HOW MANY GENES ARE SELECTED IN POPULATIONS OF *DACUS OLEAE*[1]

S. TSAKAS AND C. B. KRIMBAS

Department of Genetics, Agricultural College of Athens, Greece

Manuscript received January 10, 1974

ABSTRACT

Three natural populations of *Dacus oleae* have been sampled for six electrophoretically detected polymorphic genes. The distributions of the allele frequencies in the different populations were analyzed by the method suggested by LEWONTIN and KRAKAUER (1973) (the variation in space one) in order to test for selective neutrality. This method, however, which in our case showed that one or more genes are or have been subjected to selection, does not permit estimation of the number of selected loci. An improvement of this method is proposed that permits a minimum estimation of the selected genes. About a third of them seem to be selected in our material. In our case however, where a cytological analysis is not possible, genes are not differentiated from possible inversions including these genes.

VARIATIONS in gene frequencies can be used to decide whether selection or drift is operating. This is true for temporal changes within a population (KRIMBAS and TSAKAS 1971) as well as for spatial differentiation between populations of the same species (LEWONTIN and KRAKAUER 1973). While natural selection can operate differently at each locus and for each allele at a locus, the effect of population size or of the population's breeding structure is uniform over all loci and all alleles. The methods devised have a broad applicability since both mutation and migration are taken into consideration. Furthermore, statistical tests are available in order to decide between the alternatives of selection and drift (LEWONTIN and KRAKAUER 1973).

We intend to use the method of spatial differentiation for natural populations of the fruit fly *Dacus oleae* to inquire whether we could explain the gene frequency distribution observed solely by the breeding structure or the size of the population examined. Three Greek populations have been sampled: that of Corfu island, in the northwest, that of Thassos island, in the northeast, and that from Crete, in the south.

Dacus oleae has been found to be polymorphic for several electrophoretically detected enzyme systems, and their genetics has been studied: all are controlled by autosomal genes having several alleles, which mendelize normally. In this study we used the esterase gene *Est-A* and *Est-B* (ZOUROS, TSAKAS and KRIMBAS 1970; KRIMBAS and TSAKAS 1971), *APH* (adult alkaline phosphatase), *ADH* (alcohol dehydrogenase), *TO* (Tetrazolim oxidase), and *ODH* (Octanol dehy-

[1] This is paper No. VI of the series "The Genetics of *Dacus oleae*".

TABLE 1

Allelic frequencies in the three populations studied and $\hat{F}$ *statistics*

The numbers of genes examined in each case are indicated by N.

Gene	Allele	Population: Thassos	Corfu	Crete	$\hat{F}$
Est-A	A_0	.00661	.02273	.01555	.0046
	A_{01}	.00000	.00227	.00000	.0015
	A_1	.18982	.11837	.13077	.0078
	A_{12}	.00000	.00227	.00000	.0015
	A_2	.47738	.46933	.55778	.0064
	A_{23}	.00220	.00000	.00000	.0015
	A_3	.02003	.01136	.02222	.0013
	A_{34}	.00000	.00227	.00000	.0015
	A_4	.02643	.03705	.05249	.0031
	A_5	.03133	.04178	.03391	.0006
	A_6	.04046	.04178	.02889	.0009
	A_7	.09481	.09296	.06429	.0025
	A_8	.05434	.08547	.06905	.0025
	A_9	.00661	.00682	.00444	.0002
	A_s	.04998	.06554	.02061	.0072
	N	454	440	450	
Est-B	B_2	.00870	.01307	.00833	.0005
	B_3	.02410	.01299	.01250	.0018
	B_{34}	.00000	.00216	.00000	.0014
	B_4	.62780	.49892	.69723	.0283
	B_{45}	.02381	.02410	.05351	.0020
	B_5	.00870	.02381	.02960	.0038
	B_6	.04965	.07185	.04257	.0030
	B_7	.11237	.10751	.07805	.0026
	B_8	.00000	.00216	.00417	.0014
	B_9	.00000	.00216	.00000	.0014
	B_s	.14487	.24127	.07404	.0362
	N	462	462	480	
APH	1.00	.56866	.52597	.50208	.0031
	.96	.43134	.47403	.49792	.0031
	N	466	462	480	
ADH	1.00	.59442	.60606	.58958	.0002
	1.30	.40558	.39394	.41042	.0002
	N	466	462	480	
TO	1.00	.98290	.98238	.98333	.0000
	.65	.01710	.01762	.01667	.0000
	N	468	462	480	
ODH	1.12	.00427	.00000	.03497	.0188
	1.05	.23504	.17808	.18881	.0038
	1.00	.76069	.78083	.77622	.0004
	.95	.00000	.04109	.00000	.0278
	N	234	146	286	

drogenase). The techniques used for detection do not differ essentially from the ones usually mentioned, except for the esterases we described earlier (TSAKAS and KRIMBAS 1970.)

Alleles have been characterized in the usual way, except for *Est-A* and *Est-B*, where we followed the same symbols used in previous publications.

In Table 1 we indicate the frequencies of the 36 alleles of these six genes in the three populations sampled, as well as the $\hat{F}$ values for each allele (uncorrected for sample size). These $\hat{F}$'s would correspond to the inbreeding coefficients of a "Wahlundian" population—that is, of an ideal population subdivided into many populations and of equal size each. These $\hat{F}$'s are expected to be equal in the case of selective neutrality. The mean $\hat{F}$ equals .00507. The ratio of the observed variance of the $\hat{F}$ distribution ($S^2_{\hat{F}}$) to the theoretically expected one in the case of neutrality ($\sigma^2_{\hat{F}}$) is used in order to decide whether the results could be explained by selective neutrality or if we need to invoke selection (LEWONTIN and KRAKAUER 1973). The theoretically expected variance in the case of neutrality is calculated according to LEWONTIN and KRAKAUER (1973) as $\sigma^2_{\hat{F}} = \frac{2\bar{F}^2}{n-1}$, where $\bar{F}$ is the mean of the calculated $\hat{F}$'s from our data, and n the number of populations sampled.

This ratio is equal to 2.791 (for 30 d.f.) and is highly significant ($P \ll .001$) indicating that at least one gene, if not more, is or has been selected some time in the history of these populations after their splitting apart.

Although this method does not allow us to answer directly the question of how many genes are selected, we can try an approach to this problem in spite of the danger of overexploiting the raw data. From an inspection of Table 1 it seems that some *Est-B* alleles (B_4 and B_8) provide exceptionally large $\hat{F}$ values. In fact all eleven alleles of *Est-B* give an $\hat{F}$ distribution with a ratio $\frac{S^2\hat{F}}{\sigma^2_{\hat{F}}} = 2.331$ (for 10 d.f.) that is statistically highly significant ($.001<P<.01$). We can also perform the following tests: exclude one gene at a time and calculate the same ratio for the remaining five genes. Six such tests could be made. In all cases (Table 2) the results are statistically significant. This means that at least two genes are selected. We can then proceed to further tests: exclude two genes at a time, and examine the $\hat{F}$ distributions of the alleles of four remaining genes. Fifteen such tests can be made, fourteen of which have significant results. This would indicate that *Est-B* and *ODH* (that is to say at least two out of the six genes examined) are selected ($= .33 \pm .19$). Of course this estimate is a minimum one, since it could be argued that the remaining four genes are selected, the *ODH* and *Est-B* genes being neutral. Furthermore it should be pointed out that when genes are mentioned they could be genes or chromosome segments, e.g. inversions including these genes.

EWENS (1972) proposed another method to approach this problem. He examined the distribution of the allelic frequencies of one gene at a time in one population. Under some strict conditions (stable population size, closed population, possibility of detection of every new allele generated by mutation, etc.) these

TABLE 2

Statistical tests using $\frac{s_F^2}{\sigma_F^2}$ *ratios of the* $\hat{F}$ *distribution of the alleles for different combinations of the genes examined*

	s_F^2/σ_F^2	d.f.	P
A) For alleles of all genes examined	2.791	30	<.001
B) For alleles of five of the examined genes			
a) excluding *Est-A*	2.561	16	<.001
b) excluding *Est-B*	2.386	20	<.001
c) excluding *APH*	2.560	29	<.001
d) excluding *ADH*	2.597	29	<.001
e) excluding *TO*	2.581	29	<.001
f) excluding *ODH*	3.370	27	<.001
C) For alleles of four of the examined genes			
a) excluding *Est-A* + *Est-B*	2.518	6	.01 –.02
b) excluding *Est-A* + *APH*	2.503	15	.001–.01
c) excluding *Est-A* + *ADH*	2.242	15	.001–.01
d) excluding *Est-A* + *TO*	2.223	15	.001–.01
e) excluding *Est-A* + *ODH*	3.607	13	<.001
f) excluding *Est-B* + *APH*	2.815	19	<.001
g) excluding *Est-B* + *ADH*	2.141	19	.001–.01
h) excluding *Est-B* + *TO*	2.115	19	.001–.01
i) excluding *Est-B* + *ODH*	0.943	17	.50 –.70
j) excluding *APH* + *ADH*	2.606	28	<.001
k) excluding *APH* + *TO*	2.588	28	<.001
l) excluding *APH* + *ODH*	3.471	26	<.001
m) excluding *ADH* + *TO*	2.387	28	<.001
n) excluding *ADH* + *ODH*	2.869	26	<.001
o) excluding *TO* + *ODH*	3.097	26	<.001

distributions at stationary state would tend to certain values. EWENS also provides two statistics, L and F, (a variance ratio differing from the coefficient of inbreeding previously characterized by the same symbol) with its non-integer degrees of freedom and a computer program to compute them. A statistically significant result would indicate selection and also provide some indication of the kind of selection (heterozygous superiority, or selection for one allele).

These statistics have been computed and recorder in Table 3 for every gene and population sampled. In two cases only, for the *APH* gene, the F statistics would indicate selection. However it should be remarked that, as EWENS pointed out, this method is not very powerful. Furthermore many of the assumptions of the model are unrealistic, especially the possibility of detecting every allele: two different allozymes can be confused when they have the same charge. From these considerations and the lack of concordance of results between the EWENS method and the LEWONTIN and KRAKAUER one, we conclude that for the time being the

TABLE 3

L and F statistics of the Ewens test for every gene and population examined

Genes		*L*	*F*	df	Comment
Est-A	Thassos	0.45	1.26	20.75 , 12.29	non-significant
Est-A	Corfu	0.35	1.18	26.54 , 14.10	non-significant
Est-A	Crete	0.34	1.20	18.21 , 11.38	non-significant
Est-B	Thassos	0.17	1.11	11.11 , 8.58	non-significant
Est-B	Corfu	—0.08	0.96	18.07 , 11.38	non-significant
Est-B	Crete	—0.39	0.80	13.16 , 9.52	non-significant
APH	Thassos	1.90	133.58	0.73 , 1.35	non-significant
APH	Corfu	1.93	945.52	0.73 , 1.35	*F* significantly large
APH	Crete	1.94	148571.45	0.79 , 1.34	*F* significantly large
ADH	Thassos	1.86	69.58	0.73 , 1.35	non-significant
ADH	Corfu	1.84	54.57	0.73 , 1.35	non-significant
ADH	Crete	1.87	78.06	0.72 , 1.34	non-significant
TO	Thassos	—0.68	0.26	0.73 , 1.34	non-significant
TO	Corfu	—0.66	0.28	0.73 , 1.35	non-significant
TO	Crete	—0.68	0.26	0.72 , 1.34	non-significant
ODH	Thassos	0.27	1.32	2.56 , 3.11	non-significant
ODH	Corfu	0.36	1.43	2.98 , 3.09	non-significant
ODH	Crete	0.52	1.71	2.42 , 3.09	non-significant

most powerful and reliable method actually at our disposal is the first one used in this study.

We are grateful to MR. J. SOURDIS for computing the data of Table 3, to PROF. J. HUBBY, PROF. R. C. LEWONTIN, DR. B. CHARLESWORTH and MISS E. KRITICOS for corrections in the manuscript.

LITERATURE CITED

EWENS, W. J., 1972 The sampling theory of selectively neutral alleles. Theoret. Pop. Biol. **3:** 87–112.

KRIMBAS, C. B. and S. TSAKAS, 1971 The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control—selection or drift? Evolution **25:** 454–460.

LEWONTIN, R. C. and J. KRAKAUER, 1973 Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. Genetics **74:** 175–195.

TSAKAS, S. and C. B. KRIMBAS, 1970 The genetics of *Dacus oleae*. IV. Relation between adult esterase genotypes and survival to organophosphate insecticides. Evolution **24:** 807–815.

ZOUROS, E., S. TSAKAS and C. B. KRIMBAS, 1968 The genetics of *Dacus oleae*. II. The genetics of two adult esterases. Genet. Res. **12:** 1–9.

ZOUROS, E. and C. B. KRIMBAS, 1969 The genetics of *Dacus oleae*. III. Amount of variation at two esterase loci in a Greek population. Genet. Res. **14:** 249–258.

Corresponding editor: R. C. LEWONTIN