, we emphasize results for the first five distance classes, since this is where the effects of the isolation-by-distance process are most noticeable. We carried out the computations for spatial autocorrelation by the SAAP program developed by D. E. W.

TABLE 1

Average pairwise correlations between surfaces at generation 200, within and between sets

Set no.	Within set	Between it and other sets
1	−0.008	−0.026
2	0.034	−0.023
3	0.002	0.007
4	0.056	−0.004
5	−0.008	0.006

CHANGE OVER SPACE

The isolation-by-distance model results in statistically significant spatial autocorrelations. Figure 1 shows the average correlogram for each set over the first five distance classes at generation 200. The expected values for the case of no autocorrelation are $-1/(n-1)$, which, in the present instance, equals −0.0025. This is so close to zero that the difference would be imperceptible and, hence, impractical to indicate in this figure. Most spatial autocorrelation coefficients up to distance class 4 in sets 1 through 4 are positive and substantial to moderate in magnitude, and almost all are highly significant. Average autocorrelation for distance class 5 is small (0.021). These results indicate areas of homogeneity representing the hills or valleys of the gene frequency surfaces. These are caused by inbreeding due to the limited vagility of each individual; matings in any one area tend to be among relatives, increasing the homogeneity of local areas. By contrast, the correlogram resulting from set 5, the panmictic example, has very low autocorrelation coefficients, close to the expected values. Since, for set 5, the correlogram at any one generation is fully independent of that at a previous generation, we chose five replicates at random from the single run of set 5 for computing a mean correlogram for this design.

Because spatial autocorrelations for different distance classes in one correlogram are not independent of each other, the significance of an entire correlogram cannot be computed directly from the significance levels of the individual coefficients. That is, we cannot infer, from the results of the separate significance tests for each autocorrelation coefficient, the correctness of the null hypothesis that sample values of I for all distance classes do not differ from the expected value $-1/(n-1)$. Rather, we must employ a significance test for correlograms against the null hypothesis of no spatial autocorrelation such as those developed by ODEN (1983) or KOOIJMAN (1976). When Oden's Q-test was used, every correlogram, except that for set 5, was significant at $P < 0.0005$. Thus, significant spatial autocorrelation has been generated by the isolation-by-distance process in all sets except the panmictic one, and we may consider the average correlograms of Figure 1 to deviate significantly from the null hypothesis of no pattern in the gene frequency surfaces.

Figure 1 indicates differences among some of the average correlograms of

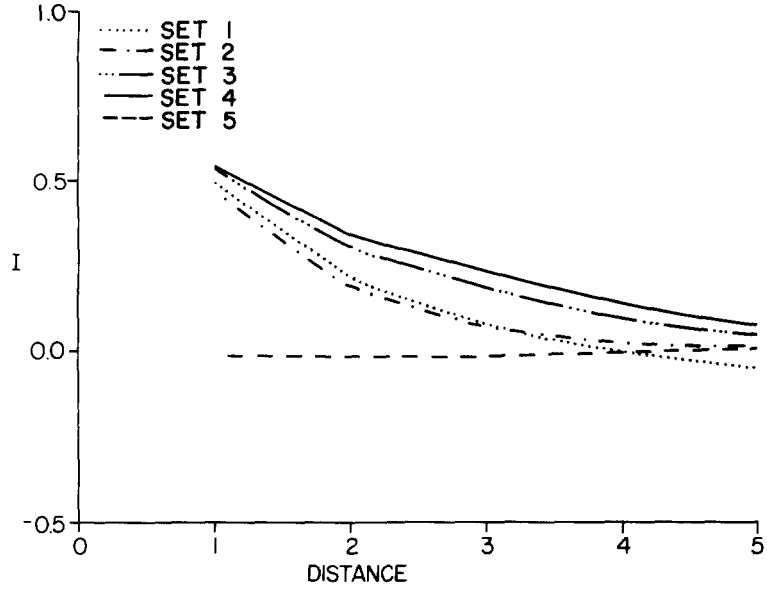

FIGURE 1.—Average spatial correlograms summarizing results at generation 200 for five different sets of parameters. Each set is based on five replicates. Sets 1 and 2, males taken at random from a neighborhood of size $N = 9$, females sessile. Set 3, both sexes vagile, taken at random from a neighborhood of size $N = 9$. Set 4, both sexes vagile, taken at random from a neighborhood of size $N = 25$. Set 5, both sexes vagile, taken at random from the entire population (neighborhood of size $N = 10{,}000$; panmixia). Set 5 was not replicated, and the average presented here is of five randomly chosen generations from the single run of this set. Ordinate: spatial autocorrelation coefficient (Moran's I). Abscissa: the first five distance classes representing class intervals of one quadrat unit with an upper bound of 5.5 such units for class 5.

the five sets. Before we examine these differences in detail, we shall first inspect the replicability of the results for separate runs with identical parameters. Figure 2 illustrates the results at generation 200 for each run in each set. In this figure we feature correlograms for all eight distance classes to permit readers to note the generally low autocorrelation coefficients starting at the fifth distance class (5 quadrat units). The correlograms depart appreciably from expectation for some runs at the eighth distance class (25.5 quadrat units). In the four sets other than the panmictic set 5, these departures increase with neighborhood size, being greatest in set 4. We attribute this phenomenon to the "edge effect" already noted by ROHLF and SCHNELL (1971). All of the largest distances in the 20 × 20 surfaces are between quadrats along the margin of the surface. These areas in the lattice are more limited in choice of parents for the next generation than are more central locations; hence, the neighborhoods of these areas effectively are smaller than their nominal values. This effective reduction in neighborhood size produces greater effects in set 4 than in set 1, which already has small homogeneous patches, even in the central portion of its area. The departures from expectation are caused by chance combinations of like or unlike patches forming along the periphery of the 20 × 20 surface. These combinations lack permanence (they usually last

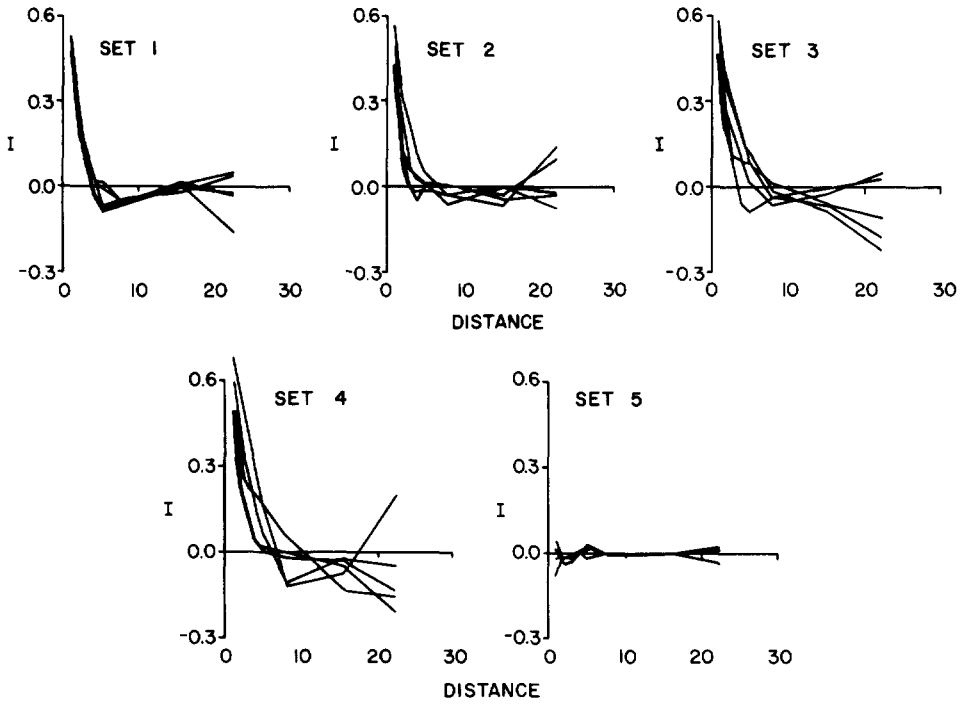

FIGURE 2.—Spatial correlograms representing replicate surfaces at generation 200 for five different sets of parameters. For definition of sets, see legend to Figure 1 and text. Set 5 was not replicated, and the average presented here is of five randomly chosen generations from the single run of this set. Ordinate: spatial autocorrelation coefficient (Moran's *I*). Abscissa: distance in quadrat units. Note that the intervals of the distance classes are unequal, as described in the text. All eight distance classes are shown here to furnish a summarization of the entire surface.

for 20 to 30 generations), although they are statistically significant at the time of their appearance. We could have avoided edge effects by connecting the edges of the lattice in such a way as to make it into an unbounded surface. However, we did not do so because, given the size of the lattice, edge effects are relatively unimportant, and, also, because we wanted our results to resemble those that might be obtained in real populations.

The scatter among the correlograms of the five replicates in each set is least for set 5 and greatest for set 4 (see Figure 2). Although individual correlograms from different sets overlap, the general trends are sufficiently differentiated so as to make the mean correlograms in Figure 1 quite distinct (except those of sets 1 and 2, as already mentioned).

To test whether the replicated runs of each set were more similar to each other than the results for the different sets, we calculated the area between each pair of correlograms over the first five distance classes as an index, *A*, of agreement. We then compared pairwise *A* values within the sets with 230 pairwise values representing runs in different sets. Such a test can be carried out by the Mantel test (MANTEL 1967; SOKAL 1979b) for which one calculates the observed and expected values of a quantity *Z*, which is the sum of the

products of the elements of the matrix of pairwise areas with the elements of another design matrix describing the intended comparison. In the Mantel test, one computes the permutational variance of Z and tests the departure from expectation as a normal deviate, relying upon the asymptotic normality of the distribution of the test statistic. However, in this example, because of the relatively small sample size, we chose not to assume asymptotic normality of the test statistic but, rather, evaluated the probability of the observed departure based on a Monte Carlo simulation of 250 random permutations of the matrix of pairwise area values.

As a preliminary test, we examined whether set 2 with fixed location of mating parents differed from set 1 where such locations were random. The observed differences within each set are not significantly less than those between the two sets ($P = 0.100$). A Mantel test of whether the scatter among runs of set 2 is less than among set 1 also proved not significant ($P = 0.744$). For this reason, we merged sets 1 and 2 in all subsequent Mantel test comparisons to form one set of nine runs, designated set 1 + 2.

The design of the experiment suggested five tests. For an experimentwise error rate of 0.05, the Dunn-Šidák method (SOKAL and ROHLF 1981) provides a significance level of 0.0102 for each individual test. The observed differences within sets 1 + 2, 3, 4 and 5 are considerably and highly significantly less than those among sets. The Mantel test yields $P < 0.00005$. There are appreciable differences between the significant correlograms of sets 1 + 2, 3 and 4 from those of set 5, in which there is no spatial structure ($P < 0.00005$). When we omitted set 5, the remaining three isolation-by-distance sets differed appreciably, but not quite significantly, among each other ($P < 0.012$). To test the effect of increasing neighborhood size from 5.0 to 9 (changing one parent from being sessile to being vagile in a neighborhood of size 9), we compared set 1 + 2 with set 3 and obtained an appreciable difference, which, however, is not significant by the Dunn-Šidák method ($P < 0.028$). A further test of increasing neighborhood size from 9 to 25 (set 3 *vs.* set 4) yielded no significant difference ($P = 0.924$). The last result is surprising in view of the apparent difference of the mean correlograms of these two sets, but is due to considerable variation among runs within each set.

Different parameters of the isolation-by-distance model result in distinctly different gene frequency surfaces, as already pointed out by ROHLF and SCHNELL (1971). These authors found that, whereas the panmictic model (our set 5) produced no local differentiation, no inbreeding and no increased variance of gene frequencies among quadrats over that expected within quadrats, for all other combinations of parameters, each of these variables increased as neighborhood size decreased, and within each neighborhood size as one of the parents became sessile. Thus, for a neighborhood size of nine with one sessile parent, the gene frequency surfaces became quite coarse as early as generation 30, with further increases in the unevenness of the surface texture at subsequent generations. ROHLF and SCHNELL (1971) carried out the simulations for 120 generations. Our simulations for 200 generations further accentuated the differences they observed.

At generation 200, there is a nonlinear relationship between fixation indexes

F and departure from the expectation of $-1/(n-1)$ for the spatial correlograms shown in Figure 1 (or between the latter and variances in gene frequency among quadrats; both relations shown in Table 2). As neighborhood size increases, the fixation index for the entire lattice (and the gene frequency variance among quadrats) decreases, yet the autocorrelation is greatest at intermediate neighborhood values, as shown by set 4. Thus, the autocorrelation function increases with neighborhood size up to a threshold, beyond which it decreases again. In other words, similarity between nearby quadrats is low when the neighborhood is either very small or very large. This phenomenon must be related to the size of homogeneous areas or patches on the gene frequency surface.

The X-intercept for a surface is that distance at which a spatial correlogram turns negative or zero. This quantity was termed patch size by SOKAL (1979a) but is given a more neutral appellation here to avoid confusion with the connotation of patch size in ecological theory. In square areas, the X-intercept closely approximates the length of one side (SOKAL, 1979a). For sets 1 and 2, the X-intercept is at a distance near 4 quadrat units, whereas for sets 3 and 4, with larger effective neighborhoods, it is greater than 5 quadrat units (Table 2). When tested by a Kruskal-Wallis test, these differences in X-intercepts among the sets can be shown to be significant (adjusted $H = 8.300$, $P = 0.0158$). Set 1 + 2 has significantly lower X-intercepts than sets 3 and 4, but the later two cannot be shown to differ significantly.

Why should the autocorrelation coefficients be higher in the surfaces with larger neighborhood sizes and lower overall fixation and variance? For any distance class, proportionately more of the point pairs within this distance class occur in the same area when homogeneous areas are larger. A higher proportion of within-area distances for a given distance class results in higher spatial autocorrelation for that class. In other words, the surfaces of sets 1 and 2 have formed, on the average, 25 homogeneous areas, whereas those of sets 3 and 4 have formed, on the average, fewer than 16. The fewer, hence larger, areas result in higher spatial autocorrelations. Yet, as areas grow still larger, loss of distinctness (contrast) among areas overcomes the effect of area size, and the autocorrelation decreases again, as we have seen. These relations will depend on the size of the homogeneous areas that result from the population genetic process, the size of the sampling area (quadrat size) and the overall size of the area studied. The X-intercept is always relative to the total area of study and cannot easily be compared across studies involving unequal geographic distances or different connection weights.

CHANGE OVER TIME

Without introducing recurrent mutation, this process can achieve temporal stationarity only at fixation or loss of an allele for the entire population. However, it would take a very long time (on the average in the thousands of generations) to reach either of these two absorbing states. ROHLF and SCHNELL (1971) found that the fixation indexes of the surfaces they studied were still increasing after 120 generations of simulation. Our graphs of the fixation

TABLE 2

Summary of differences among sets at generation 200

	Set				
	1 ♀ Sessile	2 ♀ Sessile	3 ♀ Vagile	4 ♀ Vagile	5 ♀ Vagile
N	5.0	5.0	9	25	10^4
F	0.302	0.294	0.257	0.106	−0.012
Variance	0.06320	0.05948	0.04868	0.02592	0.00480
X-intercept	4.0	4.2	5.4	5.8	Undefined
MD	0.166	0.152	0.234	0.267	0.011
$\bar{I}_1$	0.490	0.466	0.533	0.541	−0.010
$\bar{I}_2$	0.211	0.180	0.294	0.339	−0.016
$r_{t\text{-}5}$	0.916	0.911	0.858	0.721	0.156

N = neighborhood, population size of area from which parents of central individuals are drawn at random; F = average observed fixation index F_{IT} at generation 200; Variance = average gene frequency variance among quadrats at generation 200; X-intercept = average of X-intercepts at generation 200; MD = average Manhattan distance over five distance classes of spatial correlogram at generation 200 against expectation of −0.0025; $\bar{I}_1$, $\bar{I}_2$ = average spatial autocorrelation coefficient I for distance classes 1 and 2, respectively, at generation 200; $r_{t\text{-}5}$ = temporal autocorrelation, maximum of correlations between successive pairs of surfaces. These surfaces were taken at five-generation intervals from generation 0 to 200, and the correlations between them were averaged over the runs of each set.

indexes (Figure 3) and of the variances among quadrats (Figure 4) indicate that, even after 200 generations, these curves have not yet reached an asymptote. Nevertheless, the rapid rise in these parameters for the first 50 generations slowed considerably during later generations in the four sets with limited neighborhood sizes, because, in these later generations, the rate of decay of heterozygosity is so small. Empirically, we find autocorrelation structure to be quite stable beginning with generation 150 in all runs of these sets. Variances for sets 1 and 2 fluctuate around 0.06, whereas those for sets 3 and 4 fluctuate around 0.04 and 0.02, respectively.

By examining correlograms for each run at five-generation intervals, we were able to study their development over time. Figure 5 shows the development of the correlograms over time for all runs of set 4. By generation 5, correlograms were already significant, and the average X-intercept is 2.2 quadrat units. Homogeneous areas had already begun to form. By generation 10, the average X-intercept had increased to 2.4, rising further to 5.2, 5.2 and 5.8 for generations 50, 100 and 200, respectively. Thus, the intercept by and after generation 50 appears to be more than 5 quadrat units, leading us to infer homogeneous areas equivalent to a square area of more than five quadrat units (or 25 lattice units) along one side. These homogeneous areas are more than 25 times that of the neighborhood.

In set 3, spatial autocorrelation also develops virtually instantaneously. The average X-intercept is below that of set 4, being 2.0, 2.0, 4.2, 4.2 and 5.4 for 5, 10, 50, 100 and 200 generations, respectively. However, by generation 200, the inferred average size of the homogeneous areas is nearly that of set 4, although the neighborhood of set 3 is only 9, making for a homogeneous area

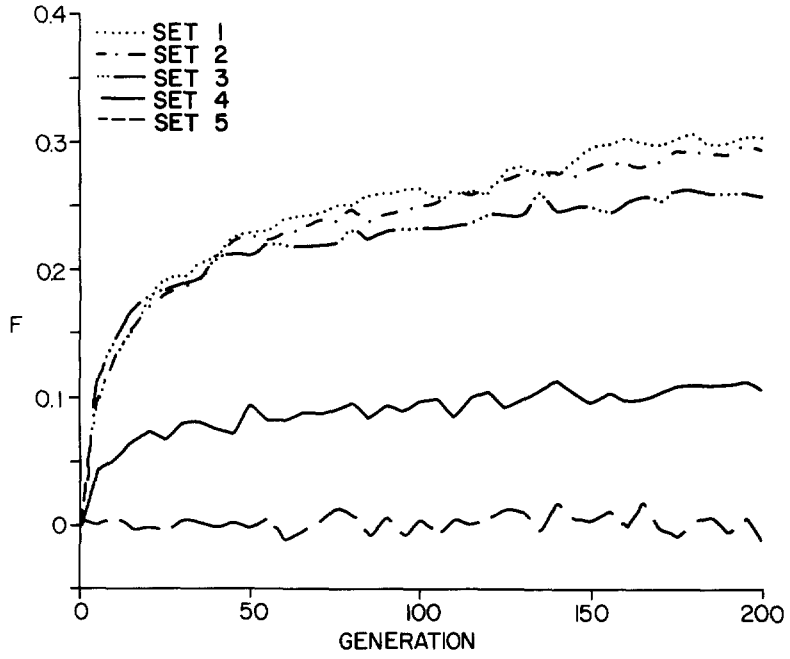

FIGURE 3.—Observed changes in F, the fixation index for the entire lattice, as a function of generation number. Mean values for the five runs of each set are furnished for every fifth generation. Set 5 is based on only a single run.

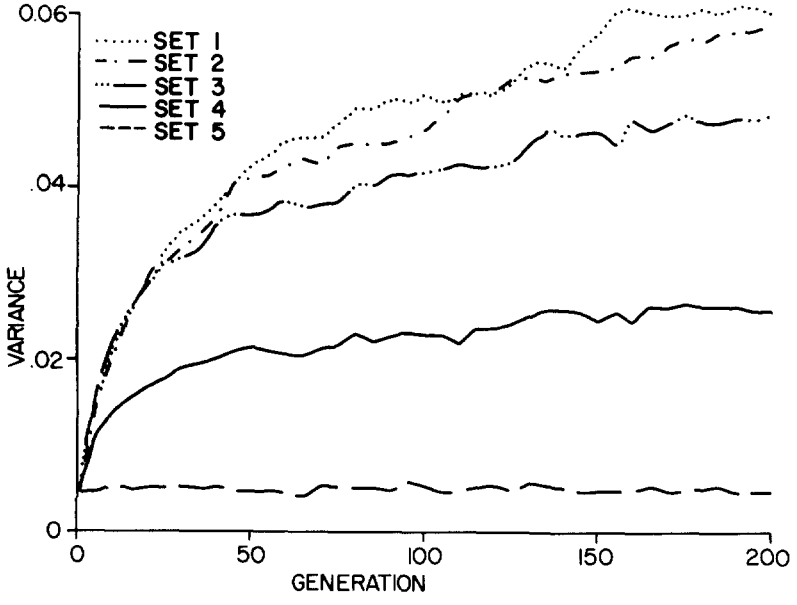

FIGURE 4.—Observed changes in variance of gene frequency among quadrats as a function of generation number. Mean values for the five runs of each set are furnished for every fifth generation. Set 5 is based on only a single run.

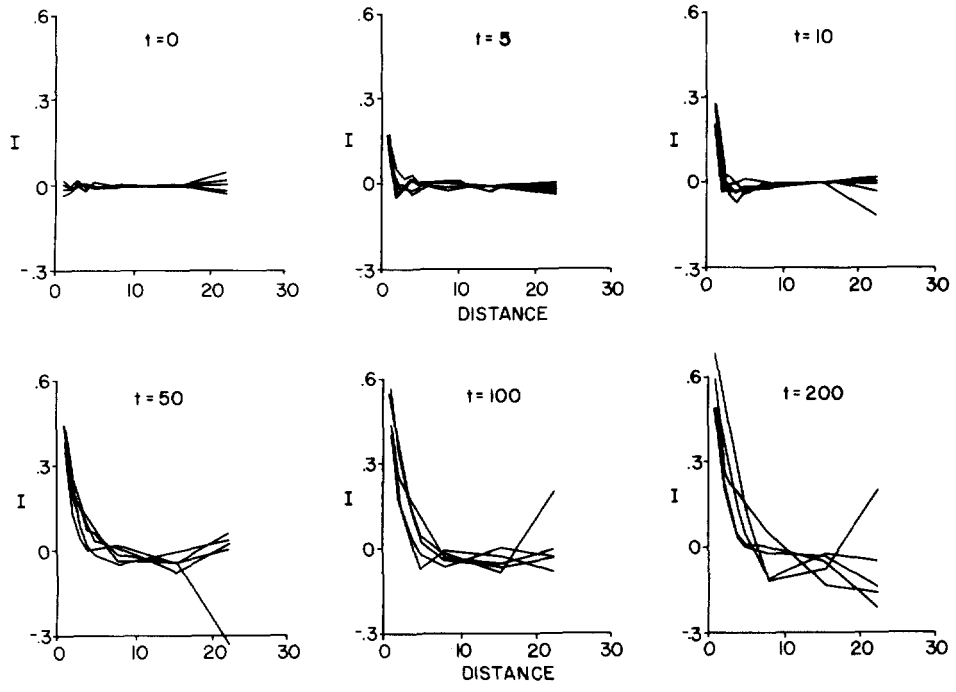

FIGURE 5.—Changes in gene frequency surfaces over time as shown by spatial correlograms. The five replicates of set 4 (both sexes vagile, taken at random from a neighborhood of size $N = 25$) are shown at six generations ranging from 0 to 200. Ordinate: spatial autocorrelation coefficient (Moran's I). Abscissa: distances in quadrat units. Note that the intervals of the distance classes are unequal, as described in the text. All eight distance classes are shown here to furnish a summarization of the entire surface.

to neighborhood ratio of 81. The eventual similarity of the X-intercepts by generation 200 for sets 3 and 4 is corroborated by their similar average correlograms for which we could not establish statistically significant differences.

Figure 6 features the average correlograms for the five sets at different generation times. The sets with different parameters differ at all generation times after the inception of the isolation-by-distance process. Thus, given adequate replication, one can distinguish different parameter sets at a very early stage (by generation 5) in the development of the isolation-by-distance pattern.

It is of interest to study the temporal autocorrelations of the surfaces. We can estimate these as matrix correlations (SNEATH and SOKAL 1973) between pairs of successive gene frequency surfaces represented by 400 quadrats at five-generation intervals. These correlations are shown in diagrammatic form, averaged over the five runs, for sets 1 and 4 in Figure 7. As we expected in a phenomenon that shows both spatial and temporal autocorrelation, there is a gradual decrease of temporal autocorrelation between the surfaces as the time lag between them increases. There is also a tendency, at any constant lag, for temporal autocorrelation to rise as the number of generations increases. As time goes on, homogeneous areas increase in extent but also in amplitude

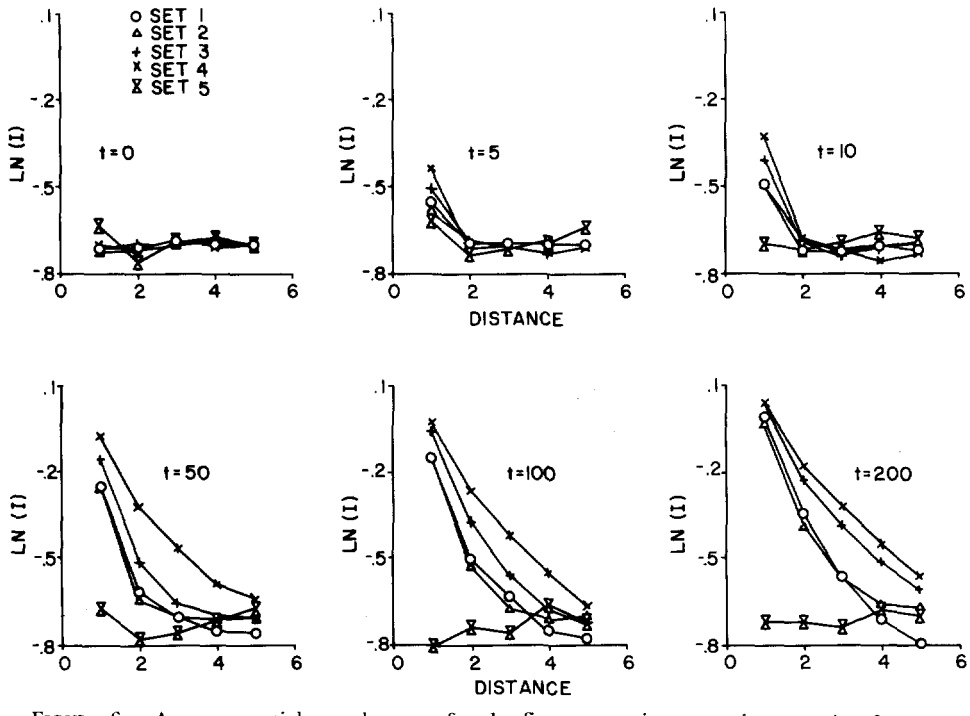

FIGURE 6.—Average spatial correlograms for the five sets at six generations ranging from 0 to 200. Set 5 is based on only a single run. Ordinate: spatial autocorrelation coefficient (Moran's *I*) transformed to natural logarithms to emphasize differences among sets. Abscissa: distances in quadrat units. Note that the intervals of the distance classes are unequal, as described in the text. All eight distance classes are shown here to furnish a summarization of the entire surface.

(maximal departure from the average gene frequency). This increase in relief of the gene frequency surfaces results in areas of near or complete fixation or loss that remain relatively stable for substantial periods of time. Thus, temporal autocorrelations are higher in set 1 than in set 4, even though the neighborhood and homogeneous areas are greater in the latter, because set 1 has a higher rate of fixation and, hence, greater surface relief.

DISCUSSION

We now return to the four questions posed in the introduction and note the information concerning them provided by the results reported in the preceding section. (1) Does isolation-by-distance result in spatial autocorrelation? The answer is clearly yes as shown in Figures 1, 2, 5 and 6 and as substantiated by significance tests of individual spatial autocorrelation coefficients and of entire correlograms.

The similarity of spatial correlograms resulting from identical generating processes (question 2) and the changes engendered in spatial correlograms by changes in the parameters of the generating processes (question 3) are illustrated in Figures 2 and 1, respectively, and generally substantiated by the

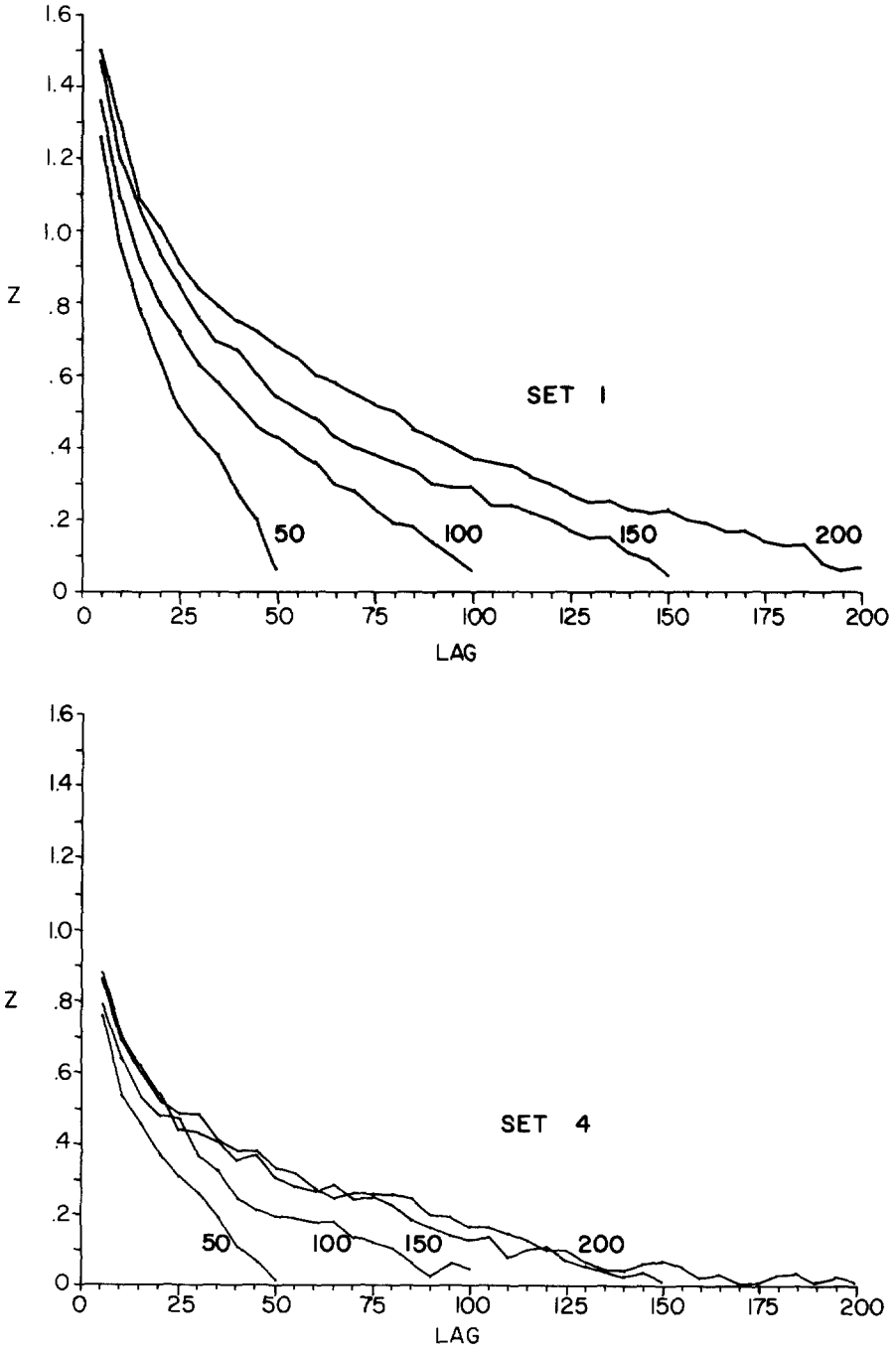

FIGURE 7.—Mean temporal autocorrelation of gene frequency surfaces plotted against lag time for sets 1 and 4. Four separate lines are shown for surfaces at generations 50, 100, 150 and 200. Ordinate: autocorrelations given in z-transformation to emphasize differences among generations for low values of lag. Abscissa: lag time in generations.

Mantel tests. Would differences among correlograms be large enough to reveal important variation in neighborhood size when different studies are compared? The replication available in this study was sufficient to distinguish differences in dispersal resulting in an approximate doubling of neighborhood size in one case (compare the neighborhoods of sets 1 + 2, with those of set 3, but note that changes in correlograms brought about by an even greater increase in neighborhood size between sets 3 and 4 were less and could not be shown to be significant). In nature, we might expect greater differences in neighborhood sizes among populations being compared by such a method and, thus, could expect significantly differing correlograms.

The relations among the various parameters describing the surfaces are summarized in Table 2. As the neighborhood size becomes larger, the mean observed fixation index F decreases from approximately 0.3 to near 0, and the variance of gene frequencies among quadrats, each based on 25 individuals or 50 genes, decreases from 0.0603 to 0.0048—quite close to the expected value of 0.005 under random mating. Thus, the smaller neighborhoods result in a higher degree of inbreeding which, in turn, causes near homozygosity and gene frequencies at, or near, loss or fixation in some of the quadrats. In the larger neighborhood sizes, the lower degree of inbreeding dampens the peaks and troughs of the gene frequency surface. However, because the decreased inbreeding is brought about by an increase in dispersal of the organisms, the peaks and troughs have wider diameters in surfaces with greater neighborhoods. Because this results in more observations becoming alike at the short distances, this tends to increase the autocorrelation. This trend is reflected in the magnitude of the departure of spatial correlograms from the expectation under the null hypothesis that each value of $I = -1/(n - 1)$. This departure from expectation is quantified as average absolute distances (Manhattan distances; SNEATH and SOKAL 1973) over the first five distance classes (Table 2). Thus, whereas for sets 1 and 2 these distances are 0.166 and 0.152, they are 0.234 and 0.267 for sets 3 and 4. However, in set 5 the entire 100×100 lattice has become the neighborhood and, consequently, there is no spatial autocorrelation whatsoever, as indicated by the minute departure from expectation, the average Manhattan distance of the correlogram over the first five autocorrelation classes being 0.011.

Another way of illustrating that an increase in neighborhood size results in an increase in the diameter of the peaks and troughs is the increase in the X-intercept shown for sets 1 through 4 in Table 2. The X-intercept cannot be estimated for the panmictic set, which constitutes a single homogeneous area, and would require embedding in a larger lattice to define it. The actual values of autocorrelation coefficients at distance class 1 and 2 are higher in the areas with larger neighborhoods and larger homogeneous areas, because proportionately more of the connections at distances less than the diameter of these areas are between similar gene frequency values. At still greater neighborhood sizes, the diameters of the homogeneous areas would increase, but their amplitude would decrease, resulting once more in a lower autocorrelation coefficient.

Note that the maximal temporal autocorrelation (averaged within sets) be-

tween surfaces for a lag of five generations is highest in the surfaces with the smallest neighborhoods (it is 0.916 in set 1 and set 2, and decreases to 0.858 for set 3 and to 0.721 for set 4). As expected, the panmictic set has essentially no temporal autocorrelation between surfaces. Its maximal value for a lag of five generations is only 0.156. Higher temporal autocorrelation for smaller neighborhoods can be explained because small neighborhoods result in very high peaks or deep troughs, *i.e.*, portions of the surface are near fixation or loss. Such regions of high relief tend to remain more stable than the wider peaks and troughs of larger neighborhoods, which show less relief. Because the standard errors of binomial frequencies near fixation or loss are very small, the probabilities of change in areas of high relief are diminished. Once a really high but narrow peak (or deep but narrow trough) has become established, a considerable amount of time is required for the joint action of chance sampling of the genome and of vagility to alter it. Note that this is not a property of smaller neighborhoods *per se* but rather a property of portions of the surface that are at, or near, fixation or loss.

Thus, when describing surfaces of gene frequencies or of morphometric variables, one must consider many parameters. For gene frequencies, it is insufficient to examine only their variance or the fixation index. Because surfaces can take on so many different shapes and forms and these can indeed reveal different population biological processes, one must characterize the surfaces in as many different ways as possible. Spatial correlograms are only limited characterizations of surface structure. As we have seen, correlations between surfaces for different gene frequencies and time periods are also necessary. In other cases (*e.g.*, SOKAL and MENOZZI 1982), correlograms need to be tailored to test directional trends or other special hypotheses.

Can the findings of this study be extended to serve as a basis for inference from empirically obtained data (question 4)? If, in a study of allozyme frequencies, we had obtained results such as those that were recorded in the replicates of any one set, spatial autocorrelation analysis would have yielded no similarity (correlation) between the surfaces (Table 1) and considerable similarity between the correlograms (Figure 2). Given such data for different loci in the same population, and following the line of reasoning advanced by SOKAL and ODEN (1978b), SOKAL (1979a) and SOKAL and WARTENBERG (1981), we would have concluded that this combination of dissimilar surfaces and similar correlograms suggested a stochastic process with the same parameters. Thus, the present study is consistent with this particular model. In such a data set, selection would be an improbable interpretation for the observations. Dissimilar surfaces imply dissimilar selective agents and, except for the quite improbable case of different selective agents having the same autocorrelation pattern, one would expect these dissimilar surfaces to have dissimilar correlograms, which they do not have. Furthermore, more than one locus might well be tracking a single selective agent, which should lead to similarity in small groups of surfaces and correlograms. This is not the case. Large scale migration can also be eliminated, because such a process involving populations with different initial gene frequencies would result in similar surfaces for most loci, which was not the case in this study.

But what conclusions could be obtained if the investigator, on examining some acutal data, encountered results such as those obtained from replicates of different sets in this study? In the simplest imaginable case, the investigator might study two allozyme frequencies in the same population with correlograms that differed as do replicates from sets 1 and 4. The surfaces would still be uncorrelated, but now the correlograms would differ appreciably. Isolation-by-distance could not be responsible for both correlograms since it would affect both loci equally. A more plausible explanation would be that there are two spatially autocorrelated environmental variables, with different patch sizes, being tracked by two different loci. Alternative explanations are possible here, but such a combination of observations would at least eliminate the hypothesis of chance fluctuations, which would be an important conclusion.

In relation to this argument, there remains the problem of the lack of a significance test for the difference between two correlograms. Although a test for the significance of individual correlograms has been developed (ODEN 1983), there is no corresponding test for the difference between a pair of correlograms (or among the members of a sheaf of correlograms). The test furnished earlier in the manuscript is based on replicated results for each of the sets. When we carry out such a test for the five replicates of set 1 *vs.* the five replicates of set 4, there is a significant difference between them by the Mantel test ($P = 0.012$). However, this test cannot be used for comparing two correlograms—one sampled from each of the sets. We are currently working on a significance test for the difference between two correlograms. If this can be developed with sufficient power, then one might be able to pick up fairly subtle differences between correlograms. However, even if such a test existed, it probably would not be successful here since variability among replicates, especially of set 4, is high. This is possibly due to edge effects in set 4 being more prominent in this simulation with larger neighborhood size. If we had made the lattice unbounded, or if in a field situation the patch size had been imbedded in a much larger population, such edge effects might not have occurred. This would, therefore, have given us higher replicability, which would have made it much simpler to distinguish the two types of correlograms. In other cases, moreover, it is reasonable to expect that strong selection pressures would lead to radically different correlograms, as, indeed, data analytic experience in our laboratory with various data sets has shown. Such intrapopulation heterogeneity is, of course, not limited to correlograms. A similar heterogeneity among samples of gene frequencies of two different populations would make it difficult to substantiate differences between two single samples, one from each population.

It is clear from the preceding discussion that this method for analyzing gene frequency surfaces is not a magic key for unlocking the secrets of population structure. It does, however, permit the investigator to ascertain the presence of necessary, but not sufficient, conditions for the operation of certain evolutionary factors. Thus, one can eliminate certain forces as impossible or at least improbable. Although this is not proof that the remaining forces are, indeed, responsible for the structure as observed, it does represent progress in our understanding of population structure.

We are indebted to F. JAMES ROHLF and NEAL ODEN for technical advice during the research. These colleagues, LEV R. GINZBURG, MASATOSHI NEI and two anonymous reviewers furnished useful criticisms of the manuscript. BARBARA THOMSON provided technical assistance, DIANE NAYLOR did the artwork and BARBARA MCKAY wordprocessed the typescript. Research supported by grant GM 2826202 from the National Institute of General Medical Sciences.

LITERATURE CITED

CAUGANT, D., J. S. JONES and R. K. SELANDER, 1982 Morphological and molecular polymorphism in *Cepaea nemoralis* in the Spanish Pyrenees. Genetica **57:** 177–191.

CHAKRABORTY, R., 1976 Cultural, language and geographical correlates of genetic variability in Andean highland Indians. Nature **264:** 350–352.

CHAKRABORTY, R., A. CHAKRAVARTI and K. C. MALHOTRA, 1977 Variation in allele frequencies among caste groups of the Dhangars of Maharashtra, India: an analysis of Wright's F_{ST} statistic. Ann. Hum. Biol. **4:** 275–280.

CLIFF, A. D. and J. K. ORD, 1981 *Spatial Processes: Models and Applications.* Pion, London.

JONES, J. S., R. K. SELANDER and G. D. SCHNELL, 1980 Patterns of morphological and molecular polymorphism in the land snail *Cepaea nemoralis.* Biol. J. Linn. Soc. **14:** 359–387.

KIMURA, M. and G. H. WEISS, 1964 The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics **49:** 561–576.

KOOIJMAN, S. A. L. M., 1976 Some remarks on the statistical analysis of grids especially with respect to ecology. Ann. Syst. Res. **5:** 113–132.

MALÉCOT, G, 1969 *The Mathematics of Heredity.* W. H. Freeman and Company, San Francisco.

MANTEL, N, 1967 The detection of disease clustering and a generalized regression approach. Cancer. Res. **27:** 209–220.

MARUYAMA, T, 1977 *Stochastic Problems in Population Genetics.* Lecture Notes in Biomathematics 17. Springer-Verlag, Berlin.

MORTON, N. E., C. MIKI and S. YEE, 1968 Bioassay of population structure under isolation by distance. Am. J. Hum. Genet. **20:** 411–419.

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.

ODEN, N. L., 1983 Assessing the significance of a spatial correlogram. Geogr. Anal. In press.

RICKLEFFS, R. E., 1973 *Ecology.* Chiron Press, Newton, Massachusetts.

ROHLF, F. J. and G. D. SCHNELL, 1971 An investigation of the isolation by distance model. Am. Nat. **105:** 295–324.

SNEATH, P. H. A. and R. R. SOKAL, 1973 *Numerical Taxonomy.* W. H. Freeman and Company, San Francisco.

SOKAL, R. R, 1979a Ecological parameters inferred from spatial correlograms. pp. 167–196. In: *Contemporary Quantitative Ecology and Related Ecometrics,* Edited by G. P. PATIL and M. L. ROSENZWEIG. International Cooperative Publishing House, Fairland, Maryland.

SOKAL, R. R, 1979b Testing statistical significance of geographic variation patterns. Syst. Zool. **28:** 227–231.

SOKAL, R. R., 1983 Analyzing character variation in geographic space. pp. 384–403. In: *Numerical Taxonomy,* Edited by J. FELSENSTEIN. Springer-Verlag, New York.

SOKAL, R. R., J. BIRD and B. RISKA, 1980 Geographic variation in *Pemphigus populicaulis* (Insecta: Aphididae) in eastern North America. Biol. J. Linn. Soc. **14:** 163–200.

SOKAL, R. R. and J. FRIEDLAENDER, 1982 Spatial autocorrelation analysis of biological variation on Bougainville Island. pp. 205–227. In: *Current Developments in Anthropological Genetics: Pop-*

ulation Structure and Ecology, Vol. II, Edited by M. H. CRAWFORD and J. H. MIELKE. Plenum, New York.

SOKAL, R. R. and P. MENOZZI, 1982 Spatial autocorrelations of HLA frequencies in Europe support demic diffusion of early farmers. Am. Nat. **119:** 1–17.

SOKAL, R. R. and N. L. ODEN, 1978a Spatial autocorrelation in biology. 1. Methodology. Biol. J. Linn. Soc. **10:** 199–228.

SOKAL, R. R. and N. L. ODEN, 1978b Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. Biol. J. Linn. Soc. **10:** 229–249.

SOKAL, R. R. and B. RISKA, 1981 Geographic variation in *Pemphigus populitransversus* (Insecta: Aphididae). Biol. J. Linn. Soc. **15:** 201–233.

SOKAL, R. R. and F. J. ROHLF, 1981 *Biometry,* Ed. 2. W. H. Freeman and Company, San Francisco.

SOKAL, R. R. and D. E. WARTENBERG, 1981 Space and population structure. pp. 186–213. In: *Dynamic Spatial Models,* Edited by D. GRIFFITH and R. MCKINNON. Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands.

WRIGHT, S, 1969 *Evolution and the Genetics of Populations,* Vol. 2: *The Theory of Gene Frequencies.* University of Chicago Press, Chicago.

WRIGHT, S, 1978 *Evolution and the Genetics of Populations,* Vol. 4: *Variability within and among Natural Populations.* University of Chicago Press, Chicago.

Corresponding editor: M. NEI