TEMPORAL AND MICROGEOGRAPHIC VARIATION IN ALLOZYME FREQUENCIES IN A NATURAL POPULATION OF *DROSOPHILA BUZZATII*

J. S. F. BARKER,* P. D. EAST*[,1] AND B. S. WEIR[†]

**Department of Animal Science, University of New England, Armidale, NSW 2351, Australia, and †Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695*

Manuscript received November 4, 1984
Revised copy accepted November 15, 1985

ABSTRACT

Temporal variation in allozyme frequencies at six loci was studied by making monthly collections over 4 yr in one population of the cactophilic species *Drosophila buzzatii*. Ten sites were defined within the study locality, and for all temporal samples, separate collections were made at each of these sites. Population structure over microgeographic space and changes in population structure over time were analyzed using *F*-statistic estimators, and multivariate analyses of allele and genotype frequencies with environmental variables were carried out.—Allele frequencies showed significant variation over time, although there were no clear cyclical or seasonal patterns. A biplot analysis of allele frequencies over seasons within years and over years showed clear discrimination among years by alleles at four loci. During the 4 yr, three alleles showed directional changes which were associated with directional changes in environmental variables. Significant associations with one or more environmental variables were found for allele frequencies at every locus and for both expected and observed heterozygosities (except those for *Est-1* and *Est-2*). Thus, variation in allele frequencies over time cannot be attributed solely to drift. Significant linkage disequilibria were detected among three loci (*Est-2*, *Hex* and *Aldox*), but there was no evidence for spatial or temporal patterns.—The *F*-statistic analyses showed significant differentiation among months within years for all loci, but the statistic used (coancestry) was heterogeneous among loci. Estimates of *F* (inbreeding) for all loci were significantly different from zero, with the loci in four groups, *Adh-1* (negative), *Pgm* (small positive), *Est-2* and *Hex* (intermediate) and *Est-1* and *Aldox* (high positive). The correlation of genes within individuals within populations (f) for each locus in each month by site sample differed among loci, as did the patterns of change in f over time (seasons). Heterogeneity in the *F*-statistic estimates indicates that natural selection is directly or indirectly affecting allele and genotype frequencies at some loci. However, the *F*-statistic analyses showed essentially no microgeographic structure (*i.e.*, among sites), although there was significant heterogeneity in allele frequencies among flies emerging from individual rots.—Thus, microspatial heterogeneity probably is most important at the level of individual rots, and coupled with habitat selection, it could be a major factor promoting diversifying selection and the maintenance of polymorphism. Resolution of the nature of this selection and of the apparent inbreeding detected at all loci except *Adh-1* will require detailed study of the breeding struc-

[1] Present address: Department of Genetics, University of California, Berkeley, California 94720.

ture of the population at the microhabitat level (individual rots) and of gene flow within the population.

ENVIRONMENTAL heterogeneity in space or in time can maintain genetic polymorphisms (reviewed by FELSENSTEIN 1976; HEDRICK, GINEVAN and EWING 1976), and KARLIN (1982) showed that the existence of a protected polymorphism is more likely in more heterogeneous environments and that spatial variation is more effective than temporal in protecting a polymorphism.

For enzyme loci in a variety of species, evidence supporting the hypothesis of environmental heterogeneity as a factor in maintaining polymorphism has been obtained in both laboratory studies and studies of natural populations. In laboratory populations of Drosophila, positive correlations between genetic and environmental variation have been demonstrated (POWELL 1971; MCDONALD and AYALA 1974; POWELL and WISTRAND 1978), although no correlation was detected by YAMAZAKI *et al.* (1983). In natural populations, correlations of allele or genotype frequencies with spatial or temporal variation in environmental factors provide presumptive evidence of selection, although they do not prove direct selection at the enzyme locus. LEWONTIN (1974) suggested that correlations with temporal variation would provide more information than those with spatial variation, in that some estimate of the magnitude of selective differences could be made.

Nevertheless, in natural populations of Drosophila there have not been many studies of temporal variation in allozyme frequencies, and only STEINER (1979) has estimated correlations with environmental variables. Two detailed studies of temporal variation have been carried out. FRANKLIN (1981) studied six loci in six winery populations of *Drosophila melanogaster* over 5 yr, with results suggesting seasonal variation in *Est-6* and *Adh* allele frequencies. For 12 polymorphic loci in one population of *D. melanogaster* studied over 2 yr, CAVENER and CLEGG (1981) found little allele frequency variation either within or between years. Seasonal changes in allele frequency also have been reported by BERGER (1971) for *α-Gpdh* in *D. melanogaster* and by DOBZHANSKY and AYALA (1973) for *Pgm-1* in *Drosophila pseudoobscura*, although PRAKASH (1974) suggested that the latter case may have been due to seasonal changes in frequency of the sex-ratio gene arrangement. Most other reports have included too few allele frequency estimates to give definitive data on temporal variation.

In contrast, there have been a number of reports of significant allele-frequency variation among spatially separate populations and significant correlations with environmental variables [*e.g.*, see NEVO (1983) for a number of species in Israel]. To the extent that spatial environmental parameters relate to long-term averages (climatic variables), whereas temporal ones relate to short-term weather or seasonal differences, it might be expected that significant correlations would be found more readily in the former case.

Significant genotype-environment associations for spatial variation have been reported for *Drosophila buzzatii* in Australia (MULLEY, JAMES and BARKER 1979). In this paper, we present results of a 4-yr study of temporal variation in allozyme frequencies in one natural population of *D. buzzatii*, and of their

correlations with environmental variables. As *D. buzzatii* breeds and feeds in rotting cladodes and fruit of Opuntia species, and is specific to this cactus niche (BARKER and MULLEY 1976), these environmental variables included ones relating to individual rots.

In addition, spatial heterogeneity at the microgeographic level (*i.e.*, within the habitat of a single population) could contribute to the maintenance of polymorphisms, particularly if coupled with habitat selection (SHORROCKS and NIGRO 1981). Thus, ten sites were defined within the study locality, and for all temporal samples, separate collections were made at each of these sites. The effects of microgeographic heterogeneity on population structure and changes in population structure over time were analyzed using the *F*-statistic estimators developed by WEIR and COCKERHAM (1984).

MATERIALS AND METHODS

The study was done at "Yarrawonga" (locality 5 of BARKER and MULLEY 1976), toward the southern end of the main *Opuntia stricta* and *D. buzzatii* distributions in Australia. At this locality, *O. stricta* grows in open sclerophyll woodland on a ridge with northeast to southwest orientation, surrounded on three sides by open grazing land with no *O. stricta*. On the remaining side, the ridge runs into higher wooded hills with some *O. stricta*. Hills on the other side of the valley (minimum of 1.2 km distant) also have *O. stricta* and *D. buzzatii* populations. The ten collection sites were defined within the area infested with *O. stricta*, five in each of two approximately linear transects. The collection area at each site was within a circle of approximately 20 m radius from the designated and marked center of the site. Within transects, sites were approximately 60–70 m apart, and the transects were approximately 100–120 m apart (Figure 1). The transects represent different microhabitats, with sites 1–5 (transect 1) running along the southeastern slope of the ridge, and sites 6–10 (transect 2) running along the crest. Sites 1–5 were more shaded and cooler, with the average daily maximum temperature at site 2 being about 1° less than that at site 8, and the average extreme maximum temperature recorded between collections being about 2° less at site 2, but with no difference between these sites in average daily minimum or in average extreme minima. However, in the winter months (June, July, August), it was often not possible to collect flies at sites 1–4 and, to a lesser extent, at site 5.

Collections were done once each month, starting in February 1974, for 4 yr (except August and November 1977) for a total of 46 samples. The average time between collections was 31 days, with a range from 21 to 46 (excluding the intervals for the two missing collections). Variation among years has been considered in some of the analyses, with the 4 yr defined as February 1974 to January 1975, February 1975 to January 1976, and so on. Seasons were defined as summer (December–February), autumn (March–May), winter (June–August), and spring (September–November).

The six enzyme loci known to be polymorphic in this population (*viz.*, *Esterase-1*, *Esterase-2*, *Hexosaminidase* [called *Pyranosidase* by BARKER and MULLEY (1976); see BARKER 1981], *Phosphoglucomutase*, *Aldehyde oxidase* and *Alcohol dehydrogenase-1*), were routinely assayed, originally using the procedures of BARKER and MULLEY (1976). All loci are autosomal, and *Est-1*, *Est-2* and *Aldox* have been located to chromosome *II*, but the linkage distances are not known. *Pgm* is on chromosome *IV* (A. FONTDEVILA personal communication), *Adh-1* is on chromosome *III* (by homology with *D. mojavensis*, ZOUROS 1976), and *Hex* is on either chromosome *IV* or *V*. For *Adh-1*, gene frequency data were obtained only from assays of larval progeny of wild-caught females for collections from February 1974 to November 1974 and from January 1975 to July 1975. In December 1974, a large collection was made, and many of these flies were used in a first attempt to get *Adh-1* phenotypes in adults using a modified assay. From August 1975, adults

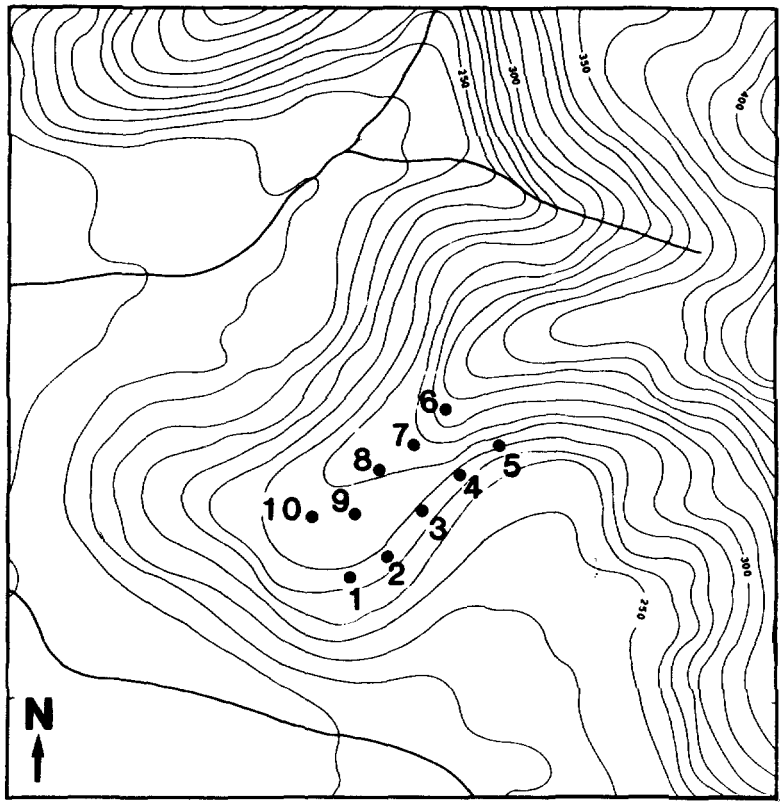

FIGURE 1.—Topographical map of the collection area, showing the location of the ten collection sites.

were scored routinely for all six loci, the assay procedures, where different from BARKER and MULLEY (1976), being as follows: All enzymes were assayed by horizontal starch gel electrophoresis. The esterase-1, esterase-2, phosphoglucomutase and alcohol dehydrogenase enzymes were run on a continuous Tris-borate-EDTA system (Buffer III of SHAW and PRASAD 1970), with the addition of 1 mg/ml $MgCl_2$ to the gel buffer. Hexosaminidase and aldehyde oxidase were assayed on a continuous Tris-citric acid system (Buffer I of SHAW and PRASAD 1970). Staining methods were as described in BARKER and MULLEY (1976), except that ADH was as follows: 4.5 ml 2-propanol, 50 mg β-NAD, 30 mg MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide), 3 mg PMS (phenazine methosulfate) in 100 ml Tris-HCl, pH 8.5, incubate, uncovered, at 37° for 1–2 h.

A total of 15,313 flies were assayed for up to six loci. The actual numbers of adult flies assayed per locus and overall allele frequencies are given in Table 1. For each locus, numbers assayed are less than the maximum possible, primarily because not all flies were assayed for all six loci in some of the larger collections.

Collection methods: When abundant, flies were collected by net from fermenting banana bait. On cooler or windy days, when flies were not so active and would not fly readily when disturbed, they were aspirated directly from the bait container. In the warmer months, *D. buzzatii* are active only in the early morning and late afternoon. As larger numbers usually are active in the afternoon, most collections were done at this time. However, for seven of the samples, collections were made on one afternoon and again the following morning. For two of these, there was significant heterogeneity in

TABLE 1

Numbers of adult flies assayed per locus and overall allele frequencies

Locus	No. of flies	Allele frequencies				
		a	*b*	*c*	*d*	*e*
Est-1	14,363	0.179	0.797	0.024		
Est-2	14,271	0.328	0.317	0.113	0.243	0.0004
Hex	13,701	0.855	0.145			
Pgm	12,175	0.022	0.973	0.005		
Aldox	11,069	0.980	0.020			
Adh-1	8,365		0.492	0.508		

Est-2 allele frequencies between the two times, but with different alleles contributing to the difference in each case. As there were no significant differences between morning and afternoon collections for other loci, we conclude that no bias was introduced when only afternoon collections were done. In winter months, when flies were in low abundance and quite inactive, they were collected by aspiration after disturbance from rots or from the underside of cladodes, and collections were done throughout the day for 2 or 3 days. All Drosophila collected were classified to species, and the numbers of each species recorded.

Collections from individual rots: During the study period, 80 rots collected in the field (55 from the study locality and 25 from another site 8 km distant—locality 60, Martindale) were returned to the laboratory, all emerging flies were collected and the *D. buzzatii* were assayed for the six polymorphic allozyme loci.

Eighteen of these rots were collected at the study location during Collection 45 (October 1977), and at collection, all had Drosophila larvae present. As each rot was found, pH and temperature within the rot and ambient temperature were recorded. In the laboratory on the day after collection, samples from each rot were taken for identification of yeast species present [methods as described by BARKER *et al.* (1983)]. For these 18 rots, no flies emerged from one rot, *D. buzzatii* emerged from 17 (15 to 672 flies) and *D. aldrichi* emerged from four (one to 31 flies). The data for these rots were used to test for genotype-environment associations.

Weather records: Data for the 28 days preceding each collection were obtained from the nearest Bureau of Meteorology station (29 km distant from the study location) for daily maximum and minimum temperature, rainfall (28-day total and number of rain days) and 1500-hr relative humidity. From collection 13 on, extreme maximum and minimum temperatures since the preceding collection were recorded at the study location; before this, these extreme values were taken from the weather station records.

Statistical analyses: As the number of flies collected in each month varied widely (from nine to 1648), standardized gene-frequency deviations (CHRISTIANSEN *et al.* 1976) were used in analysis of variance (ANOVA) of allele frequencies and in genotype-environment association analyses. For each locus in each month, these deviations were estimated as

$$(p_i - p_o)\sqrt{2N_i/p_o(1 - p_o)}$$

where p_i is the frequency of one allele in month i and $(1 - p_i)$ is the frequency of the one or more other alleles, p_o and $(1 - p_o)$ are the frequencies in the total of all the monthly samples and N_i is the number of individuals assayed for the locus in month i.

For genotype-environment association analyses, multivariate associations were tested by canonical correlation and multiple regression (MULLEY, JAMES and BARKER 1979), using the SPSS statistical package (NIE *et al.* 1975). For analyses of complete two-locus disequilibria, multi-allelic loci were treated as having two alleles (the most common

allele and other alleles pooled), in order to minimize problems due to very small numbers in some genotypic classes.

RESULTS

Variation in allele frequencies: For data pooled over the ten sites in each month, the frequencies of each allele for *Est-1* and *Est-2*, and for one allele at each of the other loci are shown in Figure 2. Although there are no obvious seasonal changes in allele frequencies, ANOVAs of standardized gene frequency deviations (Table 2) show significant variation over time for every allele (year × month interaction), as well as differences among years for *Est-1*^{c}, *Est-2*^{a}, *Est-2*^{d}, *Hex*^{a} and *Pgm*^{b}. For the last three, the changes are directional, with *Est-2*^{d} and *Hex*^{a} decreasing and *Pgm*^{b} increasing over time. Preliminary analyses of the gene frequencies (BARKER 1981) showed that *Est-2* did not vary randomly over the 4 yr, but there was no strong evidence of nonrandomness for any of the other loci.

Only two alleles show significant microgeographic variation, with *Est-2*^{b} having a higher average frequency in transect 1 than in transect 2, and *Hex*^{a} having a lower frequency at site 6 than at the other sites in transect 2.

On average, there were no significant differences in gene frequency between the sexes, but for *Aldox*^{a}, there were significant interactions of sex with transect, year and month. In addition, however, the lack of a significant sex effect for *Adh-1*^{b} may be a function of the data set. The activity of ADH is lower in males than in females, and a higher proportion of males showed gel patterns that were classed as unscorable. Thus, in some months where large collections were made, females only were assayed for *Adh-1*. For the 24 collections where both males and females were assayed, the frequency of *Adh-1*^{b} in females was higher than in males in 17 months, was lower in 4 and was equal (difference ≤ 0.01) in 3. However, if the gel scoring failures were primarily *Adh-1*^{b} homozygotes, the sex difference in gene frequency would be reduced. This does not seem likely, as heterozygotes are clearly identified by the heterodimer band, and there is no evidence for absence of the *b* band in heterozygote males. Nevertheless, to provide a conservative estimate of the apparent sex difference in gene frequency, we have used the data for 15 months where either all males were scored or the proportion of unscorable males was <10%, and these unscorable males were assumed to be *Adh-1*^{b} homozygotes. In these 15 months, the frequency of *Adh-1*^{b} was higher in females than in males in 8 months, was lower in 3 and was equal in 4. All but 1 of the 8 months where the frequency was higher in females than in males were at times of low population numbers.

A potentially useful approach to the analysis of allelic variation, which is multivariate and takes account of all loci simultaneously, is the biplot (GABRIEL 1971). The biplot relates populations to variables, defines the relative variances of the variables and correlations among them and shows how various variables discriminate among the populations. As a biplot of all allele frequencies (variables) and months (populations) did not permit simple interpretation, data were pooled to the 17 seasons of the study period, and *Adh* and the low frequency alleles (*Est-1*^{c}, *Est-2*^{e}, *Pgm*^{a} and *Pgm*^{c}) were omitted.

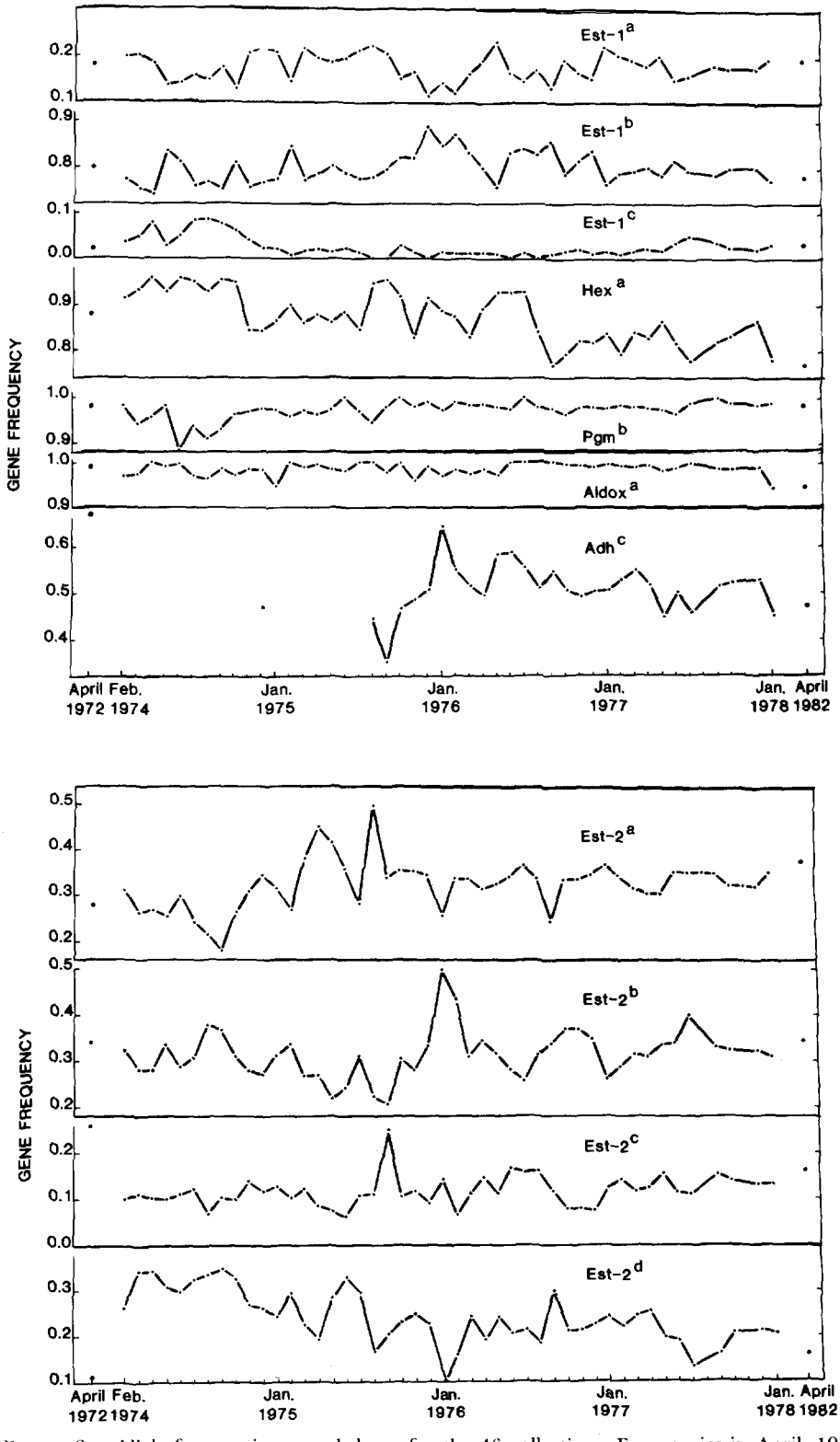

FIGURE 2.—Allele frequencies at each locus for the 46 collections. Frequencies in April, 1972 are from BARKER and MULLEY (1976), and in April, 1982 are from a collection of 1710 wild-caught flies.

TABLE 2

Analyses of variance of standardized gene-frequency deviations

Source of variation	d.f.	Mean squares											
		Est-1[a]	*Est-1*[b]	*Est-1*[c]	*Est-2*[a]	*Est-2*[b]	*Est-2*[c]	*Est-2*[d]	*Hex*[a]	*Pgm*[b]	*Aldox*[a]		*Adh-1*[b]
Sex	1	0.845	1.135	0.247	0.057	0.171	0.466	1.506	1.125	0.170	0.268		0.494
Transect	1	0.874	0.281	0.896	3.804	7.397*	0.544	0.062	0.188	0.542	0.005		1.237
Within transect 1	4	2.105	2.108	0.979	1.445	0.221	1.852	2.699	0.015	1.519	2.344		0.574
Within transect 2	4	0.654	0.602	1.707	0.638	1.196	2.297	1.611	3.080*	1.032	2.222		0.941
Year	3	1.580	1.617	5.596**	4.212*	2.447	0.790	5.917**	5.437**	4.556**	0.454		1.105
Month	11	1.258	1.248	0.789	0.818	0.304	1.206	1.611	0.615	1.313	2.208*		0.536
Sex × transect	1	0.020	0.053	0.061	0.150	1.882	0.689	0.211	0.696	0.595	6.906*		0.806
Sex × within transect 1	4	1.658	2.007	0.714	0.789	0.218	0.993	0.462	0.593	1.249	0.305		0.965
Sex × within transect 2	4	1.828	1.461	0.210	0.339	2.727	2.734	0.780	0.160	3.197*	0.525		0.382
Sex × year	3	0.544	0.375	6.650**	1.547	1.917	2.345	3.436*	0.329	0.049	3.985*		1.157
Sex × month	11	0.618	0.815	2.643*	0.522	1.939	1.387	1.657	1.609	0.857	5.079***		0.834
Year × month	31	2.699***	1.998*	2.369**	3.160***	3.230***	3.920***	3.211***	6.311***	3.015***	(30)† 2.939***	(14)	1.937*
Transect × year	3	1.164	0.851	0.107	0.314	2.383	1.800	0.525	1.264	0.524	2.395		0.981
Within transect 1 × year	12	1.136	1.217	0.846	0.763	1.474	1.628	1.605	0.562	0.979	0.686	(11)	0.526
Within transect 2 × year	12	1.434	1.599	1.455	2.071	1.562	0.911	1.033	1.042	1.105	0.659		0.825
Transect × month	11	1.575	1.896	1.422	1.059	1.768	0.817	0.782	1.246	0.919	1.312		0.603
Within transect 1 × month	38	1.097	0.865	1.107	1.410	1.484	1.212	1.488	0.753	0.999	1.586	(35)	0.920
Within transect 2 × month	44	1.109	1.133	1.540	1.501	1.268	1.344	1.596	0.951	1.044	(43) 0.863		0.941
Error		1.163	1.229	1.255	1.177	1.275	1.229	1.211	1.129	1.166	1.215		0.970
(d.f.)		(532)	(532)	(532)	(533)	(533)	(533)	(533)	(515)	(508)	(458)		(221)

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

† Degrees of freedom, where different from other analyses.

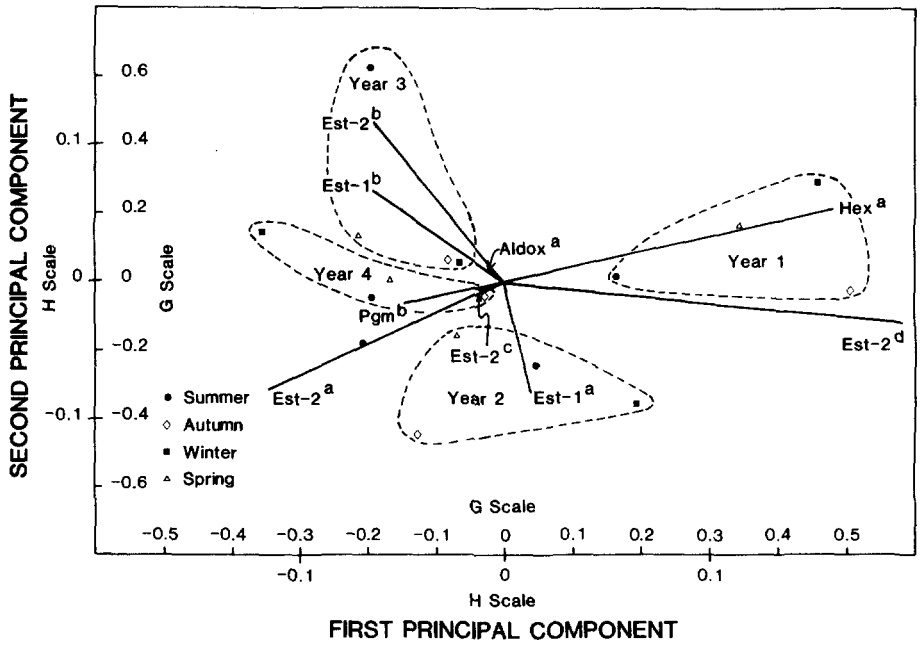

FIGURE 3.—Biplot for allele frequencies and seasons. The summer point on the *Est-2*^{a} line not defined by year is for summer, 1974, *i.e.*, the February 1974 collection.

The methodology rests on the singular value decomposition of the data matrix **Y**. In the present case, **Y** consists of the deviations of each allele frequency from the mean frequency over all seasons and, thus, has 17 rows for the 17 seasons and nine columns for the nine alleles. Column totals are zero.

The singular value decomposition of **Y** is

$$\mathbf{Y} = \sum_{i=1}^{9} \lambda_i \mathbf{p}_i \mathbf{q}_i,$$

where the vectors **p** and **q** are such that $\mathbf{p}_i'\mathbf{p}_j = \mathbf{q}_i'\mathbf{q}_j = \delta_{ij}$, with δ_{ij} being Kronecker's delta. For a two-dimensional representation of the data, we use only the first two terms of this sum.

In particular, the variance-covariance matrix of allele frequencies is proportional to **Y'Y**, and

$$\mathbf{Y'Y} = \lambda_1^2 \mathbf{q}_1 \mathbf{q}_1' + \lambda_2^2 \mathbf{q}_2 \mathbf{q}_2'$$

with

$$(\mathbf{Y'Y})\mathbf{q}_i = \lambda_i^2 \mathbf{q}_i.$$

We see then that λ_i^2 is the ith eigenvalue of **Y'Y**, and $\mathbf{q}_i$ is the associated eigenvector. In the present case, the first two eigenvalues account for 76% of the variation in allele frequencies, so that the biplot in Figure 3 gives a good description of the data. If q_{1j} and q_{2j} are the jth elements of the vectors $\mathbf{q}_1$ and $\mathbf{q}_2$, then the lines in the biplot, for the jth allele, join the origins to the

point ($\lambda_1 q_{1j}$, $\lambda_2 q_{2j}$). Thus, the length of the line is proportional to the variance of the frequency for that allele, and the cosine of the angle between two lines represents the correlation between the frequencies of those two alleles (all in the space spanned by the first two principal components $\mathbf{q}_1$ and $\mathbf{q}_2$). Thus, *Aldox*^{a} and *Est-2*^{c} show little variation, while *Hex*^{a} and *Est-2*^{d} show the most, and there are two groups of substantially correlated allelic frequencies, *viz.* *Est-2*^{d}, *Hex*^{a}, *Pgm*^{b} and *Est-2*^{a} and *Est-1*^{a}, *Est-2*^{b} and *Est-1*^{b}, with the two groups essentially uncorrelated. These generally confirm the correlations shown in Table 3, except for *Hex*^{a}-*Est-2*^{a} within the first group and some significant correlations between the two groups (*e.g.*, *Est-2*^{a}-*Est-2*^{b}).

We also see that

$$(\mathbf{YY}')\mathbf{p}_i = \lambda_i^2 \mathbf{p}_i$$

so that the $\mathbf{p}_i$ are principal components for the variance-covariance matrix $\mathbf{YY}'$ for seasons. If p_{1j} and p_{2j} are the jth elements of the vectors $\mathbf{p}_1$ and $\mathbf{p}_2$, then the points in the biplot representing the jth seasons have coordinates (p_{1j}, p_{2j}), and the distance between these points is the (appropriate) Mahalanobis multivariate distance between the seasons based on all nine allele frequencies. The grouping of seasons in the biplot clearly shows that there is no consistency between seasons in different years but that the four seasons of each year are grouped.

The great power of the biplot is that it relates populations and variables. If we set $g_i = \mathbf{p}_i$ and $h_i = \lambda_i \mathbf{q}_i$, then we see that the data matrix is being approximated by

$$\mathbf{Y} = GH' = [g_1 g_2] \begin{bmatrix} h_1' \\ h_2' \end{bmatrix}.$$

Every element in the data matrix is therefore the inner product of the allele vector and the projection of the population point onto that vector. This enables us to say exactly how the various alleles are discriminating between the seasons. Thus, we see that there is clear discrimination among years by some alleles, *viz.*, *Est-1*^{a}, *Est-1*^{b} and *Est-2*^{b} distinguish year 2 from the other years, whereas *Est-2*^{d}, *Hex*^{a}, *Pgm*^{b} and *Est-2*^{a} distinguish between year 1 and years 3 and 4.

Correlations among allele frequencies: For the data pooled over sites within each month (*i.e.*, 46 observations, except for *Ahd-1* with 29), correlations of allele frequencies were calculated (Table 3), excluding the low-frequency alleles *Est-1*^{c} and *Est-2*^{e} and including only the common allele for *Hex*, *Pgm*, *Aldox* and *Adh-1*. A number of correlation coefficients between alleles at different loci are significant, but none is very high. *Aldox*^{a} is exceptional, however, in showing no significant correlations with alleles at other loci. For *Est-2* alleles, *Est-2*^{a} shows strong negative correlation with *Est-2*^{b} and *Est-2*^{d}, but *Est-2*^{c} appears exceptional in showing little correlation with the other alleles at this locus.

Genotype-environment associations: In order to determine if there were any significant associations between sets of genetic variables and sets of environmental variables (including both abiotic and biotic variables), canonical cor-

TABLE 3

Correlation coefficients over months among allele frequencies, for data pooled over sites within each month

	Est-1^b	*Est-2*^{a}	*Est-2*^{b}	*Est-2*^{c}	*Est-2*^{d}	*Hex*^{a}	*Pgm*^{b}	*Aldox*^{a}	*Adh-1*^{c}
Est-1^a	−0.77***	0.34*	−0.48***	0.26	−0.03	−0.12	0.06	−0.14	−0.31
Est-1^b		0.07	0.32*	−0.07	−0.32*	−0.11	0.30*	0.27	0.33
Est-2^a			−0.47***	0.02	−0.54***	−0.23	0.36*	0.21	−0.42*
Est-2^b				−0.25	−0.31*	−0.21	0.01	−0.08	0.50**
Est-2^c					−0.35*	0.04	0.18	−0.02	−0.26
Est-2^d						0.39**	−0.46**	−0.11	−0.07
Hex^a							−0.45**	−0.04	−0.02
Pgm^b								0.05	0.08
Aldox^a									0.10

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

TABLE 4

Environmental variables defined for multivariate analyses

COLLUNIT—Number of collecting units, higher numbers indicating more time and effort expended and fewer flies in the population
NOBUZZ—Number of *D. buzzatti* collected
NOALD—Number of *D. aldrichi* collected
POPSIZE—NOBUZZ/COLLUNIT
RATIO—NOBUZZ/NOALD
H—Shannon-Weaver species diversity index
VARH—Variance of H
MAXTEMP[a]—Mean daily maximum temperature
MINTEMP—Mean daily minimum temperature
RANGETEMP—Mean daily temperature range
EXTMAX—Highest maximum temperature
EXTMIN—Lowest minimum temperature
EXTRANGE—(EXTMAX-EXTMIN)
RAIN—Total rainfall
RAINDAY—Number of days on which rain fell
RELHUM—Mean daily 1500 h relative humidity
ROTINC—Subjective scoring of incidence of rots in the study location (coded 1, 2, 3)
MAXSD, MINSD, RANGESD, RELHUMSD—Standard deviations of daily MAXTEMP, MINTEMP, RANGETEMP and RELHUM
MAXCV, MINCV, RANGECV, RELHUMCV—Coefficients of variation of same four variables

[a] All weather variables are for the 28-day period before each collection.

relation analyses were done first. These analyses were for each of standardized gene-frequency deviations, expected ($1 - \sum p_i^2$) and observed heterozygosities with each of the sets of environmental variables COLLUNIT to VARH, MAXTEMP to ROTINC, and standard deviations (SDs) and coefficients of variation (CVs) of MAXTEMP, MINTEMP, RANGETEMP and RELHUM (Table 4).

For standardized gene-frequency deviations, the only significant correlation was that with COLLUNIT to VARH ($R = 0.757$, $P = 0.013$). However, the loadings for RATIO, H and VARH were very small, and although that for POPSIZE was similar in magnitude to those of the other three variables, POPSIZE is a ratio of two of them and therefore may not have added significantly to the description of the association. Testing the significance of omitting variables (MARRIOTT 1974) showed that none of the variables POPSIZE to VARH contributed significantly to the correlation, so that COLLUNIT, NOBUZZ and NOALD adequately describe the relationship (Table 5).

Canonical correlation analyses of expected heterozygosities with the same sets of environmental variables gave consistent results. The correlation coefficient with COLLUNIT to VARH was close to being significantly different from zero ($R = 0.709$, $P = 0.082$), but again the relationship was adequately described by COLLUNIT, NOBUZZ and NOALD (Table 5). In addition, the correlation coefficient with the SDs was close to being significantly different from zero (Table 5), with high loadings for *Aldox* expected heterozygosity and for MINSD and RELHUMSD. Although the canonical correlation for standardized gene-frequency deviations with SDs was not significant, the loadings were high on these

TABLE 5

Canonical correlations (R) between sets of genetic and environmental variables, and coefficients for each variable

Gene frequencies	C.C.	Expected heterozygosities	C.C.		Observed heterozygosities	C.C.	
Est-1[a]	0.57	*Est-1*	0.34	−0.25	*Est-1*	0.13	−0.01
Est-1[b]	0.04						
Est-2[a]	−0.42	*Est-2*	−0.24	−0.05	*Est-2*	−0.17	−0.18
Est-2[b]	−0.35						
Est-2[c]	−0.49						
Hex[a]	0.81	*Hex*	−0.61	0.06	*Hex*	0.12	0.16
Pgm[a]	0.37	*Pgm*	−0.95	−0.42	*Pgm*	−0.85	−0.71
Pgm[b]	1.07						
Aldox[a]	−0.32	*Aldox*	0.58	0.90	*Aldox*	0.84	0.81
Adh-1[b]	0.09	*Adh-1*	0.10	−0.03	*Adh-1*	−0.81	
COLLUNIT	−0.76		−1.10				
NOBUZZ	−0.47		−0.52				
NOALD	0.66		0.09				
MAXSD				0.23		0.23	0.64
MINSD				−0.81		−0.21	−0.50
RANGESD				0.29		0.87	0.22
RELHUMSD				−0.60		−0.91	−0.51
R	0.695		0.671	0.660		0.720	0.645
P	0.016		0.007	0.076		0.80 to <0.001	0.013

C.C. = Canonical coefficients, canonical correlation (R) and its probability level (P). Coefficients ≥ 0.5 are underlined to indicate those variables making major contributions to the correlation.

same variables, and multiple regression analyses (see below) showed the same relationship.

Analyses of observed heterozygosities were complicated by the absence of data for *Adh-1* genotypes in adults for collections 1–10 and 12–18. Preliminary analyses (BARKER 1982) showed the correlation coefficient between observed heterozygosities and SDs to be near to significantly different from zero (R = 0.720, P = 0.080), but this test was conservative as the analysis used degrees of freedom dependent on the minimum number of cases, *i.e.*, the number with complete data for all variables. Assuming the full number of cases gave $P <$ 0.001. The true probability cannot be specified, but will be between these two values, and it seems reasonable to assume this correlation coefficient was significantly different from zero. To test this association further, the analysis was repeated omitting *Adh-1* observed heterozygosity, and the correlation coefficient then was 0.645 (P = 0.013, Table 5). When *Adh-1* observed heterozygosity was included, canonical coefficients were highest for observed heterozy-

TABLE 6

Multiple regression analyses of standardized gene-frequency deviations, expected and observed and SDs

Genetic variable	R^2	Regression coefficients			R^2
		COLLUNIT	NOBUZZ	NOALD	
Est-1[a]	0.10				0.14
Est-1[b]	0.17*			−0.0022**	0.11
Est-1[c]	0.09				0.34
Est-2[a]	0.02				0.15
Est-2[b]	0.12			−0.0019*	0.17
Est-2[c]	0.04				0.12
Est-2[d]	0.17*			0.0033**	0.30
Hex[a]	0.25**		−0.0034**	0.0041**	0.41*
Pgm[a]	0.27**	0.0066**		0.0018**	0.28
Pgm[b]	0.31**	−0.0116**		−0.0029**	0.37
Pgm[c]	0.23*	0.0129*			0.35
Aldox[a]	0.13	0.0084*			0.20
Adh-1[b]	0.04				0.29
Expected heterozygosity					
Est-1	0.09				0.13
Est-2	0.03				0.10
Hex	0.18*		0.000082**		0.37
Pgm	0.20*	0.000179**			0.31
Aldox	0.17*	−0.000111*			0.25
Adh-1	0.08				0.21
Observed heterozygosity					
Est-1	0.10				0.16
Est-2	0.02				0.17
Hex	0.16		0.000073*		0.43*
Pgm	0.11				0.37
Aldox	0.10	−0.000083*			0.28
Adh-1	0.04				0.42

Environmental variables with nonsignificant regression coefficients for all genetic variables are omitted from the table.

* $P < 0.05$; ** $P < 0.01$.

gosities of *Pgm*, *Aldox* and *Adh-1* and RANGESD and RELHUMSD. The important environmental variables with *Adh-1* observed heterozygosity excluded still included RELHUMSD (although the magnitude of the canonical correlation was reduced), but also included MAXSD and MINSD, suggesting that RANGESD and RELHUMSD are particularly important for *Adh-1* observed heterozygosity.

Multiple regression analyses were done for standardized gene frequency deviations and for observed and expected heterozygosities with each of the three sets of environmental variables (Table 6). Significant associations with one or more environmental variables were found for standardized gene frequency deviations at every locus and for expected and observed heterozygosities (ex-

heterozygosities on the environmental variable sets COLLUNIT to NOALD, MAXTEMP to ROTINC

Regression coefficients						Regression coefficients		
MAXTEMP	MINTEMP	RANGETEMP	EXTMIN	RELHUM	R^2	MINSD	RANGESD	RELHUMSD
					0.09			
					0.04			
			−0.610*		0.07			
					0.09			
					0.14		−1.361*	
					0.17	−1.304*	1.147*	
				−0.204*	0.06			
	4.57*			−0.228*	0.09			
				−0.108*	0.20			
				0.183*	0.25*			
			−0.806*		0.18			
					0.23*	1.751**		0.298*
				−0.099*	0.05			
					0.06			
					0.01			
					0.11			
					0.22*			
					0.31**	−0.020**		−0.004**
					0.03			
					0.02			
					0.11			
	−0.093*		0.017*	0.0054*	0.10			
					0.23*			
					0.26*			−0.0028*
0.153*	−0.148*	−0.127*			0.32*		−0.061*	

cept for *Est-1* and *Est-2*). These significant associations in general are consistent with the results of the canonical correlation analyses. Further, although the two analyses did not always both detect a particular association as significant, when an association was significant in one analysis, it was generally near to being so and in the same direction in the other analysis.

Thus, gene frequencies and expected heterozygosities for *Hex, Pgm* and *Aldox* showed associations with the population number variables (COLLUNIT, NOBUZZ, and NOALD). Gene frequencies did not show strong associations with the weather variables, although there were significant regression coefficients for gene frequencies at all loci except *Aldox*, and RELHUM was the most important weather variable. *Est-2* and *Aldox* gene frequencies were associated with variability in temperature and relative humidity. *Adh-1* gene frequency

TABLE 7

Environmental variables defined for each rot for multivariate analyses of data from collection 45

NOBUZZ—Number of *D. buzzatti* emerging
ALD—Presence (1) or absence (0) of *D. aldrichi* among emergences
ROTTEMP—Temperature in the rot
TEMPDEV—(ROTTEMP—Ambient temperature)/Ambient temperature
pH—pH of the rot
ROTDUR—Subjective scale (0, 1) estimate of the length of time the particular cladode had been rotting (rot duration)
ROTAGE—Subjective scale (0, 1) estimate of the age of that part of the cladode that was still rotting
NO YSP—Number of yeast species isolated from the rot
C.s.—Presence (1) or absence (0) of *C. sonorensis*
P.c.—Presence (1) or absence (0) of *P. cactophila*
Cl.O—Presence (1) or absence (0) of *Clavispora* sp. O

had very small coefficients in all canonical correlation analyses and had only one significant regression coefficient (with RELHUM). However, there were strong associations for *Adh-1* observed heterozygosity with temperature variables and their variability, indicating for this locus apparent effects on heterozygosity *per se*.

Genotype-environment associations: individual rots (collection 45): Multivariate analyses for genotype-environment associations in these data have been presented briefly by BARKER (1982), using the environmental variables defined in Table 7. Seven yeast species were isolated from the 17 rots from which *D. buzzatii* emerged, but these analyses used only the presence or absence of the three most common species (*viz.*, *Candida sonorensis*, *Pichia cactophila* and *Clavispora* species O).

With only 17 cases, canonical correlation analysis could not be done for standardized gene-frequency deviations with all environmental variables. These analyses were done separately with NOBUZZ to ROTAGE (excluding ROTTEMP, which was highly correlated with TEMPDEV, $r = 0.83$, $P < 0.001$) and with NO YSP to Cl.O, but only the former was significant (Table 8). Canonical correlation analyses of expected and observed heterozygosities with each of the above two sets of environmental variables and with all variables included gave a significant correlation only for expected heterozygosities with all environmental variables (Table 8). Multiple regression analyses of each genetic variable on the full set of environmental variables gave only six significant regression coefficients for individual environmental variables, but five of these associations also had high canonical coefficients. Other associations with high canonical coefficients all had regression coefficients that were consistent in sign.

Thus, for these rots, allele frequencies at *Est-1*, *Est-2* and *Adh-1* were associated with age and pH of the rot and the number of *D. buzzatii* emerging from it. The expected heterozygosity for *Hex* increased with shorter rot duration and when *Clavispora* sp. O was present in the rot.

Complete two-locus disequilibria: For each site in each month, genotypic

TABLE 8

Canonical correlation (R) between sets of genetic and environmental variables, and coefficients for each variable, for the collection 45 rots

Gene frequencies	C.C.	Expected heterozygosities	C.C.
Est-1[a]	−0.14	*Est-1*	−0.08
Est-1[c]	0.67		
Est-2[a]	−0.43	*Est-2*	0.36
Est-2[b]	−1.06		
Est-2[c]	0.16		
Hex[a]	−0.01	*Hex*	0.80
Pgm[b]	0.11	*Pgm*	−0.46
Aldox[a]	0.32	*Aldox*	0.04
Adh-1[b]	0.62	*Adh-1*	0.24
NOBUZZ	0.60		−0.24
ALD	0.35		0.14
TEMPDEV	−0.37		−0.06
pH	0.80		0.12
ROTDUR	−0.04		−0.63
ROTAGE	0.53		−0.01
NO YSP			−0.49
C.s.			0.11
P.c.			−0.41
Cl.O			0.66
R	0.99		1.00
P	0.005		0.000

C.C. = Canonical coefficients, canonical correlation (R) and its probability level (P). Coefficient ≥ 0.5 are underlined to indicate those variables making major contributions to the correlation.

data were analyzed for evidence of associations between genes at every pair of loci, using the methods of B. S. WEIR and C. C. COCKERHAM (unpublished results). The collapsed two-allele genotypic data were used, giving nine genotypic classes for each pair of loci, and the hypotheses tested were two for Hardy-Weinberg disequilibrium (one at each locus), one for linkage disequilibrium, two for three-gene disequilibria and one for four-gene disequilibrium (as defined in WEIR 1979).

In many samples, tests of the two-locus disequilibria were not possible, as either one or both loci were homozygous, or the sample size was too small. The results are summarized in Table 9, which gives, for each locus pair, the number of disequilibria significantly different from zero ($P < 0.05$), the number of possible tests and the proportion of tests that were significant. For linkage disequilibria, using the test on page 252 of WEIR (1979), only two pairs of loci (*Est-2*, *Aldox* and *Hex*, *Aldox*) show substantial increases in the proportion of tests that were significant over the expected 5%. The higher

TABLE 9

Two-locus disequilibria showing, for each pair of loci, the number of significant disequilibria as a fraction of the number of tests made and, in parentheses, the proportion of significant tests (linkage disequilibrium above the matrix diagonal, higher order disequilibria below)

	Est-1	*Est-2*	*Hex*	*Pgm*	*Aldox*	*Adh-1*
Est-1		17/287 (0.059)	12/196 (0.061)	1/37 (0.027)	1/32 (0.031)	9/160 (0.056)
Est-2	4/112 (0.036)		21/278 (0.076)	8/128 (0.063)	14/84 (0.167)	17/197 (0.086)
Hex	3/36 (0.083)	24/232 (0.103)		3/46 (0.065)	5/33 (0.152)	11/162 (0.068)
Pgm	0/0	1/18 (0.056)	0/0		0/2	2/60 (0.033)
Aldox	0/0	4/17 (0.235)	0/1	0/0		2/43 (0.047)
Adh-1	2/78 (0.026)	20/294 (0.068)	13/161 (0.081)	0/5	0/2	

TABLE 10

Numbers of significant positive and negative cases for each locus pair showing substantial disequilibria (linkage disequilibrium above the diagonal, higher order disequilibria below)

	Est-2	*Hex*	*Aldox*
Est-2		0	3+ 11−
Hex	6+ 18−		5+ 0−
Aldox	2+ 2−	0	

order disequilibria also show two pairs of loci with substantial disequilibria, again involving the same three loci but in a different combination (*Est-2*, *Hex* and *Est-2*, *Aldox*). *Esterase-2* and *Aldox* have both been located to chromosome *II*. Both positive and negative significant disequilibria were found for each of these pairs of loci (Table 10), but with an excess of one type for all except *Est-2*, *Aldox* higher order disequilibria.

G-statistic tests (SOKAL and ROHLF 1981) were done to test for heterogeneity over sites, years, months or seasons for the *Est-2*, *Aldox* (all disequilibria and negative ones only) and *Hex*, *Aldox* linkage disequilibria and for the *Est-2*, *Hex* higher order disequilibria (all and negative only). None was significant, so that there is no evidence for spatial or temporal patterns in the occurrence of significant disequilibria.

Analyses of *F*-statistics: WEIR and COCKERHAM (1984) present methods for the estimation of the parameters F, θ and f of population structure, as defined

TABLE 11

Correlations estimated in the *F*-statistics analyses

Symbol	Interpretation
F	Correlation between genes within individuals ("inbreeding") $\equiv F_{IT}$ (WRIGHT 1951).
θ_1	Correlation of genes between individuals within the the same isolate ("coancestry") $\equiv F_{ST}$ (WRIGHT 1951). Here, isolate = month or site.
θ_2	Correlation of genes between isolates in the same subpopulation. Here, subpopulation = year or transect.
$f_1 = \frac{(F-\theta_1)}{(1-\theta_1)}$	Correlation between genes within individuals within isolates $\equiv F_{IS}$ (WRIGHT 1951).
$f_2 = \frac{(\theta_1-\theta_2)}{(1-\theta_2)}$	Correlation between genes of different individuals within isolates within subpopulations.
$f_3 = \frac{(F-\theta_2)}{(1-\theta_2)}$	Correlation between genes within individuals relative to genes between isolates within subpopulations.

previously by COCKERHAM (1969, 1973), but extend the estimation to multiple alleles and loci. Two aspects of population subdivision were considered, *viz.*, (1) over time: years and months within years and (2) over space: transects and sites within transects. The different types of correlations estimated and their interpretation are given in Table 11, and results for the analyses over time and space are given in Tables 12 and 13, respectively.

For the analysis over time, θ_1 is significantly different from zero for all loci, *i.e.*, significant differentiation between months within years, but the magnitude of θ_1 differs among loci. The differentiation between months is greatest for *Hex* and least for *Pgm*. There appear to be no significant differences among loci for θ_2, but for *Est-2* and *Adh-1*, differentiation among years is significant. On the other hand, there is essentially no indication of microgeographic structure in the population, with only a suggestion of differentiation among sites within transects for *Est-2* and *Hex*. For these two loci, *G*-statistic tests of variation in allele numbers among sites within months were done; ten of the 46 tests were significant for *Est-2* and five of 46 for *Hex*. *F*-statistic analyses done over sites, but ignoring transects, showed significant differentiation among sites for *Aldox*, *Pgm* and *Hex*, with site 10 different for the first two loci and sites 8 and 9 different for *Hex*.

Although there is no strong evidence for microgeographic structure at the level of sites and transects, there could be microspatial heterogeneity among individual rots, and assays of flies emerging from individual rots allowed this possibility to be tested. For those collections where two or more rots were available, *G*-statistic tests of heterogeneity among rots were done for each locus (Table 14). Significant heterogeneity was detected for *Est-2* in most collections, for *Est-1* in about one-half and for a number of the loci in some collections.

The estimates of F in the analyses over space and over time (Tables 12 and 13) are very similar, as expected, and that for each locus is significantly different from zero. There are, however, clear differences among the estimates

TABLE 12

F-statistic analysis of population structure over time

Locus	F	θ_1	θ_2	f_1	f_2	f_3
Est-1	0.1606 ± 0.0106[a]	0.0038 ± 0.0010	0.0012 ± 0.0009	0.1574	0.0024	0.1594
Est-2	0.1251 ± 0.0101	0.0073 ± 0.0017	0.0045 ± 0.0016	0.1188	0.0029	0.1214
Hex	0.1288 ± 0.0171	0.0209 ± 0.0067	0.0111 ± 0.0073	0.1108	0.0102	0.1198
Pgm	0.0700 ± 0.0253	0.0028 ± 0.0013	0.0009 ± 0.0008	0.0677	0.0020	0.0696
Aldox	0.1731 ± 0.0589	0.0094 ± 0.0028	0.0046 ± 0.0026	0.1618	0.0047	0.1658
Adh-1	−0.0813 ± 0.0204	0.0029 ± 0.0012	0.0023 ± 0.0011	−0.0835	0.0006	−0.0829
All loci	0.0772	0.0074	0.0042	0.0703	0.0032	0.0733

[a] Standard deviation—jackknife estimate.

TABLE 13

***F*-statistic analysis of population structure over space**

Locus	F	θ_1	θ_2	f_1	f_2	f_3
Est-1	0.1604 ± 0.0073	0.0002 ± 0.0003	0.0000 ± 0.0003	0.1601	0.0002	0.1603
Est-2	0.1244 ± 0.0058	0.0003 ± 0.0001	0.0001 ± 0.0002	0.1241	0.0002	0.1242
Hex	0.1272 ± 0.0127	0.0010 ± 0.0006	0.0008 ± 0.0007	0.1265	0.0002	0.1267
Pgm	0.0707 ± 0.0241	0.0001 ± 0.0002	−0.0001 ± 0.0002	0.0701	0.0002	0.0702
Aldox	0.1689 ± 0.0273	0.0008 ± 0.0008	0.0002 ± 0.0006	0.1681	0.0005	0.1685
Adh-1	−0.0815 ± 0.0175	−0.0003 ± 0.0002	0.0000 ± 0.0001	−0.0811	−0.0003	−0.0814
All loci	0.0761	0.0002	0.0001	0.0759	0.0001	0.0759

TABLE 14

Tests of heterogeneity in allele frequencies for each locus among flies emerging from rots collected at the same time

Month of collection	No. of rots	Locus					
		Est-1	*Est-2*	*Hex*	*Pgm*	*Aldox*	*Adh-1*
Locality 5							
March, 1974	4		**				0
July, 1974	4						0
Feb., 1975	2		***			HOMO	0
March, 1975	3		*			HOMO	0
Sept., 1975	4	***	***		NO TEST	HOMO	***
Aug., 1976	2	**	*	*		HOMO	*
Sept., 1976	4	***	***				**
Oct., 1976	3		***				
May, 1977	3	0				HOMO	
Sept., 1977	5		***				*
Oct., 1977	17	***	***	***	*		**
Locality 60							
Nov., 1976	2						
Feb., 1977	3	**	**	*			
Mar., 1977	5	***	***	***	***		
June, 1977	5	***	***	***	NO TEST	***	***
July, 1977	3						
Aug., 1977	2	*	***	**	HOMO	***	***
Sept., 1977	5	***	***	**	*	**	***

0 = no assay; HOMO = all flies homozygous for common allele; NO TEST = test not possible because of expected numbers < 1.0.

* Significant heterogeneity, $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

for each locus, with the loci in four groups in terms of inbreeding, *viz.*, *Adh-1* (negative), *Pgm* (small positive), *Est-2* and *Hex* (intermediate) and *Est-1* and *Aldox* (highest inbreeding).

One of the *F*-statistics, f, the correlation of genes within individuals within populations, can be estimated from allele frequency data for a single population:

$$\hat{f} = \frac{\frac{2n}{2n-1} H - h}{\frac{2n}{2n-1} H - \frac{h}{2n-1}}$$

where $H = 1 - \sum_u p_u^2$; p_u = frequency of allele u; and h = observed frequency of heterozygotes.

Under the null hypothesis of interest, that f is zero, the variance of f, ignoring terms in $1/n^2$, is

TABLE 15

Numbers of tests for each locus of $f = 0$, and the numbers that were significant, with $f > 0$ or $f < 0$

Locus	No. of tests	No. significant	
		$f > 0$	$f < 0$
Est-1	351	57	0
Est-2	366	76	3
Hex	332	48	3
Pgm	214	19	0
Aldox	151	23	0
Adh-1	219	3	19

$$\text{var}\,(\hat{f}) = \frac{1}{n}\left[\sum_u p_u^2 + \left(\sum_u p_u^2\right)^2 - 2\sum_u p_u^3\right] \Big/ \left(1 - \sum_u p_u^2\right)^2.$$

We now invoke the large sample (maximum likelihood) result that the square of an estimate divided by its (estimated) variance is distributed as chi-square to give the multiple-allele single-degree-of-freedom chi-square of $f = 0$:

$$\chi^2 = n(H - h)^2 / \left(\sum_u p_u^2 + \left(\sum_u p_u^2\right)^2 - 2\sum_u p_u^3\right).$$

For each locus, f, var(f) and χ^2 were estimated from the data for each month by site sample, and the results are summarized in Table 15. *Adh-1* clearly differs from the other five loci, with 19 of 22 significant tests having a negative f (observed excess of heterozygotes). For *Est-1* and *Est-2*, all or most significant tests were for samples with positive f (overall observed deficiency of heterozygotes), but with multiple alleles, some heterozygote genotypes could be in excess and contributing to the significant departure from $f = 0$.

Although not presented here, each locus was tested for deviation from Hardy-Weinberg equilibrium for the sample collected at each site in each month, using the standard goodness-of-fit χ^2 statistic on genotypic frequencies. In calculating the χ^2, if any genotype had an expected number <1, it was pooled with the genotype of next smallest expected number, and where necessary, this pooling was continued so that the test statistic was calculated only for genotypes (or pooled genotypes) all with expected numbers >1. Thus, for small samples, valid tests often were not possible, but the number of valid tests for *Est-1* and *Est-2* were 176 and 302, respectively. The significant tests among these for *Est-1* and *Est-2* (47 and 72, respectively) were investigated in more detail by comparing the observed and expected numbers, excluding cases where the difference was ≤ 1, and counting the number of cases of observed $>$ expected ($O > E$) and observed $<$ expected ($O < E$). For *Est-1* homozygotes, there were 98 $O > E$ and zero $O < E$. The common heterozygote (*Est-1 a/b*) showed only $O < E$, but for the other two heterozygotes (*Est-1 a/c* and *Est-1 b/c*), $O > E$ was almost the same frequency as $O < E$ (13 *vs.* 20, respectively).

Homozygotes for *Est-2* commonly showed an observed excess (only nine of 199 comparisons were $O < E$). For the six heterozygous genotypes, only 57 of 287 comparisons showed an observed excess, but the frequency differed among genotypes. The three heterozygous genotypes including *Est-2*^{c} had 31 of 113 comparisons with an observed excess (0.274), whereas the other three genotypes had a significantly lower ($G_{(1)} = 6.52$, $P < 0.05$) frequency of 26 of 174 (0.149). The distributions over months and seasons for these cases of observed excess heterozygotes for each genotype showed no discernible pattern for either *Est-1* or *Est-2*.

The f estimates for each locus in each month by site sample were subjected to ANOVA to assess variation in time and space of heterozygosity (Table 16). All loci except *Adh-1* showed a significant effect for months within year 2, and there were significant effects for months within other years for *Hex, Pgm* and *Adh-1*. Weighted average f for *Est-1, Est-2* and *Hex* (the three most polymorphic loci), for all loci except *Adh-1* and for all six loci also were estimated, but ANOVA of these showed significant effects only for months within year 2.

Year 2 had the largest average f (all six loci) of all years, with a least-squares mean of 0.145, as compared with 0.119, 0.066 and 0.075 for years 1, 3 and 4. Average POPSIZE (defined in Table 4) was smallest in year 2, *viz.*, 1.33 ± 0.59, as compared with 3.07 ± 1.07, 4.27 ± 1.64 and 4.73 ± 1.88 for years 1, 3 and 4, respectively, but with the highest coefficient of variation of all years.

Given this negative relationship between average f and average POPSIZE across years and the expectation that the probability of mating between relatives would be higher at small population sizes, correlation coefficients were calculated for POPSIZE with f for each of the 46 months and with Δf (change from one month to the next). These correlation analyses were done for f and Δf for each locus separately (omitting f estimates based on ≤ 20 flies) with POPSIZE for the same month, the previous month, the average of these 2 months and the average of the previous 2 months. None of these correlation coefficients was significantly different from zero.

Values of f for each locus in each month fluctuated quite widely, and, as expected from the ANOVA (Table 16), no clear patterns of a seasonal or cyclical nature were obvious. To test for randomness in the patterns, analyses of numbers of turning points and of phase length distributions were done (KENDALL and STUART 1966). Significant deviations from randomness were found only for *Pgm*.

Although much of the variation from month to month will be due to sampling error, the patterns of change in f over time appeared to differ among loci. Two analyses were done to assess further this apparent variation over time. First, correlation analyses between all pairs of loci were done, using f for each month, but omitting f estimates based on ≤ 20 flies. Only two correlation coefficients were significantly different from zero—between *Est-2* and *Aldox* ($r_{(34)} = 0.358$, $P < 0.05$) and between *Est-2* and *Pgm* ($r_{(39)} = 0.516$, $P < 0.001$). In similar analyses among Δf estimates, two correlation coefficients were significantly different from zero—between *Est-1* and *Pgm* ($r_{(33)} = 0.349$, $P <$

TABLE 16

Analyses of variance of f for each locus

	Est-1		*Est-2*		*Hex*		*Pgm*		*Aldox*		*Adh-1*	
Source of variation	d.f.	MS	d.f.	MS	d.f.	MS	d.f.	MS	d.f.	MS	d.f.	MS
Years (Y)	3	0.077	3	0.018	3	0.074	3	0.092*	3	0.004	2	0.031
Transects (T)	1	0.015	1	0.001	1	0.054	1	0.002	1	0.054	1	0.000
Months/year 1	11	0.054	11	0.018	11	0.026	11	0.101***	10	0.062		
Months/year 2	11	0.184***	11	0.276***	11	0.121*	9	0.187***	7	0.211***	5	0.111
Months/year 3	11	0.010	11	0.031	11	0.083	10	0.019	7	0.064	11	0.145*
Months/year 4	9	0.034	9	0.040	9	0.112*	8	0.006	8	0.026	9	0.107
Sites/transect 1	4	0.116	4	0.021	4	0.090	4	0.052	4	0.033	4	0.172
Sites/transect 2	4	0.006	4	0.032	4	0.074	4	0.053	4	0.046	4	0.025
Y × T	3	0.115	3	0.027	3	0.207*	3	0.012	3	0.086	2	0.004
Error	293	0.059	308	0.043	274	0.057	160	0.027	103	0.050	171	0.073

* $P < 0.05$; *** $P < 0.001$.

TABLE 17

Estimates of f for each locus, with standard deviation in parentheses, for each season pooled over years

Locus	Season			
	Autumn	Winter	Spring	Summer
Adh-1	−0.0257	−0.0478	−0.1088***	−0.0949***
	(0.0324)	(0.0850)	(0.0372)	(0.0303)
Aldox	0.2981***	−0.4624	0.1344*	0.0827
	(0.0605)	(0.3270)	(0.0497)	(0.0538)
Est-1	0.1186**	0.2076**	0.1838***	0.1655***
	(0.0252)	(0.0513)	(0.0258)	(0.0258)
Est-2	0.1405***	0.1045	0.0997**	0.1529***
	(0.0162)	(0.0351)	(0.0183)	(0.0177)
Hex	0.1696***	−0.0079	0.1275*	0.1200**
	(0.0305)	(0.0784)	(0.0314)	(0.0266)
Pgm	0.0137	0.1730	0.0303	0.1104
	(0.0426)	(0.1146)	(0.0786)	(0.0627)

* f significantly different from zero, $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

0.05) and between *Est-2* and *Pgm* ($r_{(33)} = 0.554$, $P < 0.001$), and that between *Est-2* and *Aldox* approached significance ($r_{(27)} = 0.358$, $P = 0.056$). Second, as there were no significant effects of years (except for *Pgm*) in the ANOVAs (Table 16), estimates of f were obtained for each season, pooled over years (Table 17). There are clear differences among loci, with f for *Est-1* significantly different from zero in all seasons and for *Pgm* not different from zero in any season, but with both these loci showing no differences among seasons. For *Aldox*, f is significantly higher in autumn than in other seasons, whereas for *Hex*, it is significantly less in winter than in other seasons. *Est-2* shows highest values in summer and autumn, whereas *Adh-1* shows significant deviations from zero in spring and summer.

DISCUSSION

The main results of this study are (1) that allele frequencies show significant variation within and between the 4 yr, but with no clear cyclical or seasonal pattern; (2) that there is very little microgeographic structure, although there is significant heterogeneity in allele frequencies among flies emerging from individual rots collected in the field; and (3) that there is substantial inbreeding, but that the level of inbreeding varies among loci.

Significant variation in time was found for all alleles (Table 2), but the patterns of change over the 46 months were significantly different from random only for *Est-2*^{a} and *Est-2*^{b} (BARKER 1981). Although the standard errors of the allele frequency estimates in some months were large, so that imprecise estimates of frequency could contribute to the apparent randomness, the variation in allele frequencies at loci other than *Est-2* could be due to drift.

However, the weather variables also were tested for randomness in their patterns over the 46 months, and RANGETEMP, RAIN, RELHUM, MAXCV, MINSD, RANGESD and RELHUMCV all showed no significant deviation from randomness. As all of these variables except for RAIN, MAXCV and RELHUMCV were important in the significant genotype-environment associations, it would seem unlikely that the variation in allele frequencies among months could be attributed solely to drift.

In addition, differences among years in allele frequencies were detected in ANOVA (Table 2) and in the biplot (Figure 3), with *Est-2*^{d} and *Hex*^{a} decreasing over the 4 yr and with *Pgm*^{b} increasing. Such changes could be due to directional changes in the environment or in the ecological parameters of the study population. In the biplot, all three alleles discriminate between year 1 and years 3 and 4, whereas in the multivariate regression analyses (Table 6), the regression coefficients for *Est-2*^{d} and *Hex*^{a} on NOALD were significant and positive and those on RELHUM were significant and negative. *Pgm*^{b} had significant regression coefficients of opposite sign. Although the differences among means for each year for these two environmental variables are not significant, the means for NOALD decreased over the 4 yr (294, 145, 23 and 72), whereas those for RELHUM were lower in years 1 and 2 than in years 3 and 4 (44.9, 43.8, 48.0 and 50.6). The directional changes in these variables, or in other unidentified but correlated factors, could be sufficient to account for the directional changes in allele frequencies.

Even if allele frequencies were closely associated with weather variables, the lack of a seasonal cycle in these frequencies would not be surprising, as the weather patterns were quite different in the 4 yr. Nevertheless, significant associations were detected between genetic and environmental variables in the multivariate analyses, although these do not necessarily indicate causal relationships. Those between allele frequencies and the population number variables may indicate density-dependent effects (COLLUNIT, NOBUZZ) or effects of interspecific interactions (NOALD), but they could be due to other environmental variables that affect allele frequencies and are closely correlated with changes in the population size of *D. buzzatii* or *D. aldrichi*.

There is one further factor that may affect temporal variation in some of the allozyme frequencies, *i.e., D. buzzatii* is polymorphic for inversions on chromosomes *II* and *IV* (FONTDEVILA *et al.* 1981, 1982; BARKER *et al.* 1985). As inversion frequencies may show temporal changes (ARNOLD 1982; INOUE, WATANABE and WATANABE 1984), linkage disequilibrium between allozyme loci and inversions may contribute to temporal changes in the former. *Est-1*, *Est-2* and *Aldox* have been located to chromosome *II* and *Pgm* to chromosome *IV*, but the relationships of the enzyme loci to the inversions are not yet known. Inversion frequencies were not obtained during this study, so that their possible influence cannot be assessed. However, the chromosome *IV* inversion is at very low frequencies in Australian populations of *D. buzzatii*, and a further temporal study is in progress, with analysis of chromosome *II* inversion and *Est-1* and *Est-2* allozyme frequencies (W. R. KNIBB, personal communication).

The majority of the rots utilized by these Drosophila develop from microbial

infections in cladodes attacked by larvae of *Cactoblastis cactorum*. *C. cactorum* is bivoltine, with adult moths emerging in September–November and January–March and their larval progeny starting to hatch 4–5 wk later (MANN 1970; MURRAY 1982). Depending on seasonal conditions (temperature and rainfall), rot development follows rapidly on the larval attack. Thus, *D. buzzatii* population numbers peak twice each year, once in late spring-early summer (October–November) and once in early to mid autumn (March–April), in each case within a month of the time of most active rot development. Population numbers are low during the winter and into early spring (June–September), and also crash dramatically but briefly in summer to a minimum in either December, January or February, depending on seasonal conditions (an apparent combination of maximum temperatures and rainfall). On the other hand, *D. aldrichi* shows only one peak in population numbers, in March–April, which coincides with that for *D. buzzatii* at this time.

Most of the rots developing in late spring and late summer are in younger, aerial cladodes and are quite ephemeral, drying out in 6–8 wk and 11 wk on average, respectively (BARKER *et al.* 1983). At other times of the year, the flies are primarily dependent on rots in basal cladodes (lying on the ground, generally rooted and old and often partly protected by leaf litter and other cladodes) and basal stems (in the lower part of the main stem of the plant, where the rot may extend from above ground to 10 cm or more below ground), which may remain as suitable feeding and breeding sites for some months. Thus, there are clear differences in the types of rot utilized at the times of peak *D. buzzatii* numbers and at other times of the year. Further, there are significant differences among rot types and seasons in the relative frequencies of the ten most common yeast species (BARKER *et al.* 1983), and these yeasts are differentially attractive to *D. buzzatii* and differ in their nutritional sufficiency for larval development (BARKER *et al.* 1981a,b; VACEK 1982; VACEK *et al.* 1985). Thus, the observed associations of gene frequencies with the population number variables could be mediated through effects of the microbial flora in the rots.

Nevertheless, direct density-dependent selective effects cannot be excluded. Although the peaks in *D. buzzatii* population size follow the times of major *C. cactorum* larval activity and, therefore, of absolute rot abundance, higher larval density per rot could be expected at these times. This should be particularly so for the autumn peak, when rots last longer and may support two or three generations of *D. buzzatii*. As this is also the time of the normal single *D. aldrichi* peak, associations with NOALD could be indicators of density effects. Associations with the number of *D. buzzatii* emerging per rot thus would be predicted, and were found for *Est-1*, *Est-2* and *Adh-1* in the analyses of collections from individual rots.

In similar multivariate analyses of *D. buzzatii*, using genetic data from a number of natural populations (spatial or geographic variation) and climatic environmental variables, MULLEY, JAMES and BARKER (1979) found significant associations for *Est-2* gene frequencies with means and variabilities of the climatic factors and for *Pgm* and *Adh-1* gene frequencies with their variabilities.

Such associations could represent long-term adaptations to the different climates to which the different populations have been exposed, and need not be reflected in short-term temporal tracking of seasonal weather variation. However, *Est-2* gene frequencies do show associations with measures of temperature variability in both the geographic and temporal analyses. In addition, $Adh\text{-}1^b$ shows a significant negative regression coefficient with relative humidity in both analyses. This latter association is intriguing, although no causal explanation is obvious, as OAKESHOTT *et al.* (1982) found significant associations between *Adh* gene frequencies for *D. melanogaster* and rainfall that accounted for a large proportion of the clinal variation in *Adh* gene frequencies on each of three continents.

Observed heterozygosity for *Adh-1* also showed significant associations with environmental variables, particularly variation in relative humidity and temperature variables and their variability, which could indicate selective effects on heterozygosity *per se*. However, *Adh-1* is different from the other loci in that it shows a generally consistent observed excess of heterozygotes. Such an excess may indicate selection favoring the heterozygotes, but not necessarily (WALLACE 1958; LEWONTIN and COCKERHAM 1959). Other factors which could contribute to an observed heterozygote excess include allele frequency differences between male and female parents (ROBERTSON 1965), selection acting before the allele frequencies are determined in the offspring (LEWONTIN and COCKERHAM 1959), certain types of differential fertility among the parents (PURSER 1966), a sample comprising groups of siblings (RASMUSSEN 1979), sex-dependent migration between populations (PROUT 1981) and assortative mating. Of these, we have evidence indicating allele frequency differences between the sexes. First, in the adults emerging from 56 rots for which *Adh-1* allele frequencies were estimated for both sexes, and defining a sex difference in $Adh\text{-}1^b$ frequency ≤ 0.01 as equal, the frequency of $Adh\text{-}1^b$ was higher in females than in males for 36 (significantly so for six), equal for three, and less for 17 (sign test significant, $P < 0.01$). However, there was no evidence for a nonrandom seasonal distribution of the sex difference. Second, our most conservative estimate of 8 of 15 months where tests were possible showed a higher frequency of $Adh\text{-}1^b$ in females. All but one of these 8 months were at times of low population numbers, either during winter or at the time of the summer crash in population size, when selective differences could be magnified. Thus, for flies homozygous for $Adh\text{-}1^b$, females may have lower mortality rates than males during winter, and also be less susceptible to heat and dessication during the summer crash.

Evidence for a possible interaction between *Aldox, Adh-1* and sex was noted by BARKER (1981) in the data for flies emerging from the rots, in that flies heterozygous for *Aldox* had an excess of $Adh\text{-}1^b$, and this excess was higher in females than in males. A similar tendency is apparent in the data for all monthly collections, although the differences were not significant.

In contrast to the significant temporal variation in the genetic structure of the population, evidence for microspatial variation was limited to $Est\text{-}2^b$ and Hex^a frequencies (Table 2) and to possible differentiation among sites within

transects for these same two loci in the *F*-statistic analysis (Table 13). RICHMOND (1978) reviewed evidence for microspatial genetic heterogeneity in Drosophila species, concluding that it may be the rule rather than the exception and suggesting that behavioral habitat selection may be important in maintaining the heterogeneity. However, for most of the species considered by RICHMOND, the breeding sites and, hence, the basis for any habitat selection are unknown. Thus, environmental heterogeneity to which the populations may be exposed is described in broad ecological terms, such as moist woods, open woods or meadows. At our study site, the habitat would appear to be relatively uniform, apart from the slightly lower average temperature and greater shading at the sites in transect 1 than in transect 2. However, one other difference could account for the higher average frequency of *Est-2*^{b} in transect 1. In this transect (particularly in sites 1–4), there were very few large, old *O. stricta* plants throughout the study period, although such plants were common in sites 6–10. As noted previously, rots in younger cladodes dry out more quickly so that, on average, the rot process proceeds more rapidly in transect 1. Thus, the strong negative association between *Est-2*^{b} frequency and ROTAGE found for the collections from individual rots (Table 8), if generally true throughout the study period, would contribute to the higher frequency of *Est-2*^{b} in transect 1.

BARKER (1982) has summarized our more extensive data for collections from individual rots and has emphasized the potential importance of microenvironmental variation (*i.e.*, between rots) affecting allele and genotype frequencies in *D. buzzatii*. Microspatial heterogeneity at the level of individual rots may be a major factor promoting diversifying selection and the maintenance of polymorphism, particularly if coupled with differential habitat selection by different genotypes. For one microenvironmental factor (*viz.*, yeast species) there is significant heterogeneity among rots (BARKER *et al.* 1983), and the potential for habitat selection. Laboratory and field experiments have shown that larval and adult *D. buzzatii* can discriminate among yeast species (VACEK 1982; VACEK *et al.* 1985), and the field experiments indicated that *Est-2* genotypes are differentially attracted to yeast species (BARKER *et al.* 1981b). The significant heterogeneity among rots in allele frequencies (Table 14) is inadequate to implicate habitat selection, as such heterogeneity also could be due to a founder effect (if the flies collected from a rot were the progeny of one or a few females) or to a nonrandom sample of the flies developing in a rot being collected (*e.g.*, if flies from eggs first laid in the rot had already emerged in the field). Unless there is significant spatial heterogeneity in types of rot (as perceived by different genotypes), dispersal of flies eclosing in a rot, or dispersal as rots dry out and become unsuitable as breeding sites, will prevent spatial population structuring at the microgeographic level. Nevertheless, the heterogeneity among rots emphasizes the possible importance of effects of microhabitat variation on population structure and maintenance of polymorphism and the need for more detailed field studies at this level.

Estimates of *F* for all loci except *Adh-1* were positive and significantly different from zero. Although there was little evidence for spatial genetic struc-

ture (among sites), the heterogeneity among rots in allele frequencies could contribute to a Wahlund effect causing observed deficiencies of heterozygotes. However, BARKER (1981) compared the number of cases of observed heterozygote deficiency and observed heterozygote excess for the 80 rots collected during the study period and showed that there is an overall excess of observed heterozygote deficiency even in emergences from the individual rots. Sign tests were significant for *Est-2* ($P < 0.01$) and *Hex* ($P < 0.05$) and were near to significance for *Est-1* ($P < 0.10$). As the emergences from a single rot will comprise groups of siblings, which would lead to an observed excess of heterozygotes (RASMUSSEN 1979), the observed heterozygote deficiencies for individual rots argue strongly against attributing the significant positive estimates of F for all loci (except *Adh-1*) to a Wahlund effect. It should be noted that, in these rot emergences, *Adh-1* again contrasted strongly with the other loci, with an excess of cases ($P < 0.10$) where observed numbers of heterozygotes were greater than expected.

Other possible factors that might account for the positive estimates of F include selection against heterozygotes, positive assortative mating (both unlikely for five loci), the existence of null alleles or inbreeding. BARKER and MULLEY (1976) first drew attention to the general observed hetrozygote deficiency at these five loci. At that time, there was no information on null allele frequencies, and they argued in favor of inbreeding, suggesting that flies eclosing in a particular rot may tend to remain there as long as it remains a suitable breeding site.

Null alleles appear to be rare in natural populations, *e.g.*, for 20 autosomal loci in *D. melanogaster*, a mean frequency of 0.0024 in two populations (LANGLEY *et al.* 1981). Definitive estimates are not available for *D. buzzatii*, but preliminary data suggest higher null frequencies for *Est-1* and *Est-2* than the average for *D. melanogaster*. Of 40 lines made isogenic for *Est-2*, but with no consideration of alleles at *Est-1* which also is on chromosome *II*, two were homozygous for an *Est-1* null allele, *i.e.*, a frequency of 0.05. Maximum estimates of null frequencies from a study of linkage disequilibrium between *Est-1*, *Est-2* and chromosome *II* inversions (W. R. KNIBB, unpublished results) were 0.034 for *Est-1* and 0.017 for *Est-2*. Assuming Hardy-Weinberg equilibrium, and if null heterozygotes are scored as homozygotes and null homozygotes are inviable or classed as unscorable and excluded from the data, then $F = 2p/(1 + p)$, where p is the null allele frequency. Predicted estimates of p, using the F values in Table 12, are 0.095 for *Aldox*, 0.087 for *Est-1*, 0.069 for *Hex*, 0.067 for *Est-2* and 0.036 for *Pgm*. Clearly, the estimated null frequencies for *Est-1* and *Est-2* cannot account for all the apparent inbreeding. However, null alleles at *Est-1* and *Est-2* also could account for the observations on deviations from Hardy-Weinberg equilibrium frquencies for each genotype at these loci. As would be expected with null alleles segregating, we found a greater proportion of observed heterozygote deficiencies for the common heterozygotes (*Est-1 a/b* and *Est-2 a/b*, *a/d* and *b/d*) than for the other heterozygous genotypes. Nevertheless, at least until more precise estimates of null allele frequencies are obtained, we must conclude that there is inbreeding in this population

additional to that which would accumulate in a random mating population of finite size.

However, the average level of inbreeding and the changes in inbreeding over time vary among loci. Further, in the F-statistic analysis over time, θ_I ($\equiv F_{ST}$) is heterogeneous among loci. Since all loci have been subjected to the same breeding structure, this heterogeneity in inbreeding indicates some form of selection affecting gene and genotype frequencies at some of the loci, although no inference as to the type of selection can be made (NICHOLAS and ROBERTSON 1976).

Assuming selection is affecting some of the loci, the changes in inbreeding over time (summarized for seasonal effects on f in Table 17) should be related to changes in the effective size of the population, and differences among loci in these inbreeding changes then would be indicative of the time and possible nature of the selection. However, the variable POPSIZE showed no relationship with f for any locus, although it is only a crude measure of population size and possibly not really relevant in attempting to monitor changes in f. For example, consider a locus where heterozygotes survive better during the winter. At this time, POPSIZE is small, but the flies are either not breeding or are only breeding to a limited extent. Thus, during the winter months, f would decrease, even though POPSIZE is small. When breeding commences after winter, f might increase in the first one or two generations, but would then decrease as flies disperse to newly formed rots, even though POPSIZE was still quite small.

The results of this large study suggest that selection is affecting gene and genotype frequencies at some enzyme coding loci, but the lack of any evidence for spatial structure in this population implies that the selection is probably based at the micro-environmental level.

Resolution of the nature of the selection affecting these loci, and of the implications of the selection-inbreeding interaction, will depend on more precise knowledge of the breeding structure of the population at the microhabitat level (*i.e.*, individual rots) and of gene flow within the population.

This work was supported by a grant from the Australian Research Grants Scheme to J.S.F.B. and, in part, by grant GM 11546 from the National Institutes of Health to C. C. COCKERHAM. Our appreciation goes to JOHN HACKMAN and JOYCE POOLE (North Carolina State University) and SUSAN BROTHERSTONE (University of Edinburgh) for computing assistance, to JOHN MULLEY and GREG TOLL for field and laboratory assistance, to DON FREDLINE, JILL PARKER and FRANCES MCDONALD for assistance with data preparation, to H. J. PHAFF and MARY MIRANDA for yeast species identification and to K. GABRIEL for the biplot program. B.S.W. is grateful to the Genetics Department, University of Edinburgh, for hospitality and computing facilities during preparation of the paper while he held a Guggenheim Fellowship in that Department.

LITERATURE CITED

ARNOLD, J., 1982 *Statistics of Natural Populations: Seasonal Variation in Inversion Frequencies of Mexican Drosophila pseudoobscura.* Ph.D. Thesis, Yale University, New Haven, Connecticut.

BARKER, J. S. F., 1981 Selection at allozyme loci in cactophilic *Drosophila.* pp. 161–184. In: *Genetic Studies of Drosophila Populations. Proceedings of the 1979 Kioloa Conference,* Edited by J. B. GIBSON and J. G. OAKESHOTT. Australian National University, Canberra.

BARKER, J. S. F., 1982 Population genetics of *Opuntia* breeding *Drosophila* in Australia. pp. 209–224. In: *Ecological Genetics and Evolution. The Cactus-Yeast-Drosophila Model System*, Edited by J. S. F. BARKER and W. T. STARMER. Academic Press Australia, Sydney.

BARKER, J. S. F. and J. C. MULLEY, 1976 Isozyme variation in natural populations of *Drosophila buzzatii*. Evolution **30:** 213–233.

BARKER, J. S. F., G. J. PARKER, G. L. TOLL and P. R. WIDDERS, 1981a Attraction of *Drosophila buzzatii* and *D. aldrichi* to species of yeasts isolated from their natural environment. I. Laboratory experiments. Aust. J. Biol. Sci. **34:** 593–612.

BARKER, J. S. F., F. DE M. SENE, P. D. EAST and M. A. Q. R. PEREIRA, 1985 Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. Genetica. **67:** 161–170.

BARKER, J. S. F., G. L. TOLL, P. D. EAST M. MIRANDA and H. J. PHAFF, 1983 Heterogeneity of the yeast flora in the breeding sites of cactophilic *Drosophila*. Can. J. Microbiol. **29:** 6–14.

BARKER, J. S. F., G. L. TOLL, P. D. EAST, and P. R. WIDDERS, 1981b Attraction of *Drosophila buzzatii* and *D. aldrichi* to species of yeasts isolated from their natural environment. II. Field experiments. Aust. J. Biol. Sci. **34:** 613–624.

BERGER, E. M., 1971 A temporal survey of allelic variation in natural and laboratory populations of *Drosophila melanogaster*. Genetics **67:** 121–136.

CAVENER, D. R. and M. T. CLEGG, 1981 Temporal stability of allozyme frequencies in a natural population of *Drosophila melanogaster*. Genetics **98:** 613–623.

CHRISTIANSEN, F. B., O. FRYDENBERG, J. P. HJORTH and V. SIMONSEN, 1976 Genetics of *Zoarces* populations. IX. Geographic variation at the three phosphoglucomutase loci. Hereditas **83:** 245–256.

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

COCKERHAM, C. C., 1973 Analyses of gene frequencies. Genetics **74:** 679–700.

DOBZHANSKY, TH. and F. J. AYALA, 1973 Temporal frequency changes of enzyme and chromosomal polymorphisms in natural populations of *Drosophila*. Proc. Natl. Acad. Sci. USA **70:** 680–683.

FELSENSTEIN, J., 1976 The theoretical population genetics of variable selection and migration. Annu. Rev. Genet. **10:** 253–280.

FONTDEVILA, A., A. RUIZ, G. ALONSO and J. OCAÑA, 1981 Evolutionary history of *Drosophila buzzatii*. I. Natural chromosomal polymorphism in colonized populations of the old world. Evolution **35:** 148–157.

FONTDEVILA, A., A. RUIZ, J. OCAÑA and G. ALONSO, 1982 Evolutionary history of *Drosophila buzzatii*. II. How much has chromosomal polymorphism changed in colonization? Evolution **36:** 843–851.

FRANKLIN, I. R., 1981 An analysis of temporal variation at isozyme loci in *Drosophila melanogaster*. pp. 217–236. In: *Genetic Studies of Drosophila Populations. Proceedings of the 1979 Kioloa Conference*, Edited by J. B. GIBSON and J. G. OAKESHOTT. Australian National University, Canberra.

GABRIEL, K., 1971 The biplot graphic display of matrices with application to principal component analysis. Biometrika **58:** 453–467.

HEDRICK, P. W., M. E. GINEVAN and E. P. EWING, 1976 Genetic polymorphism in heterogeneous environments. Annu. Rev. Ecol. Syst. **7:** 1–32.

INOUE, Y., T. WATANABE and T. K. WATANABE, 1984 Evolutionary change of the chromosomal polymorphism in *Drosophila melanogaster* populations. Evolution **38:** 753–765.

KARLIN, S., 1982 Classifications of selection-migration structures and conditions for a protected polymorphism. Evol. Biol. **14:** 61–204.

KENDALL, M. G. and A. STUART, 1966 *The Advanced Theory of Statistics.* Vol. 3. Griffin, London.

LANGLEY, C. H., R. A. VOELKER, A. J. LEIGH BROWN, S. OHNISHI, B. DICKSON and E. MONTGOMERY, 1981 Null allele frequencies at allozyme loci in natural populations of *Drosophila melanogaster.* Genetics **99:** 151–156.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change.* Columbia University Press, New York.

LEWONTIN, R. C. and C. C. COCKERHAM, 1959 The goodness-of-fit test for detecting natural selection in random mating populations. Evolution **13:** 561–564.

MCDONALD, J. F. and F. J. AYALA, 1974 Genetic response to environmental heterogeneity. Nature **250:** 572–574.

MANN, J., 1970 *Cacti Naturalised in Australia and Their Control.* S. G. Reid, Government Printer, Brisbane.

MARRIOTT, F. H. C., 1974 *The Interpretation of Multiple Observations.* Academic Press, London.

MULLEY, J. C., J. W. JAMES and J. S. F. BARKER, 1979 Allozyme genotype-environment relationships in natural populations of *Drosophila buzzatii.* Biochem. Genet. **17:** 105–126.

MURRAY, N. D., 1982 Ecology and evolution of the *Opuntia-Cactoblastis* ecosystem in Australia. pp. 17–30. In: *Ecological Genetics and Evolution. The Cactus-Yeast-Drosophila Model System,* Edited by J. S. F. BARKER and W. T. STARMER. Academic Press Australia, Sydney.

NEVO, E., 1983 Population genetics and ecology: the interface. pp. 287–321. In: *Evolution from Molecules to Men,* Edited by D. S. BENDALL. Cambridge University Press, Cambridge.

NICHOLAS, F. W. and A. ROBERTSON, 1976 The effect of selection on the standardized variance of gene frequency. Theor. Appl. Genet. **48:** 263–268.

NIE, N. H., C. H. HULL, J. G. JENKINS, K. STEINBRENNER and D. H. BENT, 1975 *SPSS Statistical Package for the Social Sciences.* McGraw-Hill, New York.

OAKESHOTT, J. G., J. B. GIBSON, P. R. ANDERSON, W. R. KNIBB, D. G. ANDERSON and G. K. CHAMBERS, 1982 Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. Evolution **36:** 86–96.

POWELL, J. R., 1971 Genetic polymorphism in varied environments. Science **174:** 1035–1036.

POWELL, J. R. and H. WISTRAND, 1978 The effect of heterogeneous environments and a competitor on genetic variation in *Drosophila.* Am. Nat. **112:** 935–947.

PRAKASH, S., 1974 Gene differences between the sex ratio and standard gene arrangements of the *X* chromosome and linkage disequilibrium between loci in the standard gene arrangement of the *X* chromosome in *Drosophila pseudoobscura.* Genetics **77:** 795–804.

PROUT, T., 1981 A note on the island model with sex dependent migration. Theor. Appl. Genet. **59:** 327–332.

PURSER, A. F., 1966 Increase in heterozygote frequency with differential fertility. Heredity **21:** 322–327.

RASMUSSEN, D. I., 1979 Sibling clusters and genotypic frequencies. Am. Nat. **113:** 948–951.

RICHMOND, R. C., 1978 Microspatial genetic differentiation in natural populations of *Drosophila.* pp. 127–142. In: *Ecological Genetics: The Interface,* Edited by P. F. BRUSSARD. Springer-Verlag, New York.

ROBERTSON, A., 1965 The interpretation of genotypic ratios in domestic animal populations. Anim. Prod. **7:** 319–324.

SHAW, C. R. and R. PRASAD, 1970 Starch gel electrophoresis of enzymes—a compilation of recipes. Biochem. Genet. **4:** 297–320.

SHORROCKS, B. and L. NIGRO, 1981 Microdistribution and habitat selection in *Drosophila subobscura* Collin. Biol. J. Linn. Soc. **16:** 293–301.

SOKAL, R. R. and F. J. ROHLF, 1981 *Biometry*. W. H. FREEMAN & Co., San Francisco.

STEINER, W. W. M., 1979 Genetic variation in Hawaiian *Drosophila* VI. Seasonally dependent gene changes in *Drosophila mimica*. Evolution **33:** 543–562.

VACEK, D. C., 1982 Interactions between microorganisms and cactophilic *Drosophila* in Australia. pp. 175–190. In: *Ecological Genetics and Evolution. The Cactus-Yeast-Drosophila Model System,* Edited by J. S. F. BARKER and W. T. STARMER. Academic Press Australia, Sydney.

VACEK, D. C., P. D. EAST, J. S. F. BARKER and M. H. SOLIMAN, 1985 Feeding and oviposition preferences of *Drosophila buzzatii* for microbial species isolated from its natural environment. Biol. J. Linn. Soc. **24:** 175–187.

WALLACE, B., 1958 The comparison of observed and calculated zygotic distributions. Evolution **12:** 113–114.

WEIR, B. S., 1979 Inferences about linkage disequilibrium. Biometrics **35:** 235–254.

WEIR, B. S. and C. C. COCKERHAM, 1984 Estimating F-statistics for the analysis of population structure. Evolution **38:** 1358–1370.

WRIGHT, S., 1951 The genetical structure of populations. Ann. Eugen. (Lond.) **15:** 323–354.

YAMAZAKI, T., S. KUSAKABE, H. TACHIDA, M. ICHINOSE, H. YOSHIMARU, Y. MATSUO and T. MUKAI, 1983 Reexamination of diversifying selection of polymorphic allozyme genes by using population cages in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **80:** 5789–5792.

ZOUROS, E., 1976 The distribution of enzyme and inversion polymorphism over the genome of Drosophila: evidence against balancing selection. Genetics **83:** 169–179.

Communicating editor: W. W. ANDERSON