LINKAGE DISEQUILIBRIUM IN NATURAL POPULATIONS OF *DROSOPHILA MELANOGASTER*[1]

CHARLES H. LANGLEY,[2] YOSHIKO N. TOBARI[3] AND KEN-ICHI KOJIMA[4]

Manuscript received May 9, 1973
Revised copy received May 3, 1974

ABSTRACT

Two large, stable populations (Texas and Japan) of *Drosophila melanogaster* were surveyed at 21 allozyme loci on the second and third chromosomes and for chromosomal gene arrangements on those two chromosomes. Over 220 independent gametes were sampled from each population. The types and frequencies of the surveyed genetic variation are similar to those observed previously and suggest only slight differentiation among geographically distant populations. Linkage disequilibrium among linked allozymes loci is only slightly, if at all, detectable with these sample sizes. Linkage disequilibrium between linked inversions and allozymes loci is common especially when located in the same arm. These disequilibria appear to be in the same direction for most comparisons in the two population samples. This result is interpreted as evidence of similar selective environments (ecological and genetic) in the two populations. It is also noted that the direction of these linkage disequilibria appears to be oriented with respect to the gene frequencies at the component loci.

THE importance of linkage and epistasis in evolutionary change has been considered by many population geneticists (see FRANKLIN and LEWONTIN 1970). The consideration of these complex phenomena has been fostered by a realization that the selective effect of a gene substitution may vary greatly about the mean, depending on genetic background. On the other hand, conceptual simplicity and mathematical tractability have combined with the statistical evidence of quantitative inheritance to nurture the rich-field, single-locus theory. The application of electrophoretic technique now allows the investigation of nonrandom associations involving allozymes. Such nonrandomness, where it exists, could be the result of epistatic selection on various loci (LEWONTIN 1964) or random drift of gametic frequencies (HILL and ROBERTSON 1968; OHTA and KIMURA 1969).

Descriptive efforts are seldom definitive in population studies because of the many parameters that go unmeasured despite their obvious importance. In the case of allozymic variation, parameters such as mutation rates, recombination fractions, population size, and migration rates are rarely measured. All of these are possibly critical parameters in the interpretation of a survey of genetic varia-

[1] This study was supported by PHS Grant GM-15769.

[2] Mutagenesis Branch, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, North Carolina 27709.

[3] Department of Biology, Tokyo Metropolitan University, Tokyo, Japan.

[4] Department of Zoology, University of Texas at Austin, Texas; deceased November 14, 1971.

tion. However, the first step in understanding a natural phenomenon is its detailed description. There is now much data on allozymic variability in many organisms. It is more difficult to obtain data that is appropriate for observing correlations between allozymes and other genetic elements. Several surveys of this kind have used *Drosophila melanogaster* as the material. KOJIMA, GILLESPIE and TOBARI (1970) reported linkage disequilibrium between allozymes and closely linked inversions. No significant linkage disequilibrium was observed for other comparisons. MUKAI, METTLER and CHIGUSA (1971) also noted linkage disequilibrium between inversions and linked allozyme loci. Recently CHARLESWORTH and CHARLESWORTH (1973) reported several instances of significant linkage disequilibrium between allozymes in three population samples of *D. melanogaster*. PRAKASH and LEWONTIN (1968, 1971) and NAIR and BRNCIC (1971) reported associations between polymorphic inversions and allozyme loci. ZOUROS and KRIMBAS (1973) found linkage disequilibrium between two physiological and genetical linked allozyme loci.

The following report is the result of a survey of independent second- and third-chromosome gametes from two large, stable populations of *Drosophila melanogaster*. Up to twenty-one loci on the second and third chromosomes were observed electrophoretically and each gamete was also observed cytologically for chromosomal arrangements. The goal of this study is primarily to better describe the genetic structure of natural populations and secondarily to eliminate several possible population genetic models simply on the basis of descriptive results.

MATERIALS AND METHODS

Two populations of *Drosophila melanogaster* were sampled. The Brownsville sample consisted of males collected from a large, year-round population at the city dump in Brownsville, Texas, in March of 1970. The Katsunuma sample of males was collected in the Fall of 1970, in Katsunuma, Japan. Each wild male was mated to several virgin females from a balancer stock (balancer stocks: *SM1/Gla; Ubx^{130}/Sb* for the Brownsville lines and *SM1Pm; Ubx^{130}/Sb* for the Katsunuma lines). A single F_1 male, genotype *SM1/+; Ubx^{130}/+*, was backcrossed to the balancer stock. Males and virgin females from this mating, genotype *SM1/+; Ubx^{130}/+*, were mated to initiate the lines which were maintained in mass cultures. These lines were then assayed for karyotypes and allozymes.

The procedures of the cytological analysis and the nomenclature of the inversions were described in detail by YANG and KOJIMA (1972). To repeat briefly, for the chromosomal analysis males derived from each line were crossed with homozygous females of standard gene arrangement. The third instar larvae of the F_1 progeny were dissected for salivary gland preparations.

The electrophoresis of allozymes was according to a modified technique of JOHNSON (1966). The α-amylase technique was similar to that of MCCUNE (1969). References for the buffers and staining techniques are indicated in Table 1.

RESULTS

Allozymic variation:

Of the twenty allozymic systems scored in the two populations half were polymorphic, i.e. they exhibited a second allozyme at a frequency of 3% or greater. All these loci have been mapped by recombination (see Table 1) and several

TABLE 1

Allozyme techniques

This table shows allozymes assayed, their International Union of Biochemistry numbers, genetic map position, bridge and gel buffers, and references for the staining techniques. The references for the genetic map position are given in FOX, ABACCHEALI and URSPRUNG (1971).

Allozyme	I.U.B. number	Genetic map position	Gel* buffer	Bridge† buffer	Stain technique
Alcohol Dehydrogenase (ADH)	(1.1.1.1)	2–50.1 (*Adh*)	1	1	See STONE *et al.* (1968)
Aldehyde Oxidase (AO)	(1.2.3.1)	3–56.6 (*Aldox*)	3	3	See COURTRIGHT (1967)
Aldolase (ALD)	(4.1.2.13)	——	1	1	See GILLESPIE and KOJIMA (1968)
Alkaline Phosphotase (APH-larval); (APH-adult)	(3.1.3.1)	3–46.3 (*Aph*) larval 3–47.3 (*Aph*) adult	1	1	See JOHNSON (1966)
α-Amylase (AMY)	(3.2.1.1)	2–77.7 (*Amy*)	8	8	See MCCUNE (1969)
Esterase C (EST-C)	(3.1.1.2)	3–49.0 (*Est-C*)	1	1	See JOHNSON (1966)
Esterase-6 (EST-6)	(3.1.1.2)	3–36.8 (*Est-6*)	1	1	See JOHNSON (1966)
Fumerase (FUM)	(4.2.1.4)	——	1	1	See GILLESPIE and KOJIMA (1968)
Glutamate-oxaloacetate transaminase (GOT)	(2.6.1.1)	——	1	1	See DELORENZA and RUDDLE (1970)
α-Glycerophosphate Dehydrogenase (GPD)	(1.1.99.5)	2–20.5 (*α-Gpd*)	1	1	See GILLESPIE and KOJIMA (1968)
Hexokinase (HK)	(2.7.1.1)	——	3	3	See GILLESPIE and KOJIMA (1968)
Isocitrate Dehydrogenase (IDH)	(1.1.14.)	2–27.1 (Idh-NADP)	1	2a	See SHAW and KOEN (1968)
Leucine aminopeptidase (LAP-D and LAP-A)	(3.4.1.1)	3–98 (*Lap-A*) 3–98.3 (*Lap-D*)	1	1	See JOHNSON (1966)
Malate Dehydrogenase (MDH)	(1.1.1.3)	2–35.3 (*Mdh-1*)	1	1	See GILLESPIE and KOJIMA (1968)
Malic Enzyme (ME)	(1.1.1.40)	——	1	2a	See GILLESPIE and KOJIMA (1968)
Octanol Dehydrogenase (ODH)	(–.–.–.–)	3–49.2 (*Odh*)	1	1	See STONE *et al.* (1968)
Phosphoglucoisomerase (PGI)	(5.3.1.9)	——	2b	2b	See BREWER (1970)
Phosphoglucomutase (PGM)	(2.7.5.1)	3–43.4 (*Pgm*)	1	2a	See SHAW and KOEN (1968)
Xanthine Dehydrogenase (XDH)	(1.2.3.2)	3–52.35 (*ry*+)	1	1	See SHAW and KOEN (1968)

* Gel buffers:
1 —See POULIK (1957).
2b—See the Tris-Maleic acid-Na_2-EDTA buffer.
3 —See YOUNG, PORTER and CHILDS (1964).
8 —See MCCUNE (1959).

† Bridge buffers:
1 —See POULIK (1957).
2a—See the (0.687*M*) Tris-(0.157*M*) Citric acid buffer in SHAW and KOEN (1968).
2b—See the Tris-Maleic acid-Na_2-EDTA buffer in SHAW and KOEN (1968).
3 —See YOUNG, PORTER and CHILDS (1964).
8 —See MCCUNE (1959).

TABLE 2

A comparison of frequencies of allozymes and inversions in the Brownsville and Katsunuma populations of Drosophila melanogaster*

		f(1)	*f*(2)	*f*(3)	*f*(4)	*f*(5)	*f*(6)	*f*(7)	*N*
α-Gpd	*B*	0.109	0.891	....	....	....	....	....	239
	K	0.167	0.833	....	....	....	....	....	264
Adh	*B*	0.732	0.268	....	....	....	....	....	239
	K	0.395	0.605	....	....	....	....	....	263
Mdh	*B*	0.970	0.030	....	....	....	....	....	236
	K	1.000	....	....	....	....	....	....	263
Amy	*B*	0.902	0.017	0.060	0.004	0.004	0.009	0.004	233
	K	0.904	0.033	0.023	0.031	....	....	0.004	259
Est-6	*B*	0.442	0.558	....	....	....	....	....	240
	K	0.830	0.170	....	....	....	....	....	264
Pgm	*B*	0.051	0.898	0.051	....	....	....	....	235
	K	0.333	0.667	....	....	....	....	....	264
Est-C	*B*	0.104	0.858	0.038	....	....	....	....	239
	K	0.140	0.799	0.061	....	....	....	....	264
Odh	*B*	0.117	0.883	....	....	....	....	....	240
	K	0.163	0.837	....	....	....	....	....	264
Aldox	*B*	0.000	0.004	0.538	0.026	0.383	0.049	....	227
	K	0.008	0.008	0.784	0.023	0.177	....	....	259
Lap-D	*B*	0.224	0.776	....	....	....	....	....	228
	K	0.246	0.754	....	....	....	....	....	260
		Standard	*In(2L)t*	*In(2L)T*					*N*
2L	*B*	0.824	0.172	0.004	....	....	....	....	233
	K	0.816	0.180	0.004	....	....	....	....	233
		Standard	*In(2R)NS*	*In(2R)TA*	*2L-2R*(new)				
2R	*B*	0.743	0.253	0.004	....	....	....	....	233
	K	0.858	0.138	....	0.004	....	....	....	261
		Standard	*In(3L)P*	*In(3L)M*	*In(3L)TB*				
3L	*B*	0.747	0.236	0.013	0.004	....	....	....	233
	K	0.935	0.057	0.004	0.004	....	....	....	260
		Standard	*In(3R)P*	*In(3R)C*	*In(3R)MO*	*In(3R)K*	*In(3R)TC*		
3R	*B*	0.477	0.464	0.047	0.004	0.004	0.004	....	233
	K	0.613	0.180	0.176	0.023	0.004	0.004	....	261

Monomorphic allozyme loci	*Idh*	*Aph*	*Aph-2*	*Xdh*	*Lap-A*	*Ald*	*Fum*	*Got*	*Hk*	*Me*	*Pgi*
Brownsville	N.V.	(1)235	N.V.	N.V.	(1)225	(1) 1	N.V.	(1)234	N.V.	N.V.	N.V.
	248	(2) 1	227	239	(2) 3	(2)231	231	(2) 3	235	238	229
Katsunuma	N.V.	(1)258	—	N.V.	—	N.V.	N.V.	N.V.	N.V.	N.V.	N.V.
	265	(2) 1	—	264	—	263	262	263	263	263	261
		(3) 3									
		null 1									

* *f(i)* is the frequency of the i^{th} allozyme, ordered according to its relative migration toward the cathode. *N* is the number of gametes assayed. *K* and *B* indicate the Katsunuma and Brownsville samples respectively. N.V. indicates no variants observed.

are known by cytological mapping. The loci are distributed over most of chromosomes *2* and *3*. Heterogeneity between allozymic frequencies in the two population samples are measured by x^2.

Polymorphic loci on chromosome 2 (see Table 2):

α-Gpd—α-glycerophosphate dehydrogenase: These two allozymes are identical to those reported by KOJIMA, GILLESPIE and TOBARI (1970) and MUKAI, METTLER and CHIGUSA (1971). The slower migrating allozyme is 1 and the heterozygotes demonstrate an intermediate "hybrid" band. These are the only two *α-Gpd* allozymes seen in natural populations of *Drosophila melanogaster* and the faster migrating allozyme is the most frequent in all populations. The frequencies are not heterogeneous ($P<0.06$).

Adh—alcohol dehydrogenase: There are two allozymes at this locus; they are identical to those reported by KOJIMA, GILLESPIE and TOBARI (1970) and by MUKAI, METTLER and CHIGUSA (1971). The slower migrating band is 1 and a hybrid band is evident in heterozygotes. The faster allozyme, 2, is never found on a chromosome carrying the *In(2L)t* gene arrangement. From cytological mapping (GRELL, JACOBSON and MURPHY 1965) it is known that the *Adh* locus is located within the region, 34E3–35D1, and since the proximal breakpoint of *In(2L)t* is at 34A, recombination be tween these two elements is highly unlikely. The frequencies are heterogeneous at the $P \ll 0.001$ level.

Mdh—malic dehydrogenase: Two allozymes were detected in the Brownsville population; these are the same as those reported by MUKAI, METTLER and CHIGUSA (1971). The sample from the Katsunuma population contains only the 1 allozyme which is slower migrating. KOJIMA, GILLESPIE and TOBARI (1970) found only the 1 allozyme in the Katsunuma population in the previous year. The 2 allozyme is consistently found in the North American samples (KOJIMA, GILLESPIE and TOBARI 1970, and MUKAI, METTLER and CHIGUSA 1971). It is likely that the presence of the 2 allozyme is a distinguishing characteristic in the genetic structure of these two populations. The frequencies are heterogenous at the $P < 0.005$ level.

Amy—α-amylase: This allozymic system is an apparent gene duplication with electrophoretic variants segregating at both loci as well as the original (apparently nonduplicated) arrangement (BAHN 1967). Since the recombination between the two loci is so small the individual arrangements are treated as allozymic types: $Amy^{1} = 1$; $Amy^{1-2} = 2$; $Amy^{1-3} = 3$; $Amy^{1-6} = 4$; $Amy^{2-6} = 5$; $Amy^{2-3} = 6$; and $Amy^{3-6} = 7$. The slower migrating band of Amy^{3-6} is indistinguishable from the faster band of Amy^{1-3} and the faster band of Amy^{3-6} is indistinguishable from the faster band of Amy^{1-6}. The frequencies (pooling 5, 6, 7) are heterogeneous at the $P < 0.002$ level.

Polymorphic loci on chromosome 3:

Est-6—esterase: Two allogymes were found in the two populations. These are the same as those reported by KOJIMA, GILLESPIE and TOBARI (1970). The frequency in the Katsunuma population sample is similar to those reported by

KOJIMA, GILLESPIE and TOBARI (1970), while the frequency in the Brownsville data is considerably different. The frequency in the Brownsville sample is similar to that reported for the Raleigh, N.C. population (KOJIMA, GILLESPIE and TOBARI 1970). The frequencies are heterogeneous at the $P \ll 0.001$ level.

Pgm—phosphoglucomutase: The two slower-migrating allozymes (1 and 2) were detected in both populations, while a third faster-migrating allozyme was also found in the Brownsville population. This is similar to the results of KOJIMA, GILLESPIE and TOBARI (1970) in that the Katsunuma data show only the first two allozymes with the 1 allozyme in fairly high frequency. The North American populations (Brownsville and Raleigh, N.C.) exhibit a third allozyme. The frequencies of the allozymes in the two populations are heterogeneous at the $P \ll 0.001$ level.

Est-C—esterase: Three allozymes were detected at this locus in both the populations. These are the same as those reported by KOJIMA, GILLESPIE and TOBARI (1970). The frequencies of the allozymes are not heterogeneous ($P < 0.22$).

Odh—octanol dehydrogenase: The two allozymes found at this locus are identical to those reported by KOJIMA, GILLESPIE and TOBARI (1970) and the frequencies are similar to those they reported. These frequencies are not heterogeneous ($P < 0.14$).

Ao—aldehyde oxidase: A total of six allozymes were found: the five slower in the Katsunuma sample and the five faster (2,3,4,5, and 6) in the Brownsville sample. Allozymes 3 and 5 are, however, the most frequent types in both populations. The frequencies are heterogeneous at the $P \ll 0.001$ level. KOJIMA, GILLESPIE and TOBARI (1970) did not detect this variation because they utilized a different buffer system.

Lap-D—leucylaminopeptidase: Two alleles were detected in both populations at similar frequencies. The frequencies are not heterogeneous ($P < 0.56$).

Monomorphic loci:

The other loci as listed in Table 2 showed no variation except for a rare heterozygote (in *Lap-A*, *Ald*, and *Got*).

The two populations share all polymorphic systems (except *Mdh*) and the most frequent allozymes are found in each population. The Brownsville population does appear to be more similar to the North Carolina population and the Katsunuma population shows stability in both types and frequencies (see KOJIMA, GILLESPIE and TOBARI 1970 and MUKAI, METTLER and CHIGUSA 1971).

Variation in chromosomal arrangement:

In both populations polymorphisms for paracentric inversions were found in each arm of chromosomes *2* and *3*. A more detailed description of the inversions found is given in YANG and KOJIMA (1972). The frequencies of the various arrangements are listed in Table 2.

2L: One cosmopolitan inversion was found in both populations, *In(2L)t*. This inversion has breakpoints at 22D and 34A. One small inversion, *In(2L)T*, was found in both populations only once. This inversion has breakpoints at 30F and

36D and is previously unreported. The frequencies in the two populations are not heterogeneous ($P < 0.97$).

2R: One cosmopolitan inversion was found in both populations, *In(2R)NS*. The breakpoints of this inversion are 52A and 56F. One small and previously unreported inversion was found in the Brownsville population sample, *In(2R)TA*, with breakpoints at 50A and 53A. The frequencies in the two populations are heterogeneous in the $P < 0.006$ level.

3L: Two cosmopolitan inversions were found in both populations. *In(3L)P* was most common in both populations. The breakpoints of this inversion are 63E and 72E. *In(3L)M* has breakpoints at 66D and 71D. A single example of a previously unreported inversions, *In(3L)TB*, was found in each population. The breakpoints of this inversion are 70B and 75A. The frequencies are heterogeneous at the $P \ll 0.001$ level.

3R: Four previously reported, if not cosmopolitan, inversions were found to be present in both populations. *In(3R)P*, with breakpoints at 89C and 96A, was most common in both populations. The second most frequent inversion in both populations was *In(3R)C*, with breakpoints at 92D and 100C. *In(3R)MO* was found to be present at a frequency greater than 2% in the Katsunuma population, while only one example was found in the Brownsville population. The breakpoints of *In(3R)MO* are 93D and 98F. Only one example of inversions *In(3R)K* and *In(3R)TC* were found in each of the population samples *In(3R)TC* is previously unreported and has breakpoints at 84D and 91E. The breakpoints of *In(3R)K* are 86F and 96D. Although the standard arrangement is most frequent, in all cases the situation in 3R of the Brownsville population is unusual in that the standard and *In(3R)P* arrangements are almost equally common at 47%. The frequencies in the two populations are heterogeneous at the $P \ll 0.001$ level.

The genetic structures of the two populations are very similar in that the most frequent inversions are shared by both. It should be noted that the Brownsville population is significantly more heterozygous for chromosomal arrangements.

Nonrandomness among nonallelic elements—linkage disequilibrium:

The nonrandomness of nonallelic elements is measured in two ways in this report: linkage disequilibrium or the covariance between loci, D and the correlation coefficient, R.

$$D = x_1x_4 - x_2x_3$$

$$R = D/\sqrt{pq(1-p)(1-q)}$$

where $x_1 = f(AB)$
$x_2 = f(Ab)$
$x_3 = f(aB)$
$x_4 = f(ab)$
$p = f(A) = x_1 + x_2$
$q = f(B) = x_1 + x_3$

A pooling procedure was followed; it combined all rarer alleles. This establishes

TABLE 3

*Associations between linked allozymes in the Brownsville and Katsunuma populations**

		g_1	g_2	g_3	g_4	N	χ^2	D	R
α-Gpd—Adh	*B*	158	54	16	10	238	1.99	+0.013	+0.091
(2)-(1)	*K*	85	134	19	25	263	0.29	—0.006	—0.033
α-Gpd—Amy	*B*	188	19	22	4	233	1.00	+0.006	+0.066
(2)-(1)	*K*	196	20	39	5	260	0.19	+0.003	+0.027
Adh—Amy	*B*	150	20	60	3	233	2.53	—0.014	—0.104
(1)-(1)	*K*	92	10	142	15	259	0.00	—0.001	—0.004
Est-6—Pgm	*B*	119	12	92	12	235	0.36	+0.006	+0.039
(2)-(2)	*K*	36	9	140	79	264	4.34	+0.023	+0.128
Est-6—Est-C	*B*	114	20	91	14	239	0.12	—0.004	—0.023
(2)-(2)	*K*	38	7	173	46	264	0.69	+0.008	+0.051
Est-6—Odh	*B*	120	14	92	14	240	0.44	+0.007	+0.043
(2)-(2)	*K*	40	5	181	38	264	1.07	+0.009	+0.064
Est-6—Lap-D	*B*	98	30	79	21	228	0.19	—0.006	—0.029
(2)-(2)	*K*	35	9	161	55	260	0.49	+0.007	+0.044
Pgm—Est-C	*B*	180	31	21	3	235	0.08	—0.002	—0.019
(2)-(2)	*K*	140	36	71	17	264	0.05	—0.003	—0.013
Pgm—Odh	*B*	187	24	20	4	235	0.58	+0.005	+0.049
(2)-(2)	*K*	152	24	69	19	264	2.72	+0.018	+0.102
Pgm—Ao	*B*	108	97	14	8	227	0.96	—0.010	—0.065
(2)-(3)	*K*	133	40	70	16	259	0.69	—0.010	—0.052
Pgm—Lap-D	*B*	158	47	19	4	228	0.36	—0.005	—0.040
(2)-(2)	*K*	133	40	63	24	260	0.62	—0.010	—0.049
Est-C—Odh	*B*	183	22	28	6	239	1.35	+0.008	+0.075
(2)-(2)	*K*	172	39	49	4	264	3.72	—0.018	—0.119
Est-C—Ao	*B*	101	93	21	12	227	1.52	—0.014	—0.082
(2)-(3)	*K*	163	44	40	12	259	0.08	+0.003	+0.018
Est-C—Lap-D	*B*	151	44	26	7	228	0.03	—0.002	—0.011
(2)-(2)	*K*	163	45	33	19	260	5.00	+0.024	+0.138
Odh—Ao	*B*	111	88	11	17	227	2.69	+0.018	+0.109
(2)-(3)	*K*	182	36	21	20	259	21.20	+0.043	+0.286
Odh—Lap-D	*B*	155	45	22	6	228	0.02	—0.001	—0.008
(2)-(2)	*K*	160	58	36	6	260	2.88	—0.017	—0.105
Ao—Lap-D	*B*	95	26	81	24	226	0.06	+0.003	+0.016
(3)-(2)	*K*	150	53	45	11	259	0.99	—0.011	—0.062
Est-6—Ao	*B*	68	60	54	45	227	0.05	—0.003	—0.014
(2)-(3)	*K*	172	43	31	13	259	1.96	+0.013	+0.087

* g_1 is the frequency of the gametic type formed from the alleles indicated in the parentheses under the respective loci, e.g., *α-Gpd—Adh*:g_1 is the frequency of the *α-Gpd2—Adh1* gamete, g_2—*α-Gpd2—Adh2*, g_3—*α-Gpd1—Adh1*, and g_4—*α-Gpd1—Adh2*. In cases involving more than 2 allozymes pooling was done as described in the text. N is the total number of gametes assayed for both loci. χ^2 is the 2×2 χ^2 with 1 degree of freedom. D and R are explained in the text.

two alleles at each locus (for both allozymes and chromosome arrangements in four arms). The pooling facilitates statistical analysis through the elimination of small expected numbers in the χ^2. It should be noted that in almost every system there are only two allozymic types with significant frequencies. The effect of his pooling will be indicated later. All significant levels are determined by 2×2 χ^2 unless otherwise indicated.

Nonrandom association among unlinked elements: The lines constitute a sample of male gametes from a natural population. Although the main concern here is with elements linked to one another, there is at least the possibility that interchromosomal epistasis could generate nonrandom associations among elements on different chromosomes.

Of the 40 comparisons involving elements on different chromosomes, three (0.075) were significant from the Brownsville data (*Pgm—Amy*, $P < 0.05$; *Odh*—2R, $P < 0.05$; *Amy*—2R, $P < 0.05$). In the Katsunuma data three (0.075) comparisons were significant (*Est-C—Amy*, $P < 0.05$; *Est-C*—2R, $P < 0.05$; *Lap-D*—2L, $P < 0.05$). Although several are significant at the 0.05 level no apparent nonrandomness exists in the 80 comparisons.

Nonrandom associations among linked allozymes: Table 3 contains the gametic data for various two-locus comparisons of linked allozymes. There are no significantly nonrandom associations in the Brownsville data. The Katsunuma data contain three (0.17) significant comparisons (*Odh—Ao*, *Est-6—Pgm*, *Est-C—Lap-D*).

Figure 1 is a graph of the absolute R values from Table 3 and some similar data from the 1968 sample of METTLER and CHIGUSA (1971). The abscissa is r_e, an

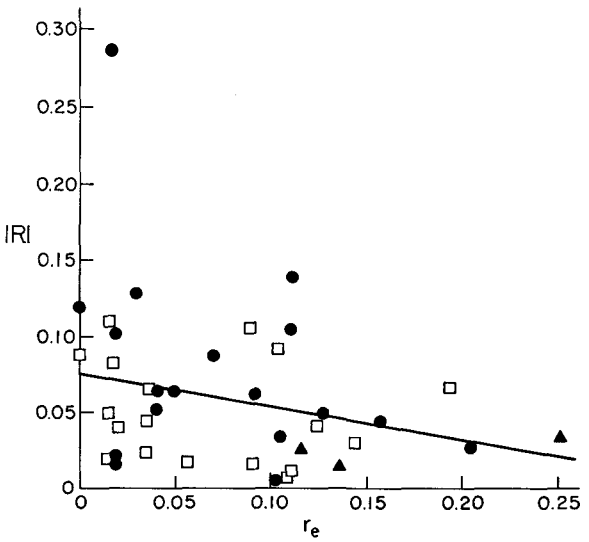

FIGURE 1.—The absolute value of the correlation coefficient for comparisons between linked allozymes plotted against r_e (see the text). The straight line is the least squares regression line ($|R| = 0.0758–0.2123r_e$). Brownsville (□), Katsunuma (●), Raleigh, N.C. 1968 sample from MUKAI, METTLER and CHIGUSA (1971) (▲).

estimate of the amount of recombination between the two allozymes in the population sampled. The effective recombination, r_e, is calculated based on the observed frequencies of gene arrangements assuming (1) no crossing over in males, (2) no crossing over in chromosome arms heterozygous for inversions, and (3) Hardy-Weinberg frequencies. The linear regression is not necessarily the appropriate model. It does, however, suggest a negative trend of the magnitude of disequilibrium with the increasing recombination ($|R| = 0.076 - 0.212r_e$); the correlation coefficient of (R) with r_e is -0.244 with $N = 39$, $P < 0.07$ if the points are assumed to be independent.

Nonrandom associations among linked allozymes and inversions: Table 4 contains the gametic data for the various comparisons involving allozymes and inversions on the same chromosome. Of the 36 combinations in Table 4, eleven (11) have a probability less than 0.05 of having occurred by chance. Nine (9) of these eleven involved allozymes and inversions in the same chromosome arm. The mean $|R|$ for allozymes and inversions in the same arm is 0.199, while the mean $|R|$ for those on opposite arms is 0.063. Also note that fourteen of eighteen paired comparisons have the same sign. If we assume that the expected frequency of positive and negative correlations equals 0.5 and that the B and K samples are independent, then the probability of fourteen out of eighteen matched pairs having the same sign is less than 0.01. If we take as given that four out of eighteen are different, we find that these four are in fact nonrandomly distributed among the four chromosome arms (all involve *3R*; $P < 0.01$). That is, *3R* appears to be less similar in the two samples (B and K) than the other three arms. MUKAI, METTLER and CHIGUSA (1971) and MUKAI, WATANBE and YAMAGUCHI (1974) found in consecutive annual samples that the linkage disequilibria for the allozymes and inversions in the same arm (chromosome 2) were all in the same direction as those found in the Brownsville and Katsunuma population samples, while those in opposite arms were not all consistent. KOJIMA, GILLESPIE and TOBARI (1971) found the direction and magnitude of the linkage disequilibria between inversions and allozyme loci (*Adh* and *α-Gpd*) on chromosome 2 to be the same as reported here for the Katsunuma population. Those on chromosome 3 were not consistent (all involved *3R*).

Three locus interactions: Interactions associated with three loci were also examined in $2 \times 2 \times 2$ χ^2's. These χ^2's were calculated for (1) all triplets of adjacently linked allozymic loci (such as *Est-6* × *Pgm* × *Est-C* and *Odh* × *Ao* × *Lap-D*), (2) all triplets of adjacently linked allozymic pairs and linked inversions (such as *Est-6* × *Pgm* × *3R* and *Pgm* × *Est-C* × *3L*), and (3) all triplets of a linked allozymic locus and the two chromosome arms (such as *Pgm* × *3L* × *3R* and *Adh* × *2L* × *2R*). In no case was χ^2 for this comparison significantly greater than the sum of the χ^2's for the three composite two-way comparisons. This indicates that the amount of third-order interaction detectable by the technique used in this study is small.

The effect of pooling: All the comparisons discussed so far were done among the two allelic systems. In cases where there were more than two alleles at a locus the rarer alleles were pooled. In only one case did this pooling "obscure an interaction"; i.e., (1) the χ^2 on the unpooled data was significant ($P < 0.05$)

TABLE 4

*Associations between inversions and linked allozymes or inversions in the Brownsville and Katsunuma populations**

Allozymes and inversions in the same arm		g_1	g_2	g_3	g_4	N	χ^2	D	R
α-Gpd—2L	*B*	165	41	26	0	232	6.29	—0.020	—0.165
(2)-(S)	*K*	174	43	39	5	261	1.74	—0.012	—0.082
Adh—2L	*B*	129	41	62	0	232	18.16	—0.047	—0.280
(1)-(S)	*K*	54	47	159	0	260	90.32	—0.111	—0.589
Amy—2R	*B*	148	59	22	1	230	6.26	—0.022	—0.165
(1)-(S)	*K*	198	36	24	1	259	2.39	—0.009	—0.096
Est-6—3L	*B*	78	52	96	6	232	35.48	—0.084	—0.391
(2)-(S)	*K*	38	7	206	10	261	7.30	—0.016	—0.167
Pgm—3L	*B*	152	56	21	2	231	3.66	—0.016	—0.126
(2)-(S)	*K*	160	15	84	2	261	3.69	—0.014	—0.119
Est-C—3R	*B*	99	100	11	22	232	3.06	+0.020	+0.115
(2)-(S)	*K*	124	84	36	17	261	1.23	—0.013	—0.069
Odh—3R	*B*	99	105	11	17	232	0.84	+0.010	+0.060
(2)-(S)	*K*	149	70	11	31	261	26.01	+0.057	+0.316
Ao—3R	*B*	80	42	29	76	227	32.57	+0.094	+0.379
(3)-(S)	*K*	142	61	17	39	259	29.03	+0.067	+0.335
Lap-D—3R	*B*	81	96	28	23	228	1.33	—0.016	—0.076
(2)-(S)	*K*	123	72	36	28	259	0.98	+0.013	+0.060
Allozymes and inversions in opposite arms									
α-Gpd—2R	*B*	151	55	21	5	232	0.67	—0.007	—0.054
(2)-(S)	*K*	185	32	39	5	261	0.34	—0.005	—0.036
Adh—2R	*B*	129	41	43	19	232	1.01	+0.013	+0.066
(1)-(S)	*K*	90	11	133	26	260	1.51	+0.012	+0.076
Amy—2L	*B*	172	35	18	5	230	0.34	+0.004	+0.038
(1)-(S)	*K*	191	43	20	5	259	0.04	+0.001	+0.012
Est-6—3R	*B*	67	63	43	59	232	2.02	+0.023	+0.093
(2)-(S)	*K*	21	24	139	77	261	4.91	—0.025	—0.137
Pgm—3R	*B*	95	113	15	8	231	3.17	—0.018	—0.177
(2)-(S)	*K*	108	67	52	34	261	0.04	+0.003	+0.012
Est-C—3L	*B*	146	53	28	5	232	1.99	—0.014	—0.093
(2)-(S)	*K*	194	14	50	3	261	0.08	—0.002	—0.017
Odh—3L	*B*	151	53	23	5	232	0.87	—0.009	—0.061
(2)-(S)	*K*	204	15	40	2	261	0.25	—0.003	—0.031
Ao—3L	*B*	92	30	78	27	227	0.04	+0.003	+0.013
(3)-(S)	*K*	191	12	51	5	259	0.65	+0.005	+0.050
Lap-D—3L	*B*	127	50	44	7	228	4.45	—0.025	—0.140
(2)-(S)	*K*	180	15	62	2	259	1.64	—0.009	—0.080
2L—2R	*B*	140	52	33	8	233	1.01	—0.011	—0.066
(S)-(S)	*K*	183	30	41	7	261	0.01	+0.001	+0.006
3L—3R	*B*	83	91	28	31	233	0.00	+0.000	+0.002
(S)-(S)	*K*	151	93	9	8	261	0.54	+0.005	+0.045

* g_1 is the frequency of the gametic type composed of the allozyme number in the first parentheses (i) and the standard gene arrangement (s). g_2—allozyme (i) and those gene arrangements other than the standard (s'). g_3—allozymes (i') other than (i) and the standard gene arrangement. g_4—(i') and (s'). N is the total number of gametes assayed for both the allozyme and gene arrangement. χ^2 is the 2×2 χ^2 with 1 degree of freedom. D and R are explained in the text.

while the pooled χ^2 was not and (2) the expected numbers were acceptably large (greater than 1). The χ^2 for the interaction between the *Est-6* and *Est-C* loci in the Katsunuma data is 5.97, df = 2 ($P < 0.05$) for the unpooled data while the pooling of the *1* and *3* alleles leads to a 2×2 $\chi^2 = 0.69$, df = 1 ($P < 0.70$). Although, in general pooling obscures information, given the sampling sizes and allele frequencies in these data it appears that the pooling did not seriously effect the estimation of the amount of nonallelic interaction. It is, however, clear that had our sample size been greater many of the interactions among rarer alleles would have been obscured by the pooling.

DISCUSSION

The results reported above provide three significant pieces of information about the genetics of natural populations. (1) Further evidence of the relatively uniform distribution of allozyme polymorphisc in Drosophila is reported. This is not indicative significant random genetic drift. (2) The linkage disequilibrium among linked allozymes are not as intense as predicted by the simulations of FRANKLIN and LEWONTIN (1970) nor do they appear to be as intense as CHARLESWORTH and CHARLESWORTH (1973) suggested. (3) The linkage disequilibria between allozyme loci and inversions in the same arm are large and show geographic and temporal stability. This is similar to the results of PRAKASH and LEWONTIN (1968, 1971), where certain gene arrangements and closely linked allozymes were found to be correlated across populations and species.

Patterns of polymorphism: The patterns of allozymic and chromosomal polymorphism in the present two population samples are similar in quantity or quality to earlier observations of *Drosophila melanogaster*. The comparison of these allozyme and inversion frequencies with those reported previously suggests some geographical differentiation in the form of temporally stable frequency differences (WATANABE 1967; BERGER 1971; KOJIMA, GILLESPLE and TOBARI 1970; MUKAI, METTLER and CHIGUSA 1971). There is no case of a frequent element (allozyme or inversion with frequency greater than 0.12) in one population being completely absent in another population. Consequently yearly samples in the Raleigh, N.C. and Katsunuma, Japan populations indicates clear temporal stability of both allozyme and inversion frequencies. The overall similarity in genetic polymorphism among these widely distributed populations suggests either common selective regimes or substantial migration and/or historically recent common origin (see later discussion) in the absence of selection.

Linkage disequilibria among linked allozymes: Linkage disequilibria among allozyme or blood group loci have been measured in several natural populations (SINNOCK and SING 1972). Because of population structure differences, the data presented here are directly comparable only to that collected in Drosophila. MUKAI, METTLER and CHIGUSA (1971) reported only one out of 12 comparisons significant in samples from two years of a stable population of *D. melanogaster* in Raleigh, N.C. None of these second chromosome loci were closely linked. CHARLESWORTH and CHARLESWORTH (1973) reported 4 out of 30

comparisons significant in three samples from 2 populations of *D. melanogaster*. Some of the third chromosome loci they examined were closely linked. Of the 36 linked allozyme-allozyme comparisons examined in this study, three were significant. Given the number of comparisons and significance levels, it is reasonable to judge only one of these (*Odh—Ao*) as a reliable observation of nonrandom association among allozymes. The results substantiate the two major conclusions to be drawn from previous reports. (1) The extreme level of linkage disequilibria observed in the simulations of FRANKLIN and LEWONTIN (1970) is not present among allozymes in natural populations of *D. melanogaster*, and (2) complete linkage equilibrium among allozymes is also not the rule.

CHARLESWORTH and CHARLESWORTH (1973) suggest that there is no correlation between linkage and magnitude of linkage disequilibrium. Although the data in Figure 1 do not refute this suggestion, they do call it into question. Since we do not know the distributions of R under random drift or selection, it is impossible to discern whether the lack of observed correlation of linkage disequilibria with linkage is a significant observation.

The causes of the observed level of linkage disequilibrium among allozymes cannot be discriminated at the present. However, simple random genetic drift appears to be an unlikely candidate because most indications are that population sizes and/or migration are large (MUKAI, METTLER and CHIGUSA 1971; MUKAI, WATANABE and YAMAGUCHI 1974; CHARLESWORTH and CHARLESWORTH 1973). Gene frequency distributions are similar and temporally stable over large geographic areas; this is not indicative of significant random genetic drift. More directly, estimates of recessive lethal frequencies and allelism rates suggest large effective population sizes (METTLER and CHIGUSA 1971; YOSHIKAWA and MUKAI 1970).

Linkage disequilibria between linked allozymes and inversions: KOJIMA, GILLESPIE and TOBARI (1970) reported several instances of nonrandomness among linked allozymes and inversions. They noted that the magnitude of the linkage disequilibrium is correlated with the linkage between the allozyme locus and the inversion. In the larger sample reported by MUKAI, METTLER and CHIGUSA (1971) a similar trend is apparent when r_e is used as the estimate of the linkage. In the results reported here this effect is most clearly seen when the linkage disequilibria between inversions and allozymes in the same chromosome arm are compared with linkage disequilibria between inversions and allozymes in opposite arms (the means of $|R|$ are 0.119 and 0.063, respectively). The amount of linkage disequilibrium between allyzomes and inversions in opposite arms is small and similar to that observed among linked allozymes. It appears that the crossing over suppression within arms containing inversion polymorphisms allows the maintenance of the linkage disequilibria between the inversions and allozyme loci in those arms.

The similarities in sign of comparisons in opposite arms may well be a sampling error; however, those between inversions and allozymes in the same arm appear to be real. The associations show geographic uniformity and temporal

stability. A similar situation has been described by PRAKASH and LEWONTIN (1968, 1971), for two other *Drosophila* species, *pseudoobscura* and *persimilis*.

The significance of these linkage disequilibria among inversions and allozymes is not clear as of those reported by PRAKASH and LEWONTIN (1968, 1971) in that a recent (last several centuries) expansion of *D. melanogaster* over the globe could possibly explain the observed similarities. Thus the significance of these observed similarities must be interpreted only as further evidence of the importance of gene arrangement polymorphisms. The evolutionary dynamics of those loci in the same arm as an inversion polymorphism can only be understood in relationship to the evolutionary behavior of the inversion polymorphism itself. This must be true whether the allozymes selectively interact with gene arrangements or not.

Directional disequilibrium: In the past there has been little consideration of the sign of linkage disequilibrium because it appeared to be completely arbitrary. There was no way to determine "coupling" and "repulsion" types *a priori*. What follows is, however, a nonarbitrary procedure for orienting linkage disequilibrium; "coupling" gametes are composed of the most frequent alleles or the least frequent alleles. The formulation is as follows:

gene frequencies	$f(A) > 0.5 > f(a)$
	$f(B) > 0.5 > f(b)$
gamete frequencies	$f(A,B) = x_1$
	$f(A,b) = x_2$
	$f(a,B) = x_3$
	$f(a,b) = x_4$

then

$$D_\omega = x_1x_4 - x_2x_3$$

$$R_\omega = D_\omega/\sqrt{f(A)f(a)f(B)f(b)}$$

The estimation of R_ω is undefined at $f(A) = 0.5$ or $f(B) = 0.5$. If $f(\mathrm{A}) = 0.5$ or $f(B) = 0.5$ the error in the sign is large. Cases where allele frequencies are near 0.5 should be avoided.

In the accompanying paper (LANGLEY and CROW 1974) it is demonstrated that D_ω can be expected to be negative at equilibrium under a very reasonable selective model. The interest here is just to note the majority of negative R_ω values among the comparisons of linked allozymes and inversions, excluding *3R*. *3R* is excluded because it is atypical in this respect also and because in the Brownsville data the frequency of the standard arrangement is approximately 0.5. In the Katsunuma and Brownsville data 0.71 of the R_ω are negative. If these comparisons were all independent and probability of this proportion would be $\mathrm{P} < 0.06$. Further investigation of nonrandom associations involving allozymes (especially in other species) will confirm or deny the suggestion of these observations. It should be noted that only a selective model could explain a significant proportion of R_ω being less than zero.

LITERATURE CITED

BAHN, E., 1967 Crossing over in the chromosomal region determining amylase isozyme in *Drosophila melanogaster*. Hereditas **58**: 1–12.

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

BREWER, G. J., 1970 *An Introduction to Isozyme Technique*. Academic Press, New York, 186.

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

COURTRIGHT, J. B., 1967 Polygenic control of aldehyde oxidase in Drosophila. Genetics **57**: 25–39.

DELORENZA, R. J. and F. H. RUDDLE, 1970 Glutamate oxlate transaminase (GOT) genetics in *Mus musculus*. Linkage, polymorphism and phenotypes of GOT-2 and GOT-1 loci. Biochem. Genet. **4**: 259–274.

FOX, D. J., ERIKA ABAECHERLI and H. URSPRUNG, 1971 *Drosophila* enzyme-genetics; A table. Experimentia **27**: 218–220.

FRANKLIN, I. and R. C. LEWONTIN, 1970 Is the gene the unit of selection? Genetics **65**: 707–734.

GILLESPIE, J. H. and K. KOJIMA, 1968 The degree of polymorphism in enzymes involved in energy production compared to that in non-specific enzymes in two *Drosophila ananassae* populations. Proc. Natl. Acad. Sci. U.S. **61**: 582–585.

GRELL, E. H., K. B. JACOBSON and J. B. MURPHY, 1968 Linkage disequilibrium in finite populations. Theoret. and Appl. Genetics **38**: 226–231.

JOHNSON, F. M., 1966 An isozyme analysis of Drosophila speciation and development. Ph.D. dissertation. University of Texas.

KOJIMA, K., J. GILLESPIE and Y. N. TOBARI, 1970 A profile of *Drosophila* species' enzymes assayed by electrophoresis. I. Number of alleles, heterozygoties, and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. Biochem. Genet. **4**: 627–637.

LANGLEY, C. H. and J. F. CROW, 1974 The direction of linkage disequilibrium. Genetics (this issue).

LEWONTIN, R. C., 1964 The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49**: 49–67.

MCCUNE, T. B., 1969 A genetic study of the amylase isozyme polymorphism in *Drosophila melanogaster*. Ph.D. dissertation. University of Hawaii.

MUKAI, T., L. E. METTLER and S. I. CHIGUSA, 1971 Linkage disequilibrium in a local population of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S. **68**: 1056–1069.

MUKAI, T., T. K. WATANABE and O. YAMAGUCHI, 1974. The genetic structure of natural populations of *Drosophila melanogaster*. XII. Linkage disequilibrium in a larger local population. Genetics (in press).

NAIR, P. S. and D. BRNCIC, 1971 Allelic variations within identical chromosome inversions. Am. Naturalist **105**: 291–294.

POULIK, M. D., 1957 Starch gel electrophoresis in a discontinuous system of buffer. Nature **180**: 1477.

OHTA, T. and M. KIMURA, 1969 Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutations. Genetics **68**: 229–238.

PRAKASH, S. and R. C. LEWONTIN, 1968 A molecular approach to the study of genetic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements in Drosophila. Proc. Nat. Acad. Sci. U.S. **59**: 398–405. ——, 1971 A molecular approach to the study of genic heterozygosity in natural populations. V. Further evidence of coadaptation in inversions of Drosophila. Genetic **69**: 405–408.

SHAW, C. R. and A. L. KOEN, 1968 Starch gel electrophoresis of enzymes. In: Chromatographic and Electrophoretic Techniques. Edited by I. SMITH. Interscience, John Wiley, New York.

SINNOCK, P. and C. F. SING, 1972 Analysis of multilocus gene systems in Tecumseh, Michigan. II. Consideration of the correlation between nonalleles in gametes. Am. J. Human Genet. **24:** 393–415.

STONE, W. S., M. R. WHEELER, F. M. JOHNSON and K. KOJIMA, 1968 Genetic variation in natural island populations of members of the *Drosophila nasuta* and *Drosophila ananassae* subgroups. Proc. Natl. Acad. Sci. U.S. **59:** 101–109.

WATANABE, T., 1967 Chromosomal polymorphism in natural populations of *Drosophila melanogaster*. Mem. Fac. Sci. Kyushu Univ. Series E (Biology) **4:** 159–178.

YANG, H. Y. and K. KOJIMA, 1972 Chromosomal polymorphism and lethal alleles in a Southwest Texas population of *Drosophila melanogaster*. Studies in Genetics VII Univ. Texas Publ. **7213:** 229–236.

YOSHIKAWA, I. and T. MUKAI, 1970 Heterozygous effects on viability of spontaneous lethal genes in *Drosophila melanogaster*. Japan J. Genet. **45:** 443–455.

YOUNG, W. J., J. E. PORTER and B. CHILDS, 1964 Glucose-phosphate dehydrogenase in *Drosophila*: X-linked electrophoretic variants. Science **143:** 140.

ZOUROS, E. and C. B. KRIMBAS, 1973 Evidence of linkage disequilibrium maintained by selection in two natural populations of *Drosophila subobscura*. Genetics **73:** 659–674.

Corresponding editor: R. C. LEWONTIN