GENETICS OF NATURAL POPULATIONS. XVII. PROOF OF OPERATION OF NATURAL SELECTION IN WILD POPULATIONS OF DROSOPHILA PSEUDOOBSCURA

THEODOSIUS DOBZHANSKY AND HOWARD LEVENE*
Columbia University, New York

Received June 6, 1948

INTRODUCTION

Most natural populations of *Drosophila pseudoobscura* are mixtures of individuals homozygous and heterozygous for gene arrangements in the third chromosome. The relative frequencies of the gene arrangements in populations show geographical and temporal variations. On Mount San Jacinto and in the Sierra Nevada in California, and in some localities in Texas, the temporal variations are cyclic and connected with the succession of the seasons (DOBZHANSKY 1943, 1948, DOBZHANSKY and EPLING 1944). It has been inferred that the seasonal changes in the frequencies of the gene arrangements are caused by natural selection. Carriers of different chromosomal types possess different adaptive values, and different types are favored at different seasons. This hypothesis has been indirectly confirmed by experiments on artificial populations which contained mixtures of different gene arrangements. Under certain conditions, changes in the relative frequencies of gene arrangements are observed from generation to generation in such artificial populations (WRIGHT and DOBZHANSKY 1946, and unpublished data). Analysis of these changes shows that, with rare exceptions, individuals of which the third chromosomes differ in gene arrangement (inversion heterozygotes) are superior in fitness to inversion homozygotes. An independent proof of the adaptive superiority of the heterozygotes to the homozygotes is as follows: Among the eggs deposited in the cages containing the artificial populations, inversion heterozygotes and homozygotes occur with relative frequencies approaching that demanded by the HARDY-WEINBERG rule. Yet, among the adult flies developed in the same population cages, excesses of heterozygotes and deficiencies of homozygotes are observed (DOBZHANSKY 1947a). It follows, then, that carriers of different gene arrangements interbreed at random in the population cages, and that a differential mortality between the egg and the adult stage upsets the HARDY-WEINBERG equilibrium proportions.

As a working hypothesis, it may be assumed that in natural populations likewise, the fitness of inversion heterozygotes is higher than that of homozygotes, and that the observed temporal changes in the natural populations are connected with this difference in fitness. If this hypothesis is valid, the homozygotes will be less and the heterozygotes will be more frequent among the flies found in nature than would be expected on the basis of the HARDY-WEINBERG equilibrium. The present article describes an attempt to test this working hypothesis.

* Observational data by TH. DOBZHANSKY, mathematical analysis by H. LEVENE.

METHOD

It should be noted that departures from the HARDY-WEINBERG equilibrium may result from causes other than selection, and that selection does not necessarily produce deviations from this equilibrium. Thus, excesses of homozygotes may be produced by inbreeding (e.g., mating of sibs shortly after emergence from pupae, or preferential mating of relatives), assortative mating, or random differentiation into sub-populations within the territory from which the samples are collected. The last contingency is excluded, because the mobility of adult *Drosophila pseudoobscura* is known to be considerable in relation to the size of collecting territories (DOBZHANSKY and WRIGHT 1943). Causes leading to deficiencies of homozygotes do not seem to be frequent in nature, though preferential mating of carriers of different genotypes might produce such deficiencies. All of the above causes of deviations from equilibrium proportions may be subsumed under the category of departures from panmixia; the deviations so produced will be discoverable at all stages of the life cycle.

Selection operating through differential mortality at any stage of the life cycle will cause departure from the HARDY-WEINBERG equilibrium at that stage and at all later stages, but not at earlier stages. On the other hand, selection operating through differential fecundity or sexual activity of adult flies would cause no disturbance of this equilibrium at any stage of the life cycle, provided that the population is panmictic. It follows that the effects of departures from panmixia and of differential mortality in the egg stage can be distinguished fro n those of differential mortality at any stage after egg hatching. For this purpose, samples of adult flies are collected in natural habitats and analyzed in two ways. First, females which had been fertilized before capture are permitted to lay eggs in individual cultures. The larvae are raised under optimal conditions at which little selection occurs (see WRIGHT and DOBZHANSKY 1946), and the chromosomes of one full grown larva from each culture are examined. These samples of larvae can be considered random samples of the eggs that would have been laid in nature by the females, and they can be used to test for panmixia. Such data will be called egg samples. Second, the males, caught in natural habitats at the same time, are individually mated to females of known chromosome constitution, and chromosomes of the resulting larvae are examined. Seven larvae (before 1945) and six larvae (during and since 1945) are examined. The first larva examined discloses one of the paternal chromosomes. The remaining larvae have a probability 63/64 or 31/32 of disclosing the other paternal chromosome. The wild adult males are thus classified as to karyotype, except that on the average 1/64 or 1/32 of the heterozygotes present are incorrectly classified as homozygotes. This type of data will be called adult male samples and can be used to test for selection, if egg samples have shown panmixia to hold. For more details about the methods used see DOBZHANSKY and EPLING (1944).

PANMIXIA IN NATURAL POPULATIONS

Extensive data on the incidence of homo- and heterozygotes in natural populations have been published by DOBZHANSKY and EPLING (1944). These

data were collected, whenever possible, by examining a single larva in the offspring of each wild female, and represent accordingly what has been called above "egg samples." However, some "adult male samples" have also been included. On the basis of such data DOBZHANSKY and EPLING concluded that panmixia obtained in natural populations of *Drosophila pseudoobscura*. This evidence needs reexamination for two reasons. First, many of the samples reported were combined egg and adult male samples, although egg samples decidedly predominated because they required much less labor in the laboratory. Second, χ^2 was calculated by an incorrect method. The correct method can be exemplified for a hypothetical sample having only ST and CH chromosomes. Let a, b, c be the number of ST/ST, ST/CH, CH/CH individuals (egg or adult males) observed. Then $p=(2a+b)/2(a+b+c)$ and $q=(2c+b)/2(a+b+c)$ are the proportions of ST and CH chromosomes observed. The "expected" numbers under the HARDY-WEINBERG equilibrium are then approximately $Ea=np^2$, $Eb=2npq$, $Ec=nq^2$, where $n=a+b+c$. We then compute

$$\chi^2 = \frac{(a - Ea)^2}{Ea} + \frac{(b - Eb)^2}{Eb} + \frac{(c - Ec)^2}{Ec}.$$

Then, if the expected values are not too small, χ^2 will have the Chi-square distribution, but not—as might appear at first glance—with two degrees of freedom. The number of degrees of freedom is the number of entries that can be made arbitrarily without changing the expected values. Now in the present case, if *a* be specified, then we must have $2a+b=2np$, which fixes b for constant n and p, and when b is fixed the relation $2c+b=2np$ fixes c. Hence if we consider all tables with the same expected values, and hence with the same n, p, q, we find that only one entry is arbitrary, and we have only one degree of freedom for χ^2. Now the method used by DOBZHANSKY and EPLING was to calculate

$$Y^2 = \frac{[(a + c) - (Ea + Ec)]^2}{Ea + Ec} + \frac{(b - Eb)^2}{Eb},$$

i.e. a "χ^2" based on only two classes, homozygotes and heterozygotes. It is easy to show that Y^2 is always less than or equal to our previous χ^2 when computed from the same sample. Hence if χ^2 is distributed as Chi-square with one degree of freedom, Y^2 cannot be so distributed, and the use of the Chi-square table with Y^2 leads to P values which are too high. In general, if there are r different gene arrangements in a sample, there will be $r(r+1)/2$ possible genotypes, and χ^2 when computed in the obvious manner will have $[r(r+1)/2]-r$, or $r(r-1)/2$ degrees of freedom. Again, in this general case, a "χ^2" based only on number of homozygotes and number of heterozygotes will not have a Chi-square distribution. The general χ^2 with $r(r-1)/2$ degrees of freedom is unsatisfactory, first because the expected numbers of some rarer genotypes are usually so small as to make the Chi-square table inapplicable, and second because significant values of χ^2 can be caused by discrepancies of many kinds other than excess or deficiency of homozygotes. A general method

for testing directly whether homozygotes are in excess or deficiency will be given below.

Before this method had been devised, one of us (LEVENE, unpublished) had tested data from fifty samples collected at Mount San Jacinto and Death Valley in California (data in DOBZHANSKY and EPLING 1944, table 9) by the following method.

The chromosomes appearing in a sample were classified into two types: A, the most common chromosome in the sample (usually Standard in these data) and B, all other chromosomes. Larvae were classified as A/A, A/B, and B/B; and χ^2 was computed, based on these three classes. Fifteen of these fifty samples have now been omitted because they included appreciable numbers of adult males. The seven samples from Wildrose A, B, C, and D have been combined into two groups, Wildrose 1939 and Wildrose 1940. Although DOBZHANSKY and EPLING (1944) have recorded apparently significant differences between the samples from the A, B, C and D stations, these stations are sufficiently close relative to the migration radius of a fly (DOBZHANSKY and WRIGHT 1943) to be treated as parts of a single locality. We then have 30 samples which are essentially egg samples and give the following results. There are 5 samples giving χ^2 values from 0 to .0642, 8 from .0643 to .2749, 4 from .2750 to .7083, 7 from .7084 to 1.642, and 6 greater than 1.642. The expected number of χ^2 values in each of these intervals is 6. A test of goodness of fit on this distribution gives $\chi^2 = 1.667$ with 4 degrees of freedom, or $P = .79$. A more sensitive test is obtained by adding all 30 χ^2 values. The sum is 29.367 and should have a Chi-square distribution with 30 degrees of freedom, giving $P = .50$. Of the 30 samples, 13 have a deficiency of homozygotes and 17 have an excess, which again conforms to expectation.

This evidence shows clearly that there was no important departure from panmixia in these samples. It should be noted that there is a drawback to the statistical method here used. In fact, A/A larvae are homozygotes, and A/B are heterozygotes; but many of the "homozygous" B/B larvae are actually inversion heterozygotes. However, in view of the clearcut nature of the results, it was not felt necessary to undertake the considerable labor of recomputing this material by the new method given below.

STATISTICAL TREATMENT OF ADULT MALE SAMPLES

Since we are now assured that there is a good approach to panmixia, it is of interest to examine adult male samples for evidence of selection. Since egg samples based on impregnated females require much less labor in the laboratory than the analysis of adult males, most of the earlier samples were exclusively egg samples, or contained, at best, relatively few adult males. Since 1945 several larger samples of adult males have been taken for the express purpose of investigating selection. Since it was expected that natural selection would lead to a deficiency of homozygotes, it became desirable to find a method that was specifically designed to test this hypothesis and one that could make use of data from small samples. This method is as follows:

Let h_r be the observed number of homozygotes in the r-th sample, let H_r be

the expected value of h_r and let σ_r be the standard error of h_r. Then if we sum the values of h_r from a number of independent samples of size n_r, the total number of homozygotes $h = \Sigma h_r$ will have expected value ΣH_r and standard error $\sqrt{\Sigma \sigma_r^2}$. Furthermore, h will be nearly normally distributed if h and $\Sigma n_r - h$ are both large, even if the individual sample sizes n_r are small. We can thus consider

$$t = \frac{\Sigma h_r - \Sigma H_r}{\sqrt{\Sigma \sigma_r^2}} \tag{1}$$

as a normal deviate, and reject the hypothesis that the HARDY-WEINBERG equilibrium holds if $P \leqq .05$ or some other conventional level of significance, where

$$P = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

can be obtained from a table of the normal curve (*e.g.* for $t = -1.64485$, $P = .05$). Note that only negative t are significant since we are only interested in the alternative hypothesis of too few homozygotes. For a large sample $t_r = (h_r - H_r)/\sigma_r$ will itself be normal and can be used to test that particular sample. For simplicity we now drop the subscript r and consider a single unique sample. Suppose there are k distinct inversions present. Let q_i be the *population* frequency of the i-th arrangement, and let y_i be the number of such chromosomes in our sample. Then $\Sigma q_i = 1$ and $\Sigma y_i = 2n$. Also let x_{ii} be the number of homozygous i/i individuals and x_{ij} the number of i/j heterozygotes in the sample. Then if the q_i were known, we would have $E(x_{ii}) = nq_i^2$, $E(x_{ij}) = 2nq_iq_j$, and $E(h) = H = \Sigma nq_i^2$, where $h = \Sigma x_{ii}$ is the number of homozygotes observed and the symbol E stands for "expected value of." Actually, however, the q_i are not known. Hence we must consider the conditional zygotic distribution in samples that have the same number of chromosomes of each type as the observed sample. It can be shown that under these conditions, the exact expected value of x_{ii} is given by

$$E(x_{ii}) = \frac{y_i^2 - y_i}{4n - 2},$$

and hence

$$E(h) = H = \Sigma \frac{y_i^2 - y_i}{4n - 2} = \frac{1}{4n - 2}\{C - 2n\}, \tag{2}$$

where

$$C = \Sigma y_i^2.$$

For a single moderate sized sample we could use the approximate formula $E(x_{ii}) = y_i^2/4n$ which corresponds to that when q_i is known, but this would lead to serious bias in combining data from a large number of small samples. Note that when $y_i = 1$ the approximate formula gives $E(x_{ii}) = 1/4n$, while the

exact formula gives $E(x_{ii}) = 0$ correctly, since a homozygous i/i individual requires two i chromosomes. It can also be shown that to a good degree of approximation we have

$$\sigma_h^2 = \frac{1}{4n^2}\left\{C(n+2) + C^2\left(\frac{2n+5}{8n^2}\right) - D\frac{n+2}{n}\right\} - \frac{1}{2} \tag{3}$$

where

$$C = \Sigma(y_i^2) \quad \text{and} \quad D = \Sigma(y_i^3).$$

Proofs of the relations (2) and (3) will be published elsewhere (LEVENE, 1949). It should be noted that relation (3) is not exact for small samples, but the error is of the order of magnitude of 1/n and, unlike a small error in H, will not become important when many small samples are summed.

We have still to correct for the small error due to failure to observe both chromosomes of some male heterozygotes (see above). Consider ST/CH heterozygotes. In the long run, for every 64 ST/CH flies observed (in data since 1945) we will classify 62 as ST/CH, 1 as ST/ST and 1 as CH/CH. Accordingly we must correct our data by subtracting one ST/ST and one CH/CH and adding two ST/CH flies for every 62 ST/CH actually recorded. In the present case we have h homozygotes and (n−h) heterozygotes recorded, and we have merely to use as our corrected figures

$$\begin{aligned}(n - h^*) &= (n - h) + e,\\ h^* &= h - e,\end{aligned} \tag{4}$$

where $e = \frac{1}{63}(n-h)$ in data before 1945, and $e = \frac{1}{31}(n-h)$ in data collected during and after 1945. As a result of the correction h* will occasionally be negative. If we were dealing with a single sample, h* would then be taken as zero, but in combining a large number of samples bias will result unless the negative values are retained. The random error of classification will not affect the expected value H, but it will increase the variance of h*. However, it is easy to show that this increase will be approximately equal to e. For our purposes it will then be sufficient to use the formula

$$\sigma^2 = \sigma_{h^*}^2 = \sigma_h^2 + e. \tag{5}$$

Ordinarily this correction will not change σ^2 very much.

As an example of the calculations required, consider the sample from Piñon Flats, June 1946. There were 7 ST/ST, 5 AR/AR, 19 CH/CH, 1 TL/TL, 20 ST/AR, 22 ST/CH, 25 AR/CH, 3 ST/TL, 5 AR/TL, and 4 CH/TL, giving n = 111 flies, of which 32 were homozygotes. The calculations are given in table 1. There is a deficiency of homozygotes. The probability of obtaining fewer homozygotes for a sample of 111 with the observed numbers of the given chromosomes, is P = .185, and so this sample shows no significant deviation from the equilibrium proportions.

TABLE 1

Computations for Piñon Flats, June 1946. ST=Standard, AR=Arrowhead, CH=Chiricahua, TL=Tree Line. y_i=no. of i chromosomes observed. $Ex_{ii}=(y_i^2-y_i)/(4n-2)$=no. of i/i homozygotes expected.

CHROMOSOME ARRANGEMENT	i	y_i	y_i^2	y_i^3	$y_i^2-y_i$	Ex_{ii}
ST	1	59	3481	205379	3422	7.7421
AR	2	60	3600	216000	3540	8.0090
CH	3	89	7921	704969	7832	17.7195
TL	4	14	196	2744	182	0.4118
Total		222 =2n	15198 =C	1129092 =D	14976	33.8824 =H

$$n = 111, \quad h = 32, \quad n - h = 79, \quad 4n - 2 = 442,$$

$$e = (n - h)/31 = 79/31 = 2.5484, \quad h^* = h - e = 29.4516,$$

$$H = (C - 2n)/(4n - 2) = (15198 - 222)/442 = 33.8824 \text{ check},$$

$$\sigma_h^2 = \frac{1}{4n^2}\left\{C(n+2) + C^2\frac{2n+5}{8n^2} - D\frac{n+2}{n}\right\} - \frac{1}{2}$$

$$= \frac{1}{49284}\left\{15198 \times 113 + \frac{230979204 \times 227}{98568} - \frac{1129092 \times 113}{111}\right\} - \frac{1}{2}$$

$$= 22.3171 - \tfrac{1}{2} = 21.8171,$$

$$\sigma^2 = \sigma_h^2 + e = 21.8171 + 2.5484 = 24.3655, \quad \sigma = 4.9361$$

$$t = \frac{h^* - H}{\sigma} = \frac{29.4516 - 33.8824}{4.9361} = -0.8976,$$

$$P = \int_{-\infty}^{-.8976} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0.1847.$$

DEPARTURE FROM EQUILIBRIUM IN ADULT MALE SAMPLES

Table 2 summarizes the data for 66 adult male samples that have been analyzed. The size of the samples ranged from 5 to 279 individuals, with only five samples having more than 100 individuals each. In table 2, the locality and date of the samples are given in the two columns on the left, followed by size of the sample (number of males analyzed), number of homozygous individuals found, corrected number of homozygotes, expected number of homozygotes, and the variance (σ^2).

The analysis can be started by observing that in 56 samples the observed numbers of homozygotes are less than the expected ones, while in only 10 samples is the relationship reversed. If no differential mortality of inversion homo- and heterozygotes occurs in natural populations between the egg stage and the stage at which the adult males are captured, positive and negative deviations should be about equally frequent. This suggests a real deficiency of homozygotes in the natural populations sampled. Among the ten largest samples (with more than 50 flies each), all show fewer homozygotes than ex-

TABLE 2

Individual adult male samples. n = no. of males analyzed. h = no. of homozygotes. h = corrected no. of homozygotes. H = no. of homozygotes expected under* HARDY-WEINBERG *equilibrium. σ^2 = variance of h*.*

LOCALITY	DATE	n	h	h*	H	σ^2
Andreas	May 39	30	17	16.794	13.169	4.769
"	June 39	17	5	4.809	5.121	3.359
"	Apr. 41	23	10	9.794	11.822	2.718
"	Apr. 42	22	13	12.857	12.163	2.227
"	Apr. 20/40	24	12	11.819	12.276	3.083
"	Oct. 39	16	3	2.794	5.516	2.708
"	Nov. 39	34	11	10.635	14.343	6.109
"	Dec. 39	22	8	7.778	10.070	3.412
"	Jan. 40	18	4	3.778	6.514	3.469
"	Feb. 40	22	8	7.778	10.930	4.328
"	Mar. 40	15	8	7.889	8.276	1.339
"	Mar. 41	22	10	9.809	12.442	2.722
"	Mar. 42	25	12	11.794	13.551	3.664
"	Oct. 40	20	10	9.841	9.846	3.990
"	Mar. 28/40	16	4	3.809	5.129	3.690
"	Sept. 6/41	29	13	12.746	14.333	3.933
"	Feb. 19/41	23	10	9.794	10.822	3.262
"	Feb. 1/41	12	5	4.889	6.304	1.417
Piñon	May 39	15	4	3.825	4.897	3.225
"	Apr. 40	40	8	7.492	11.911	8.602
"	June 40	19	7	6.809	5.838	4.131
"	Apr. 45	63	18	16.548	20.016	14.745
"	June 46	111	32	29.452	33.882	24.366
"	Apr. 47	111	38	35.645	37.195	23.709
"	Sept. 39	7	0	−0.111	2.154	1.621
"	Oct. 39	15	3	2.809	5.621	3.078
"	Mar. 40	9	0	−0.143	3.059	1.816
"	Dec. 41	20	9	8.825	7.718	3.740
"	Mar. 46	279	106	100.419	108.546	49.415
"	1940	14	4	3.841	4.148	3.002
Keen	May 39	8	2	1.805	2.600	1.822
"	June 39	38	7	6.508	12.173	8.651
"	July 39	23	3	2.682	7.444	5.382
"	Sept.-Oct. 39	15	1	0.778	4.758	3.227
"	July 40	15	2	1.794	5.241	3.212
"	Aug. 40	36	8	7.556	12.394	8.042
"	Sept. 40	22	5	4.730	7.232	2.491
"	Apr. 40	35	8	7.571	12.203	7.771
"	Apr. 45	50	16	14.903	14.141	10.609
"	Apr. 46	134	46	43.161	48.285	25.634
Mather	May 47	15	3	2.613	4.690	3.708
"	Sept. 47	114	21	18.000	30.318	23.409
Aspen	Sept. 47	9	2	1.774	2.176	1.760
Lost Claim	Aug. 46	31	8	7.258	9.836	6.745
" "	May 47	10	2	1.742	3.421	2.266
" "	Sept. 47	5	0	−0.161	1.111	1.036
Avawatz Mt.	May 39	14	4	3.841	6.407	2.180

TABLE 2—(*continued*)

LOCALITY	DATE	n	h	h*	H	σ^2
Coso Mt.	July 37	36	21	20.762	21.394	3.860
Charleston Mt.	June 39	24	15	14.857	13.468	2.631
Cottonwood Mt.	June 37	68	23	22.286	27.363	15.143
Grapevine Mt.	June 37	29	9	8.682	12.132	6.055
Kingston Mt.	May 37	40	22	21.714	22.633	4.929
Lida Nev.	June 37	29	19	18.841	18.702	2.088
Panamint Mt.	May 37	46	21	20.603	24.253	6.108
Providence Mt.	May 37	54	41	40.794	42.065	1.632
Sheep Range	June 37	58	44	43.778	45.470	2.477
Wildrose	Sept. 38	20	8	7.809	5.974	3.954
"	June 39	89	27	26.016	32.254	19.617
"	1940	25	9	8.746	8.428	5.247
Jacksonville	June 46	16	4	3.613	5.387	3.237
"	July 46	15	5	4.677	5.897	2.193
"	Sept. 47	22	9	8.581	7.535	3.915
Raton	1940	32	21	20.825	22.587	2.365
Santa Cruz	1940	21	6	5.762	6.634	3.926
Santa Lucia	1940	16	4	3.809	5.903	2.598
Sonoita	1941	12	3	2.857	4.565	2.692

pected. Nevertheless, in only one sample (taken at Mather, California, in September 1947) is the deviation significant by the usual criteria ($P = .0054$); for the other nine the P values range from .375 to .051.

The critical evidence comes from adding the values of deviations and variances from different samples as described above. We then find for the ten largest samples the deviation $h^* - H = -50.30$ and $\sigma = 14.15$ corresponding to a P value of .0002, while for all 66 samples $h^* - H = -139.30$, $\sigma = 20.35$ and P is less than 10^{-6}. The conclusion is justified that inversion homozygotes are significantly less frequent, and heterozygotes more frequent, in natural populations than they would be if no differential mortality occurred between the egg stage and the stage of capture.

Since the present data result from combining samples from widely scattered times and places, with considerable differences in numbers and relative proportions of the different chromosome types, there is no satisfactory over-all measure of the deviation from the HARDY-WEINBERG equilibrium. However, for a rough indication it may be noted that for the combined data the ratio $(\Sigma h^* - \Sigma H)/\Sigma H = -.147$, or there is a 14.7% deficiency of homozygotes. The lack of a good measure of the extent of the deviation does not interfere with the combination of the data for purposes of the test of significance.

Table 3 represents an unsuccessful attempt to group the adult male samples in categories which could be compared with the known facts about seasonal changes in the frequencies of gene arrangements. Thus, seasonal changes have been recorded at Andreas Canyon and at Piñon Flats, but not at Keen Camp on Mount San Jacinto. Changes occur in spring and summer but not in winter (DOBZHANSKY 1943). Seasonal changes have been recorded in the Sierra Ne-

vada, and changes from year to year are indicated for some localities in the Death Valley region (DOBZHANSKY and EPLING 1944, DOBZHANSKY 1948). The population of Keen underwent considerable change between 1940 and 1945 (DOBZHANSKY 1947b). P values which indicate significant deficiencies of homozygotes appear in table 3 for early samples at Keen Camp, for winter samples in Andreas Canyon, for the Sierra Nevada, for Death Valley, winter samples at Piñon Flats, and miscellaneous samples. Summer samples from

TABLE 3

Results of adult male samples grouped together (see text). n=no. of males analyzed. h=no. of homozygotes. h=corrected no. of homozygotes. H=no. of homozygotes expected under* HARDY-WEINBERG *equilibrium. σ^2=variance of h*. $t=(h^*-H)/\sigma$. P=probability of as many or fewer homozygotes if* HARDY-WEINBERG *equilibrium holds (see text).*

DESCRIPTION	n	h	h*	H	σ^2	t	P
Andreas Summer	116	57	56.073	54.551	16.156	+0.3789	0.6476
Andreas Winter	274	106	103.334	128.076	44.043	−3.7279	0.0001
Piñon Summer	359	107	99.771	113.739	78.778	−1.5737	0.0578
Piñon Winter	344	122	115.640	131.246	62.672	−1.9711	0.0243
Keen to 1940	192	36	33.424	64.045	40.598	−4.8056	$<10^{-5}$
Keen 1945–6	184	62	58.064	62.426	36.243	−0.7246	0.2344
Sierra Nevada	184	36	31.226	51.552	38.924	−3.2579	0.0006
Death Valley	532	263	258.729	281.543	75.921	−2.6184	0.0044
Miscellaneous	134	52	50.124	58.508	20.926	−1.8330	0.0334
Total	2319	841	806.385	945.686	414.261	−6.8442	$<10^{-6}$

Piñon Flats gave a P value of about 0.058, while no significant deviation is indicated from the recent Keen samples, and the summer samples from Andreas Canyon contain even an insignificant excess of homozygotes. No conclusion can be drawn about the relationships between the observed deficiencies of homozygotes and the known seasonal changes in the incidence of gene arrangements. It should be noted that while the total sample size is adequate for our purposes, the number of flies in the separate groups is too small to permit satisfactory conclusions to be drawn about them.

DISCUSSION

The purpose of the present study has been to test the hypothesis that the frequencies in natural populations of gene arrangements in the third chromosome of *Drosophila pseudoobscura* are controlled by natural selection. As stated in the Introduction, this hypothesis was originally advanced on the basis of observations on seasonal changes in the frequencies of the chromosomal types in certain localities (DOBZHANSKY 1943). This hypothesis was first confirmed by experiments which showed that some of the changes observed in nature can be reproduced in artificial populations in population cages (WRIGHT and DOBZHANSKY 1946). Next, it was shown that, in population cages, there is a differential mortality between the egg and the adult stage which favors inver-

sion heterozygotes and discriminates against homozygotes (DOBZHANSKY 1947a). The data reported in the present article complete the proof of the hypothesis by showing that for the eggs deposited in natural populations the proportions of homo- and heterozygotes are as demanded by the HARDY-WEINBERG rule, while among the adult males found in nature heterozygotes are more, and homozygotes are less, frequent than this rule requires. If follows that, at least among the male zygotes, there is differential mortality which favors inversion heterozygotes. This fact alone, without any information on seasonal or other temporal changes, would be sufficient to show that the frequencies of the gene arrangements in nature are subject to selective pressure.

SUMMARY

In wild populations of *Drosophila pseudoobscura*, the proportions of homozygotes and heterozygotes for different types of third chromosomes are, among the eggs deposited by the adult flies, in conformity with the binomial square rule (the HARDY-WEINBERG formula). Yet, among the adult male flies found in nature these proportions depart from the binomial square rule, because the homozygotes are less, and heterozygotes more, frequent than demanded by the rule. Thus, a differential mortality occurs between the egg and the adult stage which favors the heterozygotes. The chromosomal variation is controlled by natural selection.

LITERATURE CITED

DOBZHANSKY, TH., 1943 Genetics of natural populations. IX. Temporal changes in the composition of populations of *Drosophila pseudoobscura*. Genetics **28**: 162–186.

1947a Genetics of natural populations. XIV. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. Genetics **32**: 142–160.

1947b A directional change in the genetic constitution of a natural population of *Drosophila pseudoobscura*. Heredity **1**: 53–64.

1948. Genetics of natural populations. XVI. Altitudinal and seasonal changes produced by natural selection in certain populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genetics **33**: 158–178.

DOBZHANSKY, TH., and C. EPLING, 1944. Contributions to the genetics, taxonomy and ecology of *Drosophila pseudoobscura* and its relatives. Carnegie Inst. Washington Publ. **554**: 1–183.

DOBZHANSKY, TH., and S. WRIGHT, 1943 Genetics of natural populations. X. Dispersion rates in *Drosophila pseudoobscura*. Genetics **28**: 304–340.

LEVENE, H., 1949 On a matching problem arising in Genetics. Ann. Math. Statistics **20** (in press).

WRIGHT, S., and TH. DOBZHANSKY, 1946 Genetics of natural populations. XII. Experimental reproduction of some of the changes caused by natural selection in certain populations of *Drosophila pseudoobscura*. Genetics **31**: 125–156.