Genet. Res., Camb. (1982), **39**, *pp.* 41–61
With 5 text-figures
Printed in Great Britain

Multilocus nonrandom associations in *Drosophila melanogaster*

BY J. M. MALPICA* AND D. A. BRISCOE†

Institute of Animal Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, Great Britain

(*Received 24 November 1980, and in revised form 30 July 1981*)

SUMMARY

Seven third chromosome samples from laboratory populations and one from a wild population were classified by their genotypes at five allozyme loci and by the occurrence of lethals. The data were analysed for independence among classifications by a randomization method, and it was found that when the number of classifications is high the independence hypothesis does not hold. We have split the independence hypothesis among populations and have found that the lack of fit to the hypothesis comes mainly from the wild population. As one of our classifications was lethality, these findings seriously question the interpretation of multilocus fitness estimations. To split the independence hypothesis into interactions, the criterion of Bartlett for non-interaction with exact probability methods was used; interactions at the levels of pairs and triplets were not significant as a whole. It was not possible to carry the interaction analysis any further as the frequencies of chromosomal types were completely determined by their marginal totals at orders of interaction higher than that of triplets. Available parametric estimations of these high order interactions are therefore of dubious meaning. The analyses of chromosome samples and the temporal surveys of population cages described here, as well as data from the literature, suggest that the Est 6-Pgm region is prone to show disequilibrium.

INTRODUCTION

Independence among the genetic units is a common assumption in both population and quantitative genetics. Several models, either of drift or selection, conclude that nonrandom association among loci can be common (e.g. Hill & Robertson, 1968; Franklin & Lewontin, 1970).

Using allozymes as marker loci, numerous estimations of such associations have been published in the past few years (for recent reviews see Langley, 1977, and

* Present address: I.N.I.A. Departmento de Genética, Apartado 8111, Madrid Spain.
† Present address: School of Biological Sciences, Macquarie University, North Ryde, N.S.W. 2113, Australia.

Barker, 1979), with the conclusion that there is little evidence for strong linkage disequilibrium between allozymic loci per se, and that there is strong nonrandom association between such loci and closely linked chromosomal arrangements. Nevertheless, probably due to the low power of tests for linkage disequilibrium, as Barker (1979) says: 'One must conclude that the evidence is inconclusive, and that it is premature to suggest that epistatic interactions and linkage disequilibria among allozyme loci are rare in natural populations.'

We have looked at the association state of sets of third chromosomes extracted from populations of *D. melanogaster*, practically free of inversions for that chromosome, and have found that in a large natural population, when several loci are involved, the independence hypothesis is untenable.

MATERIALS AND METHODS

The populations used in this study, namely Standard Kaduna, Mancha, Riudevella, Stellenbosch and Amherst have their origins in Nigeria, the centre and NE of the Iberian Peninsual, SW of South Africa and NE of the United States, respectively. Foundation stocks were over 500 inseminated females. Amherst was kindly supplied by Dr P. T. Ives and Stellenbosch by Dr J. H. Louw.

In the laboratory, the populations were kept in cages at 25 °C with weekly changes of standard corn-agar-molasses food pots ($\simeq$ 200 c.c) that were kept in the cage for three weeks.

Three months after its arrival in the laboratory, two replicates were extracted from the original Mancha cage. Two extra pots were introduced in it and transferred to the two new cages after a week. The procedure was repeated for three consecutive weeks. Subsequently homogeneity among the three cages was tested by comparing the gene frequencies of four allozymes (Adh, αGphd, Est-6 and Pgm). In these samples, the maximum numbers of variants for each of the four allozymes were 2, 2, 2, and 4 respectively. The average number of genes scored per locus per cage was 182. The samples did not differ significantly in gene frequencies: the sum of the likelihood ratios of the four gene frequencies by cages contingency tables was not significant ($G_{(12)} = 15{\cdot}380$, $0{\cdot}3 > P > 0{\cdot}2$). The conditions of the replicates were then changed by adding ethanol to the melted standard food, to a final concentration of 10%.

Horizontal starch gel electrophoresis was used throughout this study. The buffer system was tris-versene-borate (0·05 M, pH = 8·0 for the gel, 0·5 M in the tanks), except for esterases and octanol dehydrogenase for which the Poulik (Poulik, 1957) system was used. Staining recipes used were those of Shaw & Prasad (1970), with two minor modifications: (i) in the case of phosphoglucomutase, a cutting of Watman 3 filter paper of approximately the same size as the gel was laid on the fresh cut gel surface, and the staining mixture (5 c.c.) was poured on it; and (ii) in the case of aldehyde oxidase, a distillate of commercial acetaldehyde was used as substrate, following the advice of T. Skinner, who also provided us with the product.

Extraction of nearly isogenic lines for the third chromosome was done applying the standard technique, using as balancer stock *TM3 Sb Ser/Pr* (Lindsley & Grell, 1967). In the F2 crosses, only *TM3* males and virgin females that were not *Pr* were used. The criterion for lethality was the absence of wild type individuals in a line when two successive generations were scored (about 50 flies per generation).

The standard Kaduna population is known to be free of inversions. The rest of the populations were searched for non-recombining blocks by recombination analysis against a *rucuca* chromosome (Lindsley & Grell, 1967). Frequencies of non-recombining chromosomes (usually for less than a chromosome arm) were low; Mancha (on the sample described as 1 below) 0·02 (3/125), Stellenbosch 0·16 (3/19), Riudevella 0·05 (3/58), Amherst 0·01 (1/83). A check was also made in one of the alcohol replicates toward the end of the study (sample described as 5 below) giving a frequency of 0·01 (1/102). No association was found between those non-recombining chromosomes and any particular alleles.

RESULTS

(i) *The data*

The times spent by the above described populations in the laboratory, or under alcohol treatment, before the gametic samples were taken, are given in Table 1. Seven enzymatic loci were analysed in most of these samples, isocitrate dehydrogenase (*Idh*), esterase 6 (*Est*-6) phosphoglucomutase (*Pgm*), esterase-C (*Est*-C), octanol dehydrogenase (*Odh*), xanthine dehydrogenase (*Xdh*), and aldehyde oxidase (*Aldox*). Five of these loci (*Est*-6, *Pgm Est*-C, *Odh*, *Aldox*) were found to be polymorphic in at least one of the samples by our electrophoretic methods.

Table 1 shows the numbers which symbolize these classification criteria through this work, their map locations (taken from O'Brien & MacIntyre, 1971, except for *Est*-6 which is taken from Franklin, 1971), when applicable, and the frequencies of the variants found for them in the different samples. A list of chromosomal types and their frequencies in the samples is given in Table 2.

(ii) *Independence analysis*

To answer the question of whether the classifications are independent of each other, we have used a randomization method (Sokal & Rohlf, 1969). This method overcomes the problem of the non-asymptotic behaviour of the contingency statistics, due to the small expectations in some of the cells of the contingency table.

The tests were carried out using a computer program that, assuming independence, produces random sets of chromosomes which have the same marginal totals as the observed contingency table, and calculates the likelihood ratio of each set. From this, it gives the significance of the observed test (proportion of random sets that have a likelihood equal or bigger than the one observed), and the distribution of significances of the random sets. Note that only for a continuously distributed statistic is this distribution uniform.

Table 1. *Classification criteria and frequencies of variants*

		Samples							
Months in treatment* ...	Variants†	1 (MW)‡ 0	2 (MS) 6	3 (MA1) 2	4 (MA2) 2	5 (MA1) 13	6 (St) 36	7 (R) 8	8 (A) 24
Criteria									
1									
Lethality	1	0·576	0·767	0·571	0·577	—	0·760	0·867	0·791
	2	0·424	0·233	0·428	0·423	—	0·240	0·133	0·209
2									
	1	0·024	—	0·024	—	—	—	—	—
Est 6	2	0·224	0·317	0·298	0·183	0·216	0·300	0·590	0·391
Map position 35, 9	3	0·744	0·683	0·678	0·803	0·784	0·700	0·410	0·609
	4	0·008	—	—	0·014	—	—	—	—
3									
	1	—	—	0236	0·228	—	0·100	—	0·082
Pgm	2	0·808	0·917	0·714	0·746	0·696	0·740	0·924	0·827
Map position 43, 4	3	0·184	0·067	0·238	0·225	0·160	0·160	0·076	0·091
	4	0·008	0·017	0·012	—	0·144	—	—	—
4									
Est C	1	0·976	0·983	0·988	1·000	—	1·000	1·000	1·000
Map. position 49, 0	2	0·024	0·017	0·012	—	—	—	—	—
5									
Odh	1	0·008	—	—	—	—	—	—	—
Map. position 49, 2	2	0·976	1·000	0·988	1·000	—	1·000	1·000	1·000
	3	0·016	—	0·012	—	—	—	—	—
6									
Aldox	1	—	—	—	—	—	0·010	0·010	—
Map. position 56, 6	2	0·032	0·017	0·036	—	0·048	0·190	0·010	0·255
	3	0·968	0·983	0·964	1·000	0·952	0·800	0·981	0·754
Number of lines		125	60	84	71	125	100	105	110

* 'Treatment' refers to culture in standard population conditions, except for the Mancha alcohol replicates, for which alcohol was added to the medium (see text).

† The enzymatic variants are numbered in the order of their anodic migration. With respect to lethality symbol 1 corresponds to viable chromosomes.

‡ The letters between brackets stand for the population from which the sample has been extracted MW = Mancha Wild, MS = Mancha Standard food replicate, MA1 (2) = Mancha Alcohol replicate 1 (2), St = Stellenbosch, R = Riudevella, A = Amherst.

Table 2. *Frequencies of chromosomal types*

Chromosomal types	Samples							
	1 (MW)	2 (MS)	3 (MA1)	4 (MA2)	5 (MA1)*	6 (St)	7 (R)	8 (A)
1 1 2 1 2 3	3	—	1	—	—	—	—	—
1 2 1 1 2 3	—	—	1	1	—	1	—	—
1 2 2 1 2 1	—	—	—	—	—	—	1	—
1 2 2 1 2 2	1	—	—	—	1	3	1	6
1 2 2 1 2 3	9	14	9	7	19	14	50	25
1 2 2 2 2 3	1	—	—	—	—	—	—	—
1 2 3 1 2 2	—	—	—	—	—	1	—	2
1 2 3 1 2 3	3	1	3	2	6	3	2	2
1 2 4 1 2 3	—	—	—	—	1	—	—	—
1 3 1 1 2 2	—	—	—	—	—	2	—	—
1 3 1 1 2 3	—	—	—	1	—	4	—	3
1 3 2 1 2 2	—	1	1	—	4	11	—	17
1 3 2 1 2 3	40	27	20	23	63	30	32	30
1 3 2 1 3 2	1	—	—	—	—	—	—	—
1 3 2 1 3 3	1	—	1	—	—	—	—	—
1 3 2 2 2 3	—	1	1	—	—	—	—	—
1 3 3 1 1 3	1	—	—	—	—	—	—	—
1 3 3 1 2 2	—	—	1	—	1	—	—	—
1 3 3 1 2 3	10	2	10	6	13	7	5	2
1 3 4 1 2 3	1	—	—	—	17	—	—	—
1 4 2 1 2 3	—	—	—	1	—	—	—	—
1 4 3 1 2 3	1	—	—	—	—	—	—	—
2 1 2 1 2 3	—	—	1	—	—	—	—	—
2 2 2 1 2 2	—	—	1	—	—	—	—	3
2 2 2 1 2 3	10	4	10	2	—	6	8	4
2 2 2 2 2 2	1	—	—	—	—	—	—	—
2 2 2 2 2 3	1	—	—	—	—	—	—	—
2 2 3 1 2 1	—	—	—	—	—	1	—	—
2 2 3 1 2 3	2	—	1	1	—	1	—	1
2 3 1 1 2 3	—	—	2	—	—	3	—	6
2 3 2 1 2 2	1	—	—	—	—	1	—	—
2 3 2 1 2 3	32	8	15	20	—	9	5	6
2 3 3 1 2 2	—	—	—	—	—	1	—	—
2 3 3 1 2 3	6	1	5	7	—	2	1	3
2 3 4 1 2 3	—	1	1	—	—	—	—	—
Total number	125	60	84	71	125	100	105	110

* Criteria for which this sample has not been analysed (see Table 1) are stated as if found in state 1 for criteria 1 and 4, and in state 2 for criterion 5.

The percent significance obtained by this method for the actual samples, when the different criteria of classification are included, are given in Table 3. On average, they are based on 700 simulated sets of chromosomes with a minimum of 400.

An overall picture of the validity of the independence hypothesis may be obtained by comparing the distribution of the significances of the observed samples with that of the random samples from the simulation. That comparison is represented in Fig. 1, in which both distributions have been grouped in 10%-wide significance classes.

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

Table 3. *Estimates of the significance (per cent) of the independence tests*

Criteria included in the test	Samples							
	1 (MW)	2 (MS)	3 (MA1)	4 (MA2)	5 (MA1)	6 (St)	7 (R)	8 (A)
1 2 3 4 5 6	12	—	93	—	—	—	—	—
1 2 3 4 5	27	—	83	—	—	—	—	—
1 2 3 4 6	22	95	89	—	—	—	—	—
1 2 3 5 6	33	—	90	—	—	—	—	—
1 2 4 5 6	3*	—	92	—	—	—	—	—
1 3 4 5 6	18	—	92	—	—	—	—	—
2 3 4 5 6	3*	—	98	—	—	—	—	—
1 2 3 4	23	90	71	—	—	—	—	—
1 2 3 5	45	—	71	—	—	—	—	—
1 3 4 6	76	88	81	—	—	38	65	0·19**
1 2 4 5	14	—	98	—	—	—	—	—
1 2 4 6	6	93	85	—	—	—	—	—
1 2 5 6	18	—	84	—	—	—	—	—
1 3 4 5	47	—	67	—	—	—	—	—
1 3 4 6	47	69	86	—	—	—	—	—
1 3 5 6	19	—	86	—	—	—	—	—
1 4 5 6	6	—	100	—	—	—	—	—
2 3 4 5	9	—	90	—	—	—	—	—
2 3 4 6	6	100	97	—	—	—	—	—
2 3 5 6	17	—	96	—	—	—	—	—
2 4 5 6	2*	—	100	—	—	—	—	—
3 4 5 6	6	—	100	—	—	—	—	—
1 2 3	47	78	53	72	—	74	32	< 0·06**
1 2 4	4*	87	95	—	—	—	—	—
1 2 5	42	—	91	—	—	—	—	—
1 2 6	56	88	77	—	—	16	100	4*
1 3 4	71	50	55	—	—	—	—	—
1 3 5	25	—	55	—	—	—	—	—
1 3 6	81	51	77	—	—	26	100	< 0·08**
1 4 5	60	—	100	—	—	—	—	—
1 4 6	24	100	100	—	—	—	—	—
1 5 6	12	—	100	—	—	—	—	—
2 3 4	5*	100	83	—	—	—	—	—
2 3 5	34	—	9	—	—	—	—	—
2 3 6	44	100	95	—	39	39	23	0·63**
2 4 5	7	—	100	—	—	—	—	—
2 4 6	2*	100	100	—	—	—	—	—
2 5 6	21	—	100	—	—	—	—	—
3 4 5	27	—	100	—	—	—	—	—
3 4 6	19	100	100	—	—	—	—	—
3 5 6	9	—	100	—	—	—	—	—
4 5 6	2*	—	100	—	—	—	—	—
1 2	25	100	81	17	—	78	100	82
1 3	55	25	38	46	—	67	100	< 0·06**
1 4	58	100	100	—	—	—	—	—
1 5	50	—	100	—	—	—	—	—
1 6	100	100	100	—	—	5*	100	18
2 3	33	100	74	87	12	32	6	2*
2 4	2*	100	100	—	—	—	—	—
2 5	75	—	100	—	—	—	—	—
2 6	60	100	100	—	100	25	100	100
3 4	65	100	100	—	—	—	—	—
3 5	25	—	100	—	—	—	—	—
3 6	64	100	100	—	66	38	100	11
4 5	100	—	100	—	—	—	—	—
4 6	11	100	100	—	—	—	—	—
5 6	7	—	100	—	—	—	—	—

*, ** indicate significance at 5 and 1% level, respectively.

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

There is an excess of observations in both extremes, conservative and significant, the deviation with respect to the expectations being stronger as the dimensionality goes upwards, becoming significant when the three or four classifications are involved ($G_{(9)} = 19{\cdot}397$, $P < 0{\cdot}05$ and $G_{(9)} = 21{\cdot}280$, $P < 0{\cdot}02$, respectively). The excess on the conservative side is difficult to explain. It does not appear to be a

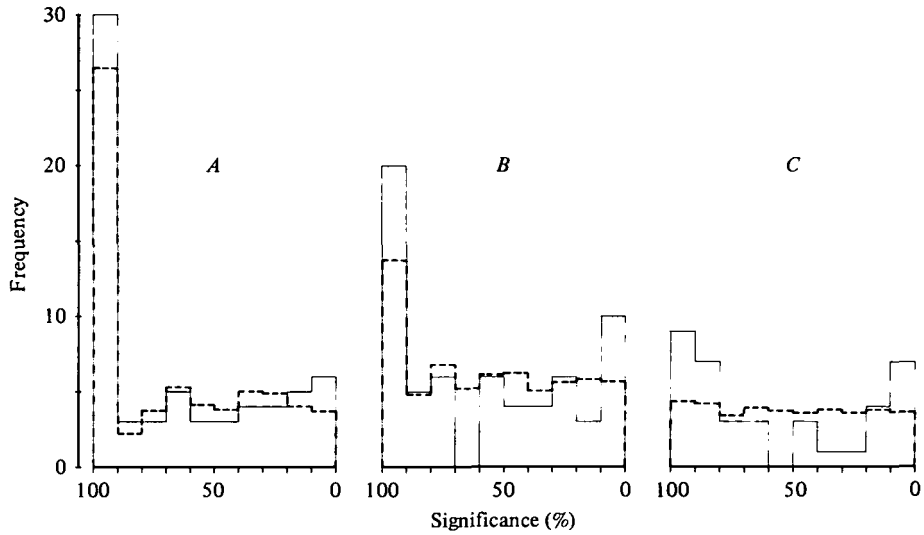

Fig. 1. Expected (dotted line) and observed (solid line) distributions of significance of the independence tests. A, B and C are two, three and four classifications respectively.

by-product of the method of analysis, because in those cases of two classification criteria with only two variants each, in which the same phenomenon shows, exact probability analysis of the same data have been carried out (see below), and the agreement of the results from both methods seems remarkably good (Fig. 2), in both significance of the observed data and expected distribution of those significances. No logical explanation of a populational type can be put forward, as no population force can reduce the sampling dispersion. Therefore, it appears that an oddity in the sample is the only likely explanation; an inspection of Table 3 tells us that the oddity would be confined to samples 2, 3 and 7, mainly to sample 3.

(iii) *Interaction analysis*

Since the independence hypothesis does not appear to fit the data when the number of classifications involved is high, a splitting of its degrees of freedom has been pursued to inquire into the level or levels at which the interaction are caused.

The partition of the independence hypothesis presents two problems, partition among multiple variants per classification within a level of interaction, and partition among levels of interaction, defining 'level of interaction' as the effect

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

due to the presence of a given number of classifications, ideally free of the effect of the necessary presence of the lower number of classification combinations.

The first type of problem is unresolved by the theory (Feldman *et al.* 1975). We have decided to split the data within samples in all the possible sets of pairs of alleles per locus, as has been done by Charlesworth & Charlesworth (1973).

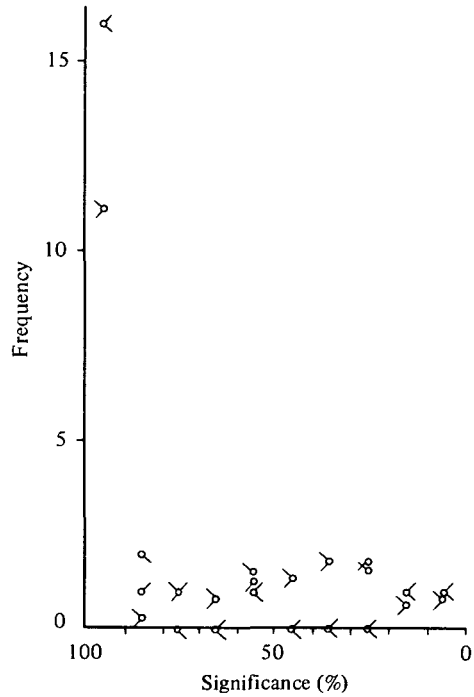

Fig. 2. Distributions of significance for the two classifications with two variants per classification tests obtained by the exact probability method (dash upwards) and by the randomization method (dash downwards). Dash on the left side = expected, dash on the right side = observed.

Two types of methods have been proposed for the splitting between levels of interaction, those based on the partition of Lancaster (Lancaster, 1951) and those that adopt the criterion of Bartlett for non-interaction (Bartlett 1935). Hill (1975*a*) has shown that linkage disequilibrium estimates based on the partitions of Lancaster under neutrality are dependent on population number, and they increase with it when the recombination fraction is small, whereas equilibrium is expected under Bartlett's criterion.

As a consequence we have chosen the criterion of Bartlett, applying the conditional distributions given by Andersen (1974), to obtain by exact probability methods the significances of the interactions in our samples and the expected distribution of these significances.

In the 2×2 interactions, some of the contigency tables may have too few chromosomes in the sample or have frequencies too extreme to have much

meaning; choosing among them the possibly meaningful ones is somewhat arbitrary. We have chosen those in which the marginal frequencies allow at least a configuration of the chromosome frequencies that have a probability equal to or smaller than 0·2. For these, the significances and the strength of the associations, measured as the correlation between allelic states, are given in Table 4. Care must be taken in making statements about the sign and strength of those associations, without taking into account its significance, because of the asymmetry of the data. For example, in the contingency table *Est*-6 (1,2)–*Pgm* (2,3) in sample 1, the state of maximum probability (actually, the observed one) gives a correlation of 0·14. No obvious pattern of sign or strength of association appears to emerge.

The comparison between the observed and expected distributions of all possible 2×2 tests is shown in Fig. 3*a*. There is a suggestion of a surplus of observations in the significant extreme, but this fails to be significant itself ($G_{(9)} = 7{\cdot}782$, $0{\cdot}7 > P > 0{\cdot}5$). There is no sign of excess of observations in the conservative extreme.

Analysis of the higher order interactions gives little insight into the problem. The reason is that the condition imposed upon the data, namely the equality of the marginals of the next inferior order to the one observed, does not allow generally, with our sample size, any possibility of freedom to the system. Out of 206 $2 \times 2 \times 2$ contingency tables in our data only 28 are allowed more than the observed configuration after fixing the marginals for pairs of classifications. Among these 28, the agreement between observed and expected distributions of significance (Fig. 2*b*) is very good ($G_{(6)} = 2{\cdot}756$, $0{\cdot}9 > P > 0{\cdot}8$). Pooling the frequencies of the rare variants, reducing the data to two variants per classification, does not help either, leaving only 19 tests with a possibility of more configurations than the observed one, out of 64 that are possible. When these pooled data are analysed, the above statements on fitting are not modified.

A value can be given for each of these higher order interactions by any of the parametric partition methods, but those values look to us very doubtful. Let us take, as an example, the data of Mukai and co-workers (Mukai, Watanabe & Yamaguchi, 1974) on the third chromosome of *D. melanogaster*, probably the largest sample yet analysed for this chromosome. After pooling, the data (ibid. Table 4) consists of 489 chromosomes classified by two alternatives at each of four classification criteria (*Est*-6, *Est*-C, *Odh* and chromosomal arrangement). These data were analysed by the chi-squared Lancaster's partition method (ibid. Table 5), with the outcome that two of the interactions, one at the level of pairs (*Esc*-C, *Odh*) and one at the triplet level (*Est*-6, *Est*-C, *Odh*), were significant. If the data are analysed using the likelihood ratio as parameters, though still using Lancaster's partition, the significance of this three-locus interaction disappears ($\chi^2_{(1)} = 5{\cdot}2718$, $P < 0{\cdot}05$ v. $G_{(1)} = 1{\cdot}2545$, $0{\cdot}3 > P > 0{\cdot}2$). If the data are analysed by the criterion of Bartlett (exact probability method), then the four-classification interaction is not allowed any possibility of freedom, as in our case, and none of the three-classification interactions pass the 35% significance level.

It could therefore be said that probably no meaningful observation of interactions

Table 4. *Summary of the results of the 2 × 2 tests of association between the variants and criteria indicated.*

(Numerals in the body of the Table are correlations between allelic states ($R \times 10^4$), and superscript letters are significance levels – blank superscript = 100, a = 50-99, b = 6-49, c = 0-5%.)

Criteria		Variants				Samples							
						1 (MW)	2 (MS)	3 (MA1)	4 (MA2)	5 (MA1)	6 (St)	7 (R)	8 (A)
1	2	1	2	1	2	2970^{b}	—	—	—	—	—	—	—
				1	3	1486^{b}	—	-305	—	—	—	—	—
				2	3	-685^{a}	367	-712^{a}	1909^{b}	—	-409^{a}	152	454^{a}
1	3	1	2	1	2	—	—	-926^{a}	1586^{a}	—	-649^{a}	—	-3821^{c}
				1	3	—	—	-2592^{a}	3162^{b}	—	132	—	-2667^{b}
				2	3	-768^{b}	193	-1319^{b}	723^{a}	—	870^{a}	-70	2053^{b}
1	4	1	2	1	2	770^{a}	-718	—	—	—	—	—	—
1	5	1	2	2	3	-1106^{a}	—	—	—	—	—	—	—
1	6	1	2	1	2	—	—	—	—	—	-5416^{b}	—	—
				1	3	—	—	—	—	—	—	389	—
				2	3	-280	718	370	—	—	1466^{b}	389	1465^{b}
2	3	1	2	1	2	—	—	-658	—	—	—	—	—
				2	3	1435	—	1231	—	—	—	—	—
		1	3	1	2	—	—	-500	—	—	—	—	—
				2	3	843^{a}	—	1217^{a}	—	—	—	—	—
				2	4	225	—	358	—	—	—	—	—
		2	3	1	2	—	—	-52	1589^{b}	—	-1511^{b}	—	-2462^{c}
				1	3	—	—	1089	2362^{b}	—	-3016^{b}	—	-5669^{c}
				2	3	68	416	1370^{b}	-160	-637^{a}	-525^{a}	1988^{b}	-498^{a}
				2	4	—	—	—	—	1642^{b}	—	—	—
				3	4	—	—	1085	—	3148^{b}	—	—	—
		2	4	1	2	—	—	—	1000	—	—	—	—
		3	4	1	2	—	—	—	227	—	—	—	—
				2	3	2128^{b}	—	—	—	—	—	—	—
				3	4	-566	—	—	—	—	—	—	—
2	4	1	2	1	2	1071	—	—	—	—	—	—	—
		1	3	1	2	—	—	246	—	—	—	—	—
		2	3	1	2	-2906^{c}	—	—	—	—	—	—	—
		2	4	1	2	-642	—	—	—	—	—	—	—
2	5	1	3	1	2	-188	—	—	—	—	—	—	—
				2	3	265	—	256	—	—	—	—	—
		2	3	2	3	718	—	—	—	—	—	—	—
		3	4	1	2	110	—	—	—	—	—	—	—
				2	3	-155	—	—	—	—	—	—	—
2	5	1	2	2	3	-860	—	-555	—	—	—	—	—
		1	3	2	3	-262	—	-351	—	—	—	—	—
		2	3	2	3	1178^{b}	—	120	—	-269	-882^{b}	—	23
		2	4	2	3	514	—	—	—	—	—	—	—
		3	4	2	3	153	—	—	—	—	—	—	—
3	4	1	2	1	2	—	—	284	—	—	—	—	—
		2	3	1	2	-751^{a}	-354	—	—	—	—	—	—
		2	4	1	2	-173	-182	-167	—	—	—	—	—
3	5	1	2	2	3	—	—	284	—	—	—	—	—
		2	3	1	2	-1886^{b}	—	—	—	—	—	—	—
				2	3	-600	—	—	—	—	—	—	—
		2	4	2	3	-141	—	-167	—	—	—	—	—
		3	4	1	2	435	—	—	—	—	—	—	—
3	6	1	2	2	3	—	—	-405	—	—	-22	—	-1864^{b}
		1	3	1	3	—	—	-826	—	—	891	—	—
		2	3	1	2	—	—	—	—	—	-5423^{b}	—	—
				1	3	—	—	—	—	—	-2419^{b}	284	—
				2	3	871^{a}	354	-380	—	127	661^{a}	284	572^{a}
		2	4	2	3	201	182	238	—	1017^{a}	—	—	—
		3	4	2	3	—	—	500	—	—	—	—	—
4	5	1	2	1	2	143	—	—	—	—	—	—	—
				2	3	-202	—	—	—	—	—	—	—
4	6	1	2	2	3	-2685^{b}	169	211	—	—	—	—	—
5	6	1	2	2	3	-143	—	—	—	—	—	—	—
		2	3	2	3	-3390^{b}	—	211	—	—	—	—	—

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

above the three-locus level has been achieved yet, and that it will be very difficult to get any, because of the labour required to obtain large sample sizes with gametic data, and of the decline in the efficiency of zygotic data as the number of classifications increases, predicted by Hill (1974*a*). Unfortunately, three-locus interactions are not expected in general (Hill, 1974*b*; Franklin & Lewontin, 1970; Feldman, Franklin & Thompson, 1974).

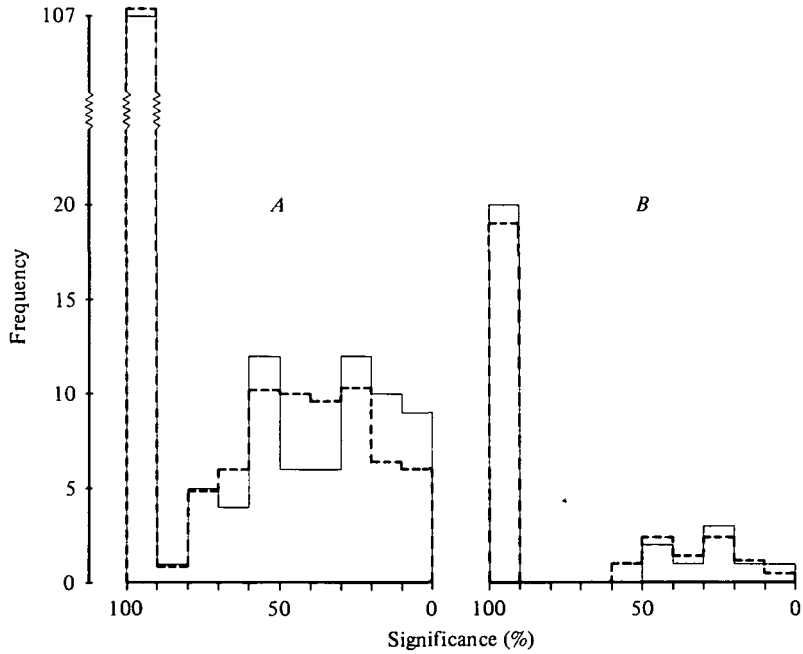

Fig. 3. Expected (dotted line) and observed (solid line) distributions of the significance of the interaction tests. A 2×2 tests, B $2 \times 2 \times 2$ tests.

We are faced now with the problem that, in our data, the independence hypothesis does not appear to hold when more than two classifications are involved. However, the interactions at the 2×2 contingency level, if any, seem to be weak, and the interactions at the $2 \times 2 \times 2$ level fit the null hypothesis very well. The interactions at levels beyond these are unable to modify the fitting of the independence hypothesis. It does appear to us that the more likely explanation is that those usually weak, pairwise interactions are the cause of the lack of fit of the independence hypothesis. It could be argued that this non-independence is a proof of the non-chance origin of the pairwise interactions, as the independence hypothesis ought to allow margin for chance deviations at any level of interaction.

(iv) *Individual associations*

When singling out as significant an individual test, taking into account the number of tests at its dimensional or interaction level, the following expression was used

$$A_i = 1 - \prod_{j=1}^{n} \{1 - pj(\alpha_i)\},$$

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

where A_i stands for the effective significance level applied to the ith test as included in a set of n, and $p_j(\alpha_i)$ for the expected probability of the significance class α_i in the jth test; α_i being the level of significance observed in the ith individual test. This formula is a straightforward generalization for discontinuity of the formula proposed by Neimann-Sørensen & Robertson (1961) for the continuous case.

Applying this formula to the independence test with the minimum $p(\alpha)$ applicable, that is, the inverse of the minimum number of random chromosomes generated per test (400), the three test with lower probability are all in sample 8 (Amherst population): lethality-*Est*-6-*Pgm*, lethality-*Pgm*-*Aldox* (both at 13%)

Table 5. *Combinations showing disequilibrium between Est 6 and Viability in a sample from the Standard Kaduna population*

		Viability		
		Viable	Lethal	Total
Esterase 6	2	48	2	50
	3	105	22	127
Total		153	24	177

and, lethality-*Pgm* (7%). Note that all three criteria combinations share the overall low probability two criteria combination, and that out of the other four possible two criteria combinations, one (*Est*-6-*Pgm*) is significant when considered by itself and two have suspiciously low probabilities.

The application of the same method to single out overall significant interactions is not formally valid, because it implies a conditional probability statement on these tests, and they are not independent within a given combination of classifications. Nevertheless, this is not a problem in practice, because the number of tests within a combination of classifications, able to reach the high level of individual significance $\{p_j(\alpha_i)\}$ required when n is large, is going to be very seldom more than one.

In this way, it is again the combination lethality-*Pgm* (1, 2) in the Amherst population that emerges with an overall 3% significance. In this population, several cases of linkage disequilibrium have been reported by Charlesworth & Charlesworth (1973). In our sample, the most significant one is that of lethality-*Pgm*. It is difficult to think of a physiological basis for such an association, as one of the criteria is not a single locus one. Nevertheless, there is some indication that, at least under the conditions of our population cages, this is not such an uncommon case. A sample from another population, Kaduna, that has been kept for a very long time by the population cage system, shows a strong linkage disequilibrium ($D' = D/Dmax = 0{\cdot}705$, $G_{(1)} = 6{\cdot}618$, $P < 0{\cdot}02$) between *Est*-6 and Lethality (Table 5).

A possible explanation for those linkage disequilibrium observations in sample 8 could be the contamination of the population by one or a few successful chromosomes a few generations back. The results of the complementation tests that

have been performed between the chromosomes carrying lethal (s) from this sample are given in Table 6. It can be seen that the complementation groups defined by the chromosomes, 1, 93, and 98, can explain roughly half of the lethality (13 out of 23 non-viable chromosomes), and that the other crosses allow us to exclude most of the rest of the chromosomes as carriers of any of those lethals or groups of them. Should the contamination have occurred, it would be expected that the locus nearer to the relatively frequent lethals would be the one showing stronger linkage disequilibrium with lethality (in our case *Pgm*) and that the locus/*i* in linkage disequilibrium with that locus (in our case *Est*-6) would show it in apposite directions for the sets of chromosomes defined by frequent and by rare lethals, respectively.

As it can be deduced from Table 7, the association between *Est*-6 (2,3) and *Pgm* (1,2), the only one to which it is possible to attach a sign in both groups of frequent and rare lethals, gives the same direction of linkage disequilibrium in both. It can thus be concluded that contamination is not a likely explanation for these results. Linkage disequilibrium between *Est*-6 and *Pgm* has been repeatedly found by Langley and co-workers (Langley, Tobari & Kojima, 1974; Langley, Ito & Voelker, 1977), and more recently by Cochrane & Richmond (1980).

An additional suggestion that the sporadic associations between *Est*-6 and *Pgm* may have other bases, apart from sampling errors, comes from the study of the temporal variation in linkage disequilibrium in the replicates of the Mancha population, represented for the association *Est*-6 (2, 3)–*Pgm* (2, 3) in Fig. 4.

The data in that figure (4), except for the points MS Dec. 1972, MA1 and MA2 March 1973 and MA1 April, 1974, which belong to the samples 1, 3, 4 and 5, were obtained on zygotes and were analysed following Hill (1974*a*), who also provided us with a computer program for performing the analysis. The arrow in the time axis marks the beginning of the alcohol treatment, so the 3 points above the arrow can be considered repeated observations on Mancha Standard food replicate (MS). As the gene frequencies diverge between treatments at the *Pgm* locus and to a lesser extent at the *Est*-6 locus (Briscoe & Malpica, unpublished results), in contrast with Cavener & Clegg (1978) results, the correlation coefficient and the relative disequilibrium ($D' = D/Dmax$) behave in a very similar way. The pattern suggests a transient linkage disequilibrium, in the Mancha alcohol replicate 1 and disequilibrium, although much less marked, in the same direction, in the replicate 2, as well as a building up of linkage disequilibrium in the opposite direction in the standard food replicate.

DISCUSSION

From these results, several tentative statements can be put forward.

(*a*) The hypothesis of independence among classifications does not hold when the number of classifications involved is large. However, it appears that the interactions at the level of pairs of loci are not strong, and that they are even weaker at the triplet level.

Table 6. *Complementation between lethal bearing lines from the Amherst population*

Lines	2	10	13	16	26	27	29	38	51	57	74	80	83	86	87	93	98	105	108	112	114	116	Chromosomal type		
																							2	3	6
1	+	+	+	0	.	+	0	+	+	+	0	+	.	.	.	.	+	.	+	+	.	0	2	2	2
2	.	+	.	+	.	.	+	.	.	+	.	.	+	.	+	+	.	+	.	.	+	.	2	2	2
10	.	.	.	+	.	.	+	.	+	.	.	.	.	.	.	+	+	.	.	.	.	+	2	3	2
13	.	.	.	.	.	.	+	+	.	.	.	.	.	+	.	+	.	.	+	.	.	.	2	3	2
16	.	.	.	.	.	.	0	+	.	+	.	.	.	.	.	+	.	.	+	.	.	0	2	1	2
26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	+	.	.	.	+	.	1	2	2
27	.	.	.	.	.	.	+	.	.	.	.	.	.	.	.	+	0	+	.	.	+	+	1	2	2
29	.	.	.	.	.	.	.	+	.	.	.	.	.	.	.	0	+	+	.	.	+	.	2	1	2
38	.	.	.	.	.	.	.	.	.	+	.	.	+	.	+	+	.	.	.	.	+	.	1	2	1
51	.	.	.	.	.	.	.	.	.	.	.	+	+	.	.	+	0	+	.	0	+	.	2	2	2
57	.	.	.	.	.	.	.	.	.	.	+	+	+	.	+	+	+	.	+	.	.	.	2	3	2
74	.	.	.	.	.	.	.	.	.	.	.	+	0	.	.	+	.	.	.	.	+	.	2	2	2
80	.	.	.	.	.	.	.	.	.	.	.	.	+	.	+	+	+	.	.	.	.	.	1	2	1
83	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.	0	2	1	2
86	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	+	.	.	.	2	1	2
87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.	0	2	1	2
93	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	+	.	0	1	2	2
98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0	+	.	1	2	1
105	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	1	2	2
108	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	1	3	2
112	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	2	2	2
114	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	2	2
116	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	1	2

Table 7. *Genotypes of frequent and rare lethals carrying chromosomes from the Amherst population*

Genotypes		Number of chromosomes	
Esterase 6	Phosophoglucomutase	Frequent	Rare
2	1	0	0
2	2	3	4
2	3	0	1
3	1	5	1
3	2	4	2
3	3	0	3

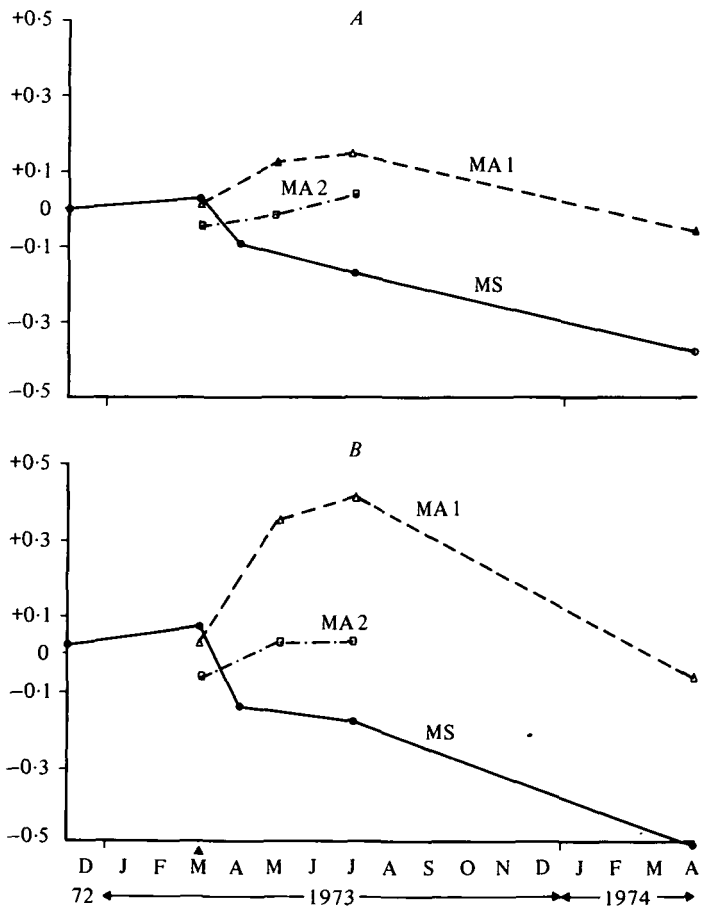

Fig. 4. Temporal changes in the association between Esterase 6 (2,3) and Phosphoglucomutase (2,3) in Manch replicates. MA1(2) = Manch alcohol replicate 1(2). A correlation, B relative linkage disequilibrium ($D' = D/Dmax$).

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

(*b*) Disequilibrium within the same pair of classifications is not uniform over populations, confirming observations of previous workers (Mukai *et al.* 1974; Langley *et al.* 1974).

(*c*) Lack of homogeneity was observed among populations, as the highly significance linkage disequilibria at the level of pairs were concentrated in a population in which a previous study by Charlesworth & Charlesworth (1973) uncovered some instances of linkage disequilibrium.

(*d*) A temporal instability of linkage disequilibrium values occurred in the populations that were subjected to that kind of check. Congruent with that is the fact that for the three loci in the Amherst population also studied by Charlesworth & Charlesworth (1973), the two significant and one suspiciously high 2×2 associations in our sample (*Est*–6–*Pgm* 23–12, 23–13, and *Pgm*–*Aldox* 12–23) failed to be significant in their sample. It is opportune to note that those temporal observations come from populations that have been subjected to a new environment. Temporal instability of linkage disequilibria in a natural population has also been reported (Langley *et al.* 1977).

The attribution of these characteristics to any of the three factors that are known as possible modifiers of the random association among loci-drift, migration, and selection seems to be difficult.

Migration, without the contribution of selection, has been generally discarded as a factor producing linkage disequilibrium in wild populations of *Drosophila* (Lewontin, 1974). This argument does not necessarily apply to laboratory conditions, where there is not homogeneity of gene frequencies and where there are considerable differences in adaptation to this environment between the long-standing populations and those newly arrived. Nevertheless, as we have seen earlier, it does appear that this explanation is unlikely to hold in our study, at least in the most striking cases of between pairs linkage disequilibrium.

Approximate predictions of disequilibrium parameters can be made under the neutral-drift hypothesis at the level of pairs of loci (Hill & Robertson, 1968; Ohta & Kimura, 1969; Sved, 1971; Sved & Feldman, 1973; Hill, 1975*b*) and, to some extent, at the level of several loci in terms of the population size and the recombination fractions.

We have estimated, by lethal allelism (Nei, 1968), the effective sizes of the populations involved in the present study (Malpica & Briscoe, 1981). For the populations Stellenbosch (sample 6), Riudevella (sample 7), Amherst (sample 8) and Kaduna those estimates are 427, 1123, 383 and 190 respectively. Those numbers contrast with the census ($\sim$ 5000) maintained by this type of population cage (Kinross & Robertson 1970).

In contrast, all 110 random crosses studied between lethal lines from sample 1 (Mancha Wild) showed complementation. That will give a minimum estimate for the population number of 3548, allowing a safety margin of 5%. Subsequent estimations made on samples taken from the same location and season, point to population numbers on the order of 10^4 (J. M. Vassallo & J. M. Malpica, unpublished results). Furthermore, there is homogeneity of allozyme frequency and

lethality over years and over considerable distances (D. A. Briscoe, J. M. Vassallo & J. M. Malpica, unpublished results). All these observations point to a reasonable size for the Mancha wild population.

We have not estimated the population number of Mancha cages, but, as it has been already pointed out, the effective population number subtained in this type of population cages is on the order of hundreds (Malpica & Briscoe, 1981).

Table 8. *Comparison between expected and observed distributions of the significance of the tests of independence in Mancha wild and the rest of the samples*

No. of classifications			Significance classes 75–100	50–75	25–50	0–25	Likelihood ratio
2	Total	Exp.	31·42	11·22	12·13	9·24	
		Obs.	35	8	8	13	4·372
	Mancha W.	Exp	7·37	3·07	2·28	2·29	
		Obs.	3	5	2	5	6·780
	Rest	Exp.	24·06	8·15	9·85	6·95	
		Obs.	32	3	6	8	8·579
3	Total	Exp.	20·61	14·81	13·99	14·60	
		Obs.	28	9	11	16	5·854
	Mancha W.	Exp.	6·05	4·68	4·47	4·81	
		Obs.	1	3	6	10	11·927
	Rest	Exp.	14·56	10·13	9·52	9·80	
		Obs.	27	6	5	6	14·754
4	Total	Exp.	10·17	9·21	9·30	9·32	
		Obs.	17	5	4	12	10·666
	Mancha W.	Exp.	3·81	3·75	3·73	3·72	
		Obs.	1	0	3	11	22·553
	Rest	Exp.	6·36	5·46	5·57	5·60	
		Obs.	16	5	1		28·625

Therefore our samples belong to two different populational types, and, one of the types has only one representative. Nevertheless, it seems worthwhile to look at the contribution of each type to the lack of fit to the independence hypothesis, as for small population sizes associations are to be expected. This is done in Table 8, where it is shown that the sample from Mancha Wild is the one causing most of the excess at the class of more significant tests, and more so when the number of classifications involved in the analysis increases.

The comparison between observed and expected distributions of significance of tests for the sample from Mancha Wild is presented graphically in Figure 5, where the data have been grouped in four 25%-wide significance classes. It can be seen that at the two-classifications level independence test (Fig. 5*A*) there is a deviation from the expected distribution under the independence hypothesis that is nearly significant ($G_{(3)} = 6{\cdot}780$, $0{\cdot}1 > P > 0{\cdot}05$, and that this deviation comes mainly from an excess of observations in the $\leqslant 25\%$-significance class ($G_{(1)} = 3{\cdot}005$, $0{\cdot}1 > P > 0{\cdot}05$). Deviations from the independence hypothesis become stronger as the number of classifications increases (Fig. 5*B*, *C*) ($G_{(3)} = 11{\cdot}926$, $P < 0{\cdot}01$ and

https://doi.org/10.1017/S0016672300020735 Published online by Cambridge University Press

$G_{(3)} = 22{\cdot}553$, $P < 0{\cdot}001$ for three and four classifications respectively). In contrast with the overall pattern, no excess of observations in the conservative end of the distribution is found.

The deviation from the expected distribution for the 2×2 interactions (Fig. 5*D*), though suggesting an excess of observations at the significant end class, fails to

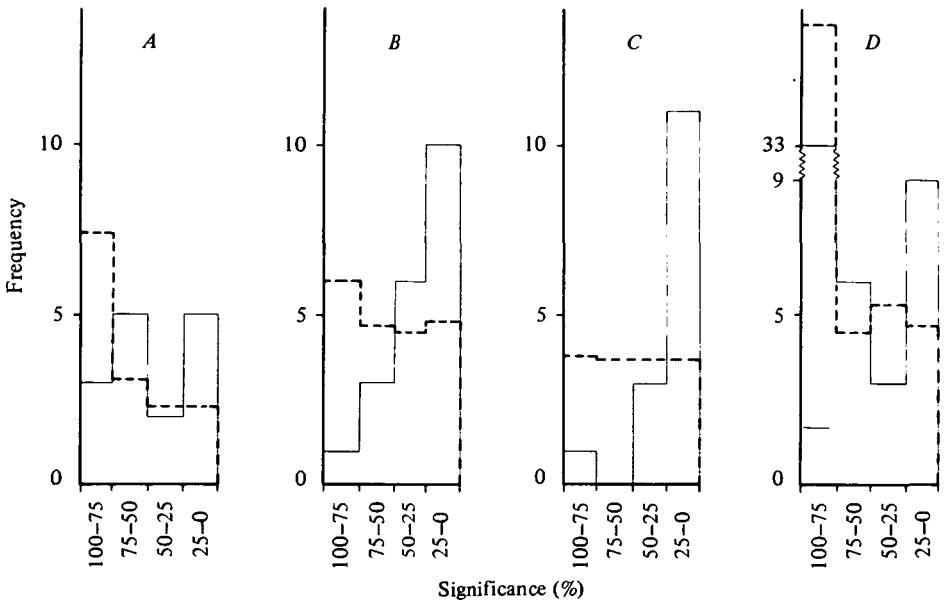

Fig. 5. Expected (dotted lines) and observed (solid lines) distributions of significance of tests in Mancha wild sample; A, B, C independence tests, with two, three and four classifications; D, 2×2 interaction tests.

be significant itself ($G_{(3)} = 5{\cdot}020$, $P < 0{\cdot}2$). Again, of the 74 possible $2 \times 2 \times 2$ interaction tests, only three are allowed more than the observed configuration, and all these three are in their most probable configuration. When the data are pooled to two classes, or when analysed for interactions by the same method, but without splitting it into 2^n tables, virtually the same results are obtained.

Then, we are faced with the fact that, in our data, although the strongest cases of linkage disequilibria are found in small populations, as neutral theory predicts, and has been previously observed (Mukai *et al.* 1974, Laurie-Ahlberg & Weir, 1979), it is the large wild population in which independence among loci is untenable.

Concerning the generality of the above observations, it should be pointed out that we have dealt with only one sample from a large wild population, and that there are few gametic samples of third chromosomes in the literature suitable for independence hypothesis testing. In fact, we have been able to find only one such set of data (Makai *et al.* 1974), and these show lack of independence among allozyme loci (three of them) within the standard chromosomal arrangement.

However, in that case the data were obtained from a laboratory population and the dependence can be explained by between-pair interactions.

As for the meaning of the present findings, the scarcity of theory for linkage disequilibrium in the case of interaction between selection and drift and for independence among several loci under the neutral hypothesis (Hill,1975*a*), does not allow further discussion either of the results from small populations or of the lack of independence among several loci in the Mancha Wild sample. The observed interactions at the level of pairs in the Mancha Wild sample are not strong enough as to overtake our sampling dispersion, so, discussion of other causes of dispersion seems unwarranted.

Nevertheless, it is not too uncommon to observe in our samples the presence of linkage disequilibrium between isozyme markers and lethality. Spurious heterosis can be attributed to neutral loci, due to association with detrimentals (Ohta, 1971). Our results indicate that this can be the case even with the high adult number yield of a population cage.

The latter argument may be generalized, because of the lack of independence between lethality and loci when there are several of them involved, and not necessarily under low population number as the Mancha Wild sample shows. Multiple locus observations of such selective-like pattern may just be a consequence of such types of associations, and in such a situation, the selective pattern could become more strong as the number of loci gets larger.

From previous reports (Langley *et al.* 1974; Langley *et al.* 1977, Cochrane & Richmond, 1980) and the present work it is becoming apparent that the region *Est*-6–*PGM* is prone to show linkage disequilibrium. This is at least a hint on where to look in future work for associations and agents causing them.

We thank Professor A. Robertson and Dr W. G. Hill for valuable discussions. The help of Professor F. García-Olmedo in the preparation of the manuscript is gratefully acknowledged.

REFERENCES

ANDERSEN, A. H. (1974). Multidimensional contingency tables. *Scandinavian Journal of Statistics* **1**, 115–127.

BARKER, J. S. F. (1979). Inter locus interactions: A review of experimental evidence. *Theoretical Population Biology* **16**, 323–346.

BARTLETT, M. S. (1935). Contingency table interactions. *Journal of the Royal Statistical Society Suppl.* **2**, 248–252.

CAVENER, D. R. & CLEGG, M. T. (1978). Dynamics of correlated genetics systems. IV. Multilocus effects of ethanol stress environments. *Genetics* **90**, 629–644.

CHARLESWORTH, B. & CHARLESWORTH, D. (1973). A study of linkage disequilibrium in populations of *Drosophila melanogaster*. *Genetics* **73**, 351–359.

COCHRANE, B. J. & RICHMOND, R. C. (1980). The effects of lethal selection on the EST-6 to PGM region of chromosome 3 in *Drosophila melanogaster*. *Genetics* **96**, 491–506.

FELDMAN, M. W., FRANKLIN, I. R. & THOMSON, G. (1974). Selection in complex genetic systems: I. The symmetric equilibria of the three locus symmetric viability model. *Genetics* **76**, 135–162.

FELDMAN, M. W., LEWONTIN, R. C., FRANKLIN, I. R. & CHRISTIANSEN, F. B. (1975). Selection

in complex genetic systems. III. An effect of allele multiplicity with two loci. *Genetics* **79**, 333–347.

FRANKLIN, I. R. (1971). Genetic variation at the Esterase-6 locus in *Drosophila melanogaster*. *Drosophila Information Service* **47**, 113.

FRANKLIN, I. R. and LEWONTIN, R. C. (1970). Is the gene the unit of selection? *Genetics* **65**, 703–734.

HILL, W. G. (1974*a*). Estimation of linkage disequilibrium in randomly mating populations. *Heredity* **33**, 229–239.

HILL, W. G. (1974*b*). Disequilibrium among several linked neutral genes in finite population. I. Mean changes in disequilibrium. *Theoretical Population Biology* **5**, 366–392.

HILL, W. G. (1975*a*). Non-random association of neutral linked genes in finite populations. In *Population Genetics and Ecology* (ed. S. Karlin and E. Nevo), pp. 339–376. New York: Academic Press.

HILL, W. G. (1975*b*). Linkage diequilibrium among multiple neutral alleles produced by mutation in finite population. *Theoretical Populations Biology* **8**, 117–126.

HILL, W. G. and ROBERTSON, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics* **38**, 226–231.

KINROSS, J. & ROBERTSON, A. (1970). Egg laying and survival rates in population cages of *Drosophila melanogaster*. *Drosophila Information Service* **45**, 83.

LANCASTER, H. O. (1951). Complex contingency tables treated by the partition of χ^2. *Journal of the Royal Statistical Society* B **13**, 242–249.

LANGLEY, C. H. (1977). Nonrandom associations between allozymes in natural populations of *Drosophila melanogaster*. In *Measuring Selection in Natural Populations* (ed. F. B. Christiansen and T. M. Fenchel), pp. 263–274. Berlin: Springer-Verlag.

LANGLEY, C. H., ITO, K. and VOELKER, R. A. (1977). Linkage disequilibrium in natural populations of *Drosophila melanogaster*. Seasonal variation. *Genetics* **86**, 447–454.

LANGLEY, C. K., TOBARI, Y. N. & KOJIMA, K. (1974). Linkage disequilibrium in natural populations of *Drosophila melanogaster*. *Genetics* **78**, 921–936.

LAURIE-AHLBERG, C. C. & WEIR, B. S. (1979). Allozymic variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster*. *Genetics*. **92**, 1295–1314.

LEWONTIN, R. C. (1974). *The Genetics Basis of Evolutionary change*. Columbia University Press: New York.

LINDSLEY, D. L. & GRELL, E. H. (1968). *Genetic variations of Drosophila melanogaster*. Carnegie Institute of Washington. Publication No 627.

MALPICA, J. M. & BRISCOE, D. A. (1981). Effective population number estimates of laboratory populations of *Drosophila melanogaster*. *Experientia*. (In the Press.)

MUKAI, T., WATANABE, T. K. & YAMAGUCHI, O. (1974). The genetic structure of natural populations of *Drosophila melanogaster*. XII. Linkage disequilibrium in a large local population. *Genetics* **77**, 771–793.

NEI, M. (1968). The frequency distribution of lethal chromosomes in finite populations. *Proceedings of the National Academy of Science, U.S.A.* **60**, 517–524.

NEIMANN SØRENSEN, A. & ROBERTSON, A. (1961). The association between blood groups and several production characteristics in three danish cattle breeds. *Acta Agriculturæ Scandinavica* **11**, 163–196.

O'BRIEN, S. J. & MACINTYRE, R. J. (1971). A biochemical genetic map of *Drosophila melanogaster*. *Drosophila Information Service* **46**, 89–93.

OHTA, T. (1971). Associative overdominance caused by linked detrimental mutations. *Genetical Research* **18**, 277–286.

OHTA, T. & KIMURA, M. (1969). Linkage disequilibrium due to random genetic drift. *Genetical Research* **13**, 47–55.

POULIK, M. D. (1957). Starch gel electrophoresis in a discontinuous system of buffers. *Nature* **180**, 1477–1479.

SHAW, C. R. & PRASAD, R. (1970). Starch gel electrophoresis of enzymes – A compilation of recipes. *Biochemical Genetics* **4**, 297–320.

SOKAL, R. F. & ROHLF, F. I. (1969). *Biometry*. San Francisco: Freeman.
SVED, J. A. (1971). Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theoretical Population Biology* **2**, 125–141.
SVED, J. A. & FELDMAN, M. W. (1973). Correlation and probability models for one and two loci. *Theoretical Population Biology* **4**, 129–132.