# LINKAGE DISEQUILIBRIUM IN NATURAL POPULATIONS OF *DROSOPHILA MELANOGASTER.* SEASONAL VARIATION

## CHARLES H. LANGLEY **AND KAZUKO** ITOl

*National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, North Carolina 27709* 

**AND** 

## ROBERT **A.** VOELKER2

## *Department of Genetics, North Carolina State University, Raleigh, North Carolina 27607*

Manuscript received June **20, 1976**  Revised copy received December **21, 1976** 

#### **ABSTRACT**

Linkage disequilibrium among ten polymorphic allozyme loci and polymorphic inversions on chromosomes *2* and *3* in a natural population **of**  *Drosophila melanogaster* was examined early and late in the annual season. Similar to previous studies, little linkage disequilibrium was observed among allozymes. The two significant cases that were observed in the first sample behaved in a contradictory way. One declined much more rapidly than expected due simply to recombination; the other declined slowly as expected. There was little change in allozyme or inversion frequencies during the season.

LEWONTIN ( **1974)** has attempted to dichotomize the conceptualizations of genetic structure of natural populations of mendelizing organisms. On one hand is the view that most genetic variation in fitness is due to low frequency alleles maintained by selection-mutation balance. Alternatively, the genetic variation in fitness is viewed as arising from the segregation of polymorphic alleles at many loci that are maintained by some form of balancing selection. These loci are thought to interact strongly in their effect on fitness. Thus, one view holds that most polymorphic variation may be isoallelic with respect to fitness. The **other** holds that polymorphic variation may well be selected and, furthermore, should show nonrandom association among alleles at linked loci. An apparently clearcut difference between these two views is the expected nonallelic association or linkage disequilibrium: either strong linkage disequilibrium between linked loci or only that expected from random genetic drift.

Polymorphic allozymes are obvious "candidates." Do they show large nonrandom association when closely linked? Several survey studies of *D. melanogaster* populations have attempted to answer this question **(KOJIMA, GILLESPIE**  and **TOBARI 1970; MUKAI, METTLER** and **CHIGUSA 1971; CHARLESWORTH** and

**Genetics 86: 44.7-454 June, 1977.** 

<sup>&</sup>lt;sup>1</sup> Present address: Zoological Laboratory, Faculty of Science, Hiroshima University, Hiroshima 730, Japan.

<sup>&</sup>lt;sup>2</sup> Present address: National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triaugle Park, **North Carolma 27709** 

**CHARLESWORTH** 1973; **MUKAI, W.4TANARE** and **YAMAGECHI** 19744; **LANGLEY, TO-BARI** and **KOJIMA** 1974; **MUKAI** and **VOELKER** 1977; **VOELKER, MUKAI** and **JOHNSON** 1977). The answer is clearly ''no." Occasional instances occur, but the magnitude and frequency suggests rather weak interaction at most. Another result of some of these studies is that allozymes are often associated with polymorphic inversions in the same arm.

In this report we examine several new pairs of loci for linkage disequilibria and compare the gametic structure early and late in the season (three to four generations apart).

#### **MATERIALS AND METHODS**

*Preparation* **of** *materials:* Adult male flies were collected in summer (early July) and in autumn (the middle of September) of *1975* near **a** rural store in Carpenter, Wake County, North Carolina. They were individually mated to five females from a *SMI/Pm* ; *TM2/Sb* stock *[ln(2LR)SMI/ln(ZLR)Pm* ; *In(3LR)TM2/Sb.]* A single *F,* male, genotype *SMI/+; TMZ/+,*  was backcrossed to the balancer stock. In each of the following generations several  $SMI/\dot{+}$ ; *TMZ/+* males were backcrossed. This procedure maintains both second and third chromosomes without recombination. *186* and **202** second and third chromosome complements were extracted from the summer and autumn collections, respectively.

*Allozyme assayss* The following six enzyme loci in the second chromosome and five loci in the third chromosome were studied: a-glycerophosphate dehydrogenase *(a-Gpdh,* map position *2-20.5),* malate dehydrogenase *(Mdh,* map position *2-35.3),* alcohol dehydrogenase *(Adh* map position 2-50.1), dipeptidase-A *(Dip-A, 2-55.2, unpublished results)*, hexokinase-C *(Hex-C*, map position *2-74.5),* a-amylase *(Amy,* map position *2-77.7),* esterase-6 *(Est-6,* map position *j-36.8),* phosphoglucomutase *(Pgm,* map position *3-43.4),* octanol dehydrogenase *(Odh,* map position *3-49.2)* , acid phosphatase *(Acph,* map position *3-101.4),* leucine aminopeptidase-D *(Lap-D,* map position *3-98.3).* For the first 10 loci. *SMl/+; TM2/+ adult* flies were assayed. For *Lap-D,* eight pupae were sampled at random from each chromosome line and assayed individually. *(TM2/+* larvae cannot be distinguished from larvae of the other genotype, *TM2/Sb*. The probability that one of those eight larvae is *TM2/+* is more than *0.995).* The procedure for the electrophoresis of allozymes was that used previously (LANGLEY, TOBARI and KOJIMA 1974) for the eight loci other than *Acph* and *Dip-A.* Assay for *Acph* was according to MAC-INTYRE *(1966),* and *Dip-A* according to a modified technique from LEWIS and TRUSLOVE *(1969).* 

*Cytological analysis:* For the chromosomal analysis, *SMI/+* ; *TMZ/+* males derived from each line were crossed with  $cn$  *bw*;  $ri$   $e$  (cinnabar brown; radius incompletus, ebony) with the standard gene arrangement. The third instar larvae, genotype  $cn$  bw/ $+$ ;  $ri$  e/ $+$  (wild-type color in Malphighian tubules and spiracle sheaths) were collected, and the giant salivary gland chromosomes were observed. Cytological nomenclature followed that of LINDSLEY and **GRELL**  (1968).

## **RESULTS**

*Allozymic variation:* The allozymes in this study were chosen **to** be (1 ) polymorphic, (2) on chromosomes 2 or 3, (3) assayable with available equipment and techniques, and (4) simply interpretable from a genetic standpoint. Aldehyde oxidase (previously assayed) and xanthine oxidase were not assayed because of possible complex inheritance **(DICKISSON** and **SULLIVAN** 1975) and difficulty in identifications of alleles. **A** discussion of the various loci and their alleles can be found in a previous report **(LANGLEY, TOBARI,** and **KOJIMA** 1974), with the exceptions of acid phosphatase  $(Acph)$  and dipeptidase  $(Dip-A)$ . Acph *is* described by **KOJIMA, GILLESPIE** and **TOBARI** (1970). *Dip-A* appears to be a

new locus on the second chromosome with three polymorphic alleles. An additional, faster migrating allele of *a-Gpdh* **(3)** was found in low frequency.

The allozyme frequencies reported here (see Table 1) are similar to those reported previously for this area of North Carolina (MUKAI and VOELKER 1977). No significant heterogeneity was detected in the allozyme frequencies o€ the two samples.

*Variation in chromosomal arrangements:* The polymorphic inversion frequencies are similar to those previously reported for this area (see Table 1 ) . The one

|--|

*Allozyme gene arrangement frequencies from the July*  $(S;N = 186)$  *and September (A;N* = *202) collections from Carpenter, North Carolina* 



*1-* f (i) is the frequency of the ith allozyme, ordered according to its relative migration toward the cathode, *i.e.*, f(1) is the slowest-migrating allele, . . . f(5) is the fastest-migration toward<br>the cathode, *i.e.*, f(1) is the slowest-migrating allele, . . . f(5) is the fastest-migrating allele.<br> ${}_*$   ${}^+f(1) = Amy$ *Amyl.* 

exception to this is  $In(2R)$ . The frequencies reported here are considerably lower than those reported by **MUKAI** and his colleagues. This inversion is also the only one showing significant heterogeneity between the two samples.

*Linkage disequilibria between allozymes:* Table *2* is the allozyme-allozyme linkage disequilibrium data from the two samples. Rarer alleles were pooled because of the small sample sizes. The possible effects of this pooling will be discussed below. The first four columns are the gamete numbers, followed by the  $2 \times 2 \chi^2$ , the linkage disequilibrium *(D)*, and the correlation coefficient *(R)*.

<b>TABLE</b>	
--------------	--

*Associations between linked allozymes in the* S *and A collections* 



		$g_1$ <sup>+</sup>	$g_{2}$	$g_{3}$	$\mathcal{E}_4$	$\chi^2$	D	$\boldsymbol{R}$
$Pgm-Odh$	S	140	12	31	1	0.92	$-0.007$	$-0.071$
$(2)-(2)$	A	161	13	23	1	0.35	$-0.004$	$-0.042$
$Pgm-Lap-D$								
$(2)-(2)$	A	120	54	16	8	0.05	$-0.002$	$-0.016$
$Pgm-Acph$								
$(2)-(2)$	A	166	8	24	$\mathbf{0}$	1.15	$-0.005$	$-0.076$
$Odh$ -Lap-D								
$(2)-(2)$	A	126	58	10 ÷	$\overline{4}$	0.05	$-0.002$	$-0.016$
$Odh$ -Acph								
$(2)-(2)$	A	177	7	13	1	0.37	$+0.002$	$+0.043$
$Lap-D-Acph$								
$(2)-(2)$	A	131	5	59	3	0.15	$+0.003$	$+0.027$

TABLE 2-Continued

*t* g, is the frequency of the gametic type formed from the alleles indicated in the parenthesis under the respective loci.

\$All tests **of** significance are for one degree of freedom \* Significant at P < 0.05 level.

There is significant association between *Adh* and *Dip-A* in the summer sample, but it is not found in the autumn sample. This association is not due to common interaction with the polymorphic inversions since the association is still present when only standard gene arrangements are considered. The association between *Est-&* and *Pgm* is significant in both samples. Again this association does not appear to be due to interaction with the polymorphic inversions (both loci show association with the inversions) since the standard sequences alone show the *Est-&* by *Pgm* linkage disequilibrium.

There are two pairs of loci that are closely linked that have not been examined previously for linkage disequilibrium in natural populations. *Amy* by *Hex-C*  and *Lap-D* by *Acph.* Although these pairs are tightly linked ( $\sim$  3.2 cM and 3.1 CM respectively), no linkage disequilibrium was detected.

*Linkage disequilibrium between allozymes and inversions:* In order to examine linkage disequilibria involving inversion polymorphisms, unique inversions were pooled with those of the similar arrangement without unique inversions *(e.g., unique*  $(2L)$ *St with*  $(2L)$ *St and unique*  $In(2L)$  *with*  $In(2L)$ *. What follows* was also done without pooling and also with exclusion of the gametes carrying the unique inversions. The results reported for the pooled data are all consistent with those obtained by the above-mentioned alternative procedures. Three gametes carrying rare but cosmopolitan *3R* inversions were excluded from the pooled analysis (see Table 3).

*Adh* and *In2L* showed an association similar to that previously reported. Significant linkage disequilibria were also observed between  $\textit{Est-6}$  and  $\textit{In}(3L)$ . This association is similar in the two samples. It is also similar in sign to that reported previously for Brownsville, Texas **(LANGLEY,** TOBARI, and **KOJIMA** 1974) and for a Raleigh, North Carolina. derived cage population. All these previous results were, however, not statistically significant. Although *Pgm* and  $In(3L)$  are not



 $\overline{\phantom{a}}$ 

		$\pmb{g}_1$	$\mathbf{g}_2$	$\mathbf{g}_3$	$\mathfrak{g}_4$	$\chi^2$	$\boldsymbol{D}$	$\cal R$
$\alpha$ -Gpdh-2L	$\mathbf S$	147	13	24	$\boldsymbol{0}$	2.10	$-0.009$	$-0.107$
$(2)-(S)$	A	155	8	33	$\mathbf 2$	0.04	$+0.001$	$+0.014$
$\alpha$ -Gpdh-2R	$\mathbf S$	151	9	23	$\mathbf{1}$	0.09	$-0.002$	$-0.022$
$(2)-(S)$	A	145	18	30	5	0.30	$+0.005$	$+0.039$
$Mdh-2L$	S	165	13	6	$\boldsymbol{0}$	0.47	$-0.002$	$-0.051$
$(1) - (S)$	A	183	10	5	$\bf{0}$	0.27	$-0.001$	$-0.037$
$Mdh-2R$	$\mathbf S$	168	10	6	$\bf{0}$	0.36	$-0.002$	$-0.044$
$(1) - (S)$	A	170	23	5	$\bf{0}$	0.67	$-0.003$	$-0.058$
$Adh\text{-}2L$	S	116	11	55	$\mathbf 2$	1.59	$-0.011$	$-0.093$
$(1) - (S)$	A	137	10	51	$\boldsymbol{0}$	3.65	$-0.013$	$-0.136$
$Adh$ -2 $R$	$\mathbf S$	119	8	55	$\overline{2}$	0.60	$-0.006$	$-0.057$
$(1)-(S)$	A	129	18	46	$\overline{5}$	$0.22\,$	$-0.005$	$-0.033$
$Dip-A-2L$	$\mathbf S$	158	13	13	$\mathbf 0$	1.06	$-0.005$	$-0.076$
$(2)-(S)$	A	172	9	16	$\mathbf{1}$	0.03	$+0.001$	$+0.012$
$Dip-A-2R$	$\mathbf S$	161	10	13	$\bf{0}$	0.80	$-0.004$	$-0.066$
$(2)-(S)$	A	159	22	16	$\mathbf{1}$	0.60	$-0.005$	$-0.055$
$Hex-C-2L$	$\mathbf S$	151	13	20	$\boldsymbol{0}$	1.71	$-0.008$	$-0.096$
$(1) - (S)$	A	176	9	12	$\mathbf{1}$	0.20	$+0.002$	$+0.032$
$Hex-C-2R$	S	154	10	20	$\bf{0}$	1.29	$-0.006$	$-0.084$
$(1) - (S)$	A	163	22	12	$\mathbf{1}$	0.21	$-0.003$	$-0.032$
$Amy-2L$	S	163	12	8	$\mathbf{1}$	0.24	$+0.002$	$+0.036$
$(4) - (S)$	A	175	$\boldsymbol{9}$	13	$\mathbf{1}$	0.14	$+0.001$	$+0.026$
$Amr$ -2 $R$	$\mathbf S$	165	10	9	$\boldsymbol{0}$	0.54	$-0.003$	$-0.054$
$(4) - (S)$	A	161	23	14	$\bf{0}$	1.98	$-0.008$	$-0.100$
$Est-6-3L$	$\mathbf S$	110	$\bf{0}$	69	$\overline{5}$	$7.64***$	$+0.016$	$+0.204$
$(2) - (S)$	A	121	$\overline{2}$	68	$\overline{7}$	$6.38*$	$+0.018$	$+0.179$
$Est-6-3R$	$\mathbf S$	94	16	60	14	0.62	$+0.011$	$+0.058$
$(2)-(S)$	A	101	22	64	11	0.35	$-0.008$	$-0.042$
$Pgm-3L$	$\mathbf S$	147	5	32	$\boldsymbol{0}$	1.08	$-0.005$	$-0.077$
$(2)-(S)$	A	165	9	24	$\bf{0}$	1.30	$-0.006$	$-0.081$
$P$ gm-3 $R$	$\mathbf S$	127	25	$\sqrt{27}$	5	0.01	$-0.001$	$-0.008$
$(2)-(S)$	$\mathbf A$	143	31	22	$\overline{2}$	1.37	$-0.010$	$-0.083$
$Odh-3L$	$\mathbf S$	166	$\overline{5}$	13	$\bf{0}$	0.39	$-0.002$	$-0.046$
$(2)-(S)$	A	175	9	14	$\bf{0}$	0.72	$-0.003$	$-0.060$
$Odh-3R$	S	144	27	10	3	0.47	$+0.005$	$+0.051$
$(2)-(S)$	A	154	30	11	3	0.25	$+0.003$	$+0.035$
$Lap-D-3L$								
$(2)-(S)$	A	130	6	59	3	0.02	$+0.001$	$+0.010$
$Lap-D-3R$								
$(2)-(S)$	A	119	17	46	16	$5.43*$	$+0.029$	$+0.166$
$Acph-3L$								
$(2) - (S)$	A	181	9	8	$\mathbf{0}$	0.40	$-0.002$	$-0.045$
$Acph-3R$								
$(2)-(S)$	A	157	33	8	$\mathbf 0$	1.67	$-0.007$	$-0.092$
$2L$ -2 $R$	$\mathbf S$	161	10	13	$\mathbf 0$	0.80	$-0.004$	$-0.066$
$(S) - (S)$	A	166	22	9	$\mathbf{1}$	0.03	$-0.001$	$-0.012$
3L-3R	$\mathbf S$	150	29	4	$\mathbf{1}$	0.05	$+0.001$	$+0.017$
$(S) - (S)$	A	162	27	3	6	16.97***	$+0.023$	$+0.293$

TABLE 3

*Associations between gene arrangements and linked allozymes in S and A collections* 

\* Significant at the  $P < 0.05$  levels.<br>\*\* Significant at the  $P < 0.01$  level.<br>\*\*\* Significant at the  $P < 0.001$  level.

associated to a statistically significant degree, they do deviate in the same direction as that previously reported for both Japan and Texas (LANGLEY, TOBARI and KOJIMA 1974). *Lap-D* and  $In(3R)$  are associated in the autumn sample. The direction oi the disequilibrium is opposite from that reported for Texas.

*Linkage disequilibrium between inversions:* Of the four comparisons between linked inversions in this report, one was statistically significant. Although the expected number of  $In(3L)$ - $In(3R)$  gametes in the autumn sample is small  $(1.5)$  and perhaps inflates the  $x^2$ , the statistical significance of the result is surely high.

*Nonrandom associations between unlinked elements:* One of the 96 unlinked comparisons was statistically significant *(Dip-A by Est-6;*  $x^2 = 7.84$ *,*  $df = 1$ *).* One test of this statistical significance in 96 is expected. Similar to previous results (LANGLEY, TOBARI and KOJIMA 1974) , nc significant association was observed between unlinked elements.

*Nonrandoin associations among three loci:* All linked triplets *of* allozyme loci and chromosome arrangements were examined by  $2 \times 2 \times 2 \times 2$ . Of the 157 such comparisons, many were significant. After the marginal  $2 \times 2 x^2$  values were subtracted, only one remained significant *(Est-6* by  $3\overline{L}$  by  $3R$ ;  $x^2 = 5.1$ ,  $df = 1$ ). Considering the number of tests, this result is most likely trivial.

*Effect of pooling:* In no instance was there a significant  $x^2$  *(i.e., P < 0.05)* when rarer alleles were not pooled that was accompanied by an insignificant *x2* after pooling.

## DISCUSSION

The results presented above reinforce three previously reported observations about the genetic structure of *Drosophila melanogaster* populations in nature. **(1** ) Gene frequencies are temporarily stable. Although some seasonal variability may exist and yearly changes have been observed, the temporal variation in a particular area is considerably less than the total geographic variation in the species.  $(2)$  The amount of linkage disequilibrium among allozymes is, on the average. small, although occasional significant cases are obesrved. *(3)* Allozymes are often associated with polymorphic inversion in the same chromosome arm.

The uniqueness of this particular study is the sampling of the same population twice in the same year. Since little variation is seen in the two samples, perhaps we should conclude that the population is large and stable. The two cases of allozyme-allozyme linkage disequilibria do not clarify the situation. The *Est-6-Pgm*  case is just what we expect in a large, stable population: closely linked genes out of equilibrium (for whatever reason) stay that way in the short run *(3* generations). The *Adh-Dip-A* case is just the opposite; although they are similarly closely linked, the linkage disqulilibrium disappears in this short period. The expected decays in disequilibria due to recombination are as follows: for *Adh-* $Dip-A$ ,  $(1 - 0.025)^3$   $0.022 = 0.020$   $(-0.002$  observed) and for *Est-6-Pgm*,  $(1 (0.033)^3$   $(0.032 = -0.029$  (-0.026 observed). There seem to be two alternative interpretations available. The first is that the *Est-6-Pgm* case is a legitimate case of linkage disequilibrium (whatever its origin) and **the** *Adh-Dip-A* case is a sampling error. The probability of a  $x^2 > 6.11$  in the 41 remaining (ex-

c. H. LANGLET, K. 110 AND R. A. VOELKER<br>cluding *Est-6-Pgm*) linked allozyme-allozyme comparisons is  $1 - (.9852)^{41}$ cluding *Est-6-Pgm*) linked allozyme-allozyme comparisons is  $1 - (.9852)^{41}$ <br>~ 0.4. Or in the face of this we can entertain the possibility that the gamete frequencies changed in response perhaps to selection. The most reasonable view for the present is that most allozymes are in equilibrium with one another and that the causes and dynamics of those few that are out of equilibrium is unclear.

How do these and previous results bear on the question of the genetic structure of gametes in mendelizing populations? One answer might be that this picture is consistent with allozymic variation being selectively neutral or isoallelic, and thus supports the "neoclassical" vision. An alternative and more prudent answer recognizes that allozymes are collections of alleles with similar electrophoretic phenotypes and not individual and unique alleles. If we consider each allozyme to represent three alleles and we examine the calculation of *D,* we find that

$$
D = \Sigma_i \Sigma_j \Sigma_k \Sigma_l \left[ f(A_i B_j) f(A_k B_l) - f(A_i B_j) f(A_k B_j) \right].
$$

This is the sum of 81 individual linkage disequilibria,  $D_{ijkl}$  (64 independent  $D_{ijk}$ . Since there is no particular reason to think that electrophoretically similar alleles might have similar epistatic interactions. we might expect *D* to average to zero as the number of unseen alleles increases. In conclusion, we recognize that electrophoretic techniques may not be sufficiently precise to resolve the dichotomy posed by LEWONTIN (1974).

### LITERATURE CITED

- CHARLESWORTH, **B.** and D. CHARLESWORTH, 1973 A study of linkage disequilibrium in populations of *Drosophila melanogaster.* Genetics **73** : 351-359.
- DICKINSON, W. J. and D. T. SULLIVAN, 1975 *Gene-enzyme systems in Drosophila.* Springer-Verlag, New York.
- KOJIMA, K., J. GILLESPIE and Y. N. TOBARI, 1970 A profile of Drosophila species enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities, and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. Biochem. Genet. **4** : 627-637.
- LANGLEY, C.H., Y. N. TOBARI and K. KOJIMA, 1974 Linkage disequilibrium in natural populations of *Drosophila melanogaster.* Gene tics **78:** 921-936.
- LEWIS, W. H. P. and G. M. TRUSLOVE, 1969 Electrophoretic heterogeneity of mouse erythrocyte peptidase. Biochem. Genet. **3:** 493-498.
- LEWONTIN, **R.** C., 1974 *The genetic basis* of *evolutionary change.* Columbia University Press, New York.
- LINDSLEY, D. L. and E. H. GRELL, 1968 Genetic variations of *Drosophila melanogaster.* Carnegie Inst. Washington Publ. 627.
- MACINTYRE, R.J., 1966 The genetics of an acid phosphatase in *Drosophila melanogaster* and *Drosophila simulans.* Genetics **53:** 461-474.
- 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: 1065-1069.
- MUKAI, T., T. K. WATANABE and O. YAMAGUCHI, 1974 The genetic structure of natural populations of *Drosophila melanogaster.* XII. Linkage disequilibrium in a large local population. Genetics **77:** 771-793.
- MUKAI, T. and R. A. VOELKER, 1977 The genetic structure of natural populations of *Drosophila melanogaster.* XIII. Further studies on linkage disequilibrium. Genetics *86:* 175-185.
- VOELKER, R. A., T. MUKAI and F. M. JOHNSON, 1977 Genetic variation in populations of *Drosophila melanogaster* from the western United States. Genetica (in press).

Corresponding editor: R. C. LEWONTIN