GENETICS OF NATURAL POPULATIONS. V. RELATIONS BETWEEN MUTATION RATE AND ACCUMULATION OF LETHALS IN POPULATIONS OF DROSOPHILA PSEUDOOBSCURA

TH. DOBZHANSKY AND SEWALL WRIGHT[1]
Columbia University, New York, and University of Chicago, Chicago

Received August 12, 1940

INTRODUCTION

NATURAL populations of *Drosophila pseudoobscura* have been shown to carry numerous recessive mutant genes concealed in heterozygous condition (STURTEVANT 1937, DOBZHANSKY and QUEAL 1938b, DOBZHANSKY 1939). Some of these genes are lethal or semilethal, others produce visible external effects, and still others modify the viability, duration of the development, or other characters. Although the lethals would seem to be least likely among the above genetic variants to play a constructive role in evolution, they are favorable material for population studies. Their detection involves no personal equation and is technically simpler than that of the visibles or of modifiers of physiological processes. Moreover, a lethal, provided it is completely recessive, is subject to natural selection only in homozygous condition, when its adaptive value is zero; experimental determination of the adaptive values of other genes is a most elusive task. The present article reports the results of a study of the identity of lethals found in natural populations as shown by their allelism; of the rate of elimination of lethals by selection; and of the rate of their *de novo* origin by mutation. The mathematical analysis is restricted to the implications from the data on the occurrence of lethals. The consideration of other sources of information regarding the population structure in *Drosophila pseudoobscura*, such as the variability of the chromosome structure, is reserved for a future publication.

The observational and experimental work on which the present paper is based was carried through at the Division of Biology of the California Institute of Technology, Pasadena, with the support of a grant from the Carnegie Corporation of New York, made through the Carnegie Institution of Washington, whose help is gratefully acknowledged. MISS M. L. QUEAL, MR. B. A. SPASSKY, and MRS. N. P. SIVERTZEV-DOBZHANSKY have shared the technical work involved. DRS. MORGAN WARD, K. MATHER, and H. J. MULLER have contributed advice concerning alternative methods of calculation of the minimum number of mutable loci.

[1] Observational and experimental data by Th. Dobzhansky, mathematical analysis by Sewall Wright.

MATERIAL AND METHOD

The technique of the detection and the general properties of lethals and semilethals in *Drosophila pseudoobscura* were described by DOBZHANSKY and QUEAL (1938b). For the present study 123 lethal-bearing third chromosomes isolated from eleven populations inhabiting as many mountain ranges in the Death Valley region of California and Nevada were used. A map showing the positions of ten of these ranges has been published (DOBZHANSKY and QUEAL 1938a); the eleventh one, Argus Range, lies southwest of Coso Mountains and west of Panamint Valley as shown on this map. The numbers of the third chromosomes analyzed and of the lethals detected in them are shown in table 1. The numbers of the lethals that were preserved and used in the present study appear in parentheses.

TABLE 1

LOCALITY	CHROMOSOMES ANALYZED	LETHALS	LOCALITY	CHROMOSOMES ANALYZED	LETHALS
Lida	55	14 (14)	Awavaz Mts.	23	2 (2)
Cottonwood Mts.	93	17 (16)	Kingston Range	101	12 (10)
Grapevine Mts.	56	7 (7)	Charleston Mts.	103	15 (12)
Coso Range	124	22 (21)	Sheep Range	90	11 (11)
Argus Range	8	4 (4)	Providence Mts.	99	15 (15)
Panamint Mts.	105	12 (11)			
			Total	857	131 (123)

In order to convey a clear idea about the extent of the territory whose population is characterized by the above samples, it will be convenient to distinguish between the "regions" such as Death Valley, Mexico, and Guatemala, "ranges" consisting of more or less continuous mountain forests isolated from each other by deserts, "localities" consisting of the areas whose populations were actually sampled (10 to 20 traps distributed over a distance from one-quarter to one mile in each case), and "sublocalities" or "stations," still smaller areas distinguished in certain studies. The eleven localities treated here are all in separate mountain ranges, but each locality covers only a small fraction of the range it represents. The eleven ranges can be inclosed in a minimum rectangle about 210 miles long and 120 miles wide, called here the Death Valley region.

The lethal-bearing third chromosomes are perpetuated in the form of balanced strains; in most of these one of the third chromosomes contains a lethal, and the other the mutant genes orange, Blade, Scute, and purple. Since *Bl* is lethal when homozygous, only two of the four classes of zygotes survive, and the strain breeds true. The *or Bl Sc pr* chromosome has the Standard gene arrangement, wherefore this system of balancing is satisfactory only for those lethal-bearing chromosomes which have gene ar-

rangements other than Standard. Another balancer, carrying the gene Emarginate in the Arrowhead arrangement, was used for keeping the lethal-bearing Standard chromosomes.

FREQUENCY OF ALLELIC LETHALS WITHIN A POPULATION

In the populations from the Death Valley region about 15 percent of the third chromosomes contain lethals or semilethals (table 1). This figure conveys no information regarding the frequency of individuals that die on account of homozygosis for lethals. It is evident that if only one kind of lethal were present and random mating is assumed, the frequency of inviable zygotes would be the square of that or the lethal-bearing chromosomes. On the other hand, if each lethal is present in a population only once, no lethal homozygotes are produced. The following experiments were designed to determine the actual situation.

Fourteen lethals were recovered from the Lida sample (table 1). Each of these lethal-bearing strains was outcrossed to the remaining thirteen strains. A similar series of crosses was made for the other localities. Since most of the lethals are balanced over *or Bl Sc pr*, these crosses may be represented thus: lethal *A/or Bl Sc pr* ♀ × lethal *B/or Bl Sc pr* ♂. If lethals *A* and *B* are alleles, the offspring of the cross consists of Blade-Scute individuals only; otherwise Blade-Scute and wild type individuals appear in a ratio approaching 2:1. Since the Blade and non-Blade flies are easily distinguishable through the glass of the culture bottle, the presence or absence of wild types can be established by inspection with a hand lens without etherizing the flies. This is important in an experiment involving thousands of crosses. In crosses in which one or both lethals were balanced over Emarginate (see above), as well as where the number of wild type flies seemed small, the flies were examined under a microscope, and counts were made.

The results of the intercrosses are shown in table 2. The number of kinds of lethals found is, in most localities, smaller than the number of lethals analyzed. This amounts to saying that some of the lethals proved to be alleles. Thus, in the Lida sample one of the lethals was encountered four times, two were found twice each, and six were represented once each. An estimate of the frequency of lethal homozygotes that must be formed in nature may be arrived at in the following manner. The average frequency of the lethal-bearing third chromosomes in the populations from the Death Valley region is 15.29 ± 0.83 percent (probable errors are used throughout this paper). The frequency of zygotes receiving two lethals, one from the father and one from the mother, is the square of 15.29, or 2.34 percent. Most of such zygotes carry non-allelic lethals and are viable. The data included in table 2 show the results of 772 intercrosses in which

both parents carried lethals recovered from the same population. Among these, 24 crosses, or 3.11 ± 0.42 percent, gave the results indicating the presence of allelic lethals (that is, absence of wild type offspring). Hence, about 3.11 percent of the zygotes with two lethals carry allelic ones. Among all zygotes produced in natural populations of the Death Valley region, 0.073 percent ($0.0234 \times 0.0311 \times 100$) may be expected to die on account of homozygosis for third chromosome lethals.

TABLE 2

Intra-locality tests. Frequency of lethals found once, twice, and four times in the same population.

LOCALITY	LETHALS ANALYZED (k)	NUMBER OF TESTS $\left\{\frac{k^2-k}{2}\right\}$	SINGLES	PAIRS	QUADRUPLETS	TOTAL KINDS OF LETHALS
Lida	14	91	6	2	1	9
Cottonwood	16	120	10	3	—	13
Grapevine	7	21	5	1	—	6
Coso	21	207*	13	2	1	16
Argus	4	6	4	—	—	4
Panamint	11	55	9	1	—	10
Awavaz	2	1	—	1	—	1
Kingston	10	45	10	—	—	10
Charleston	12	66	12	—	—	12
Sheep Range	11	55	9	1	—	10
Providence	15	105	13	1	—	14
Total	123	772	91	12	2	105

* Three of the possible 210 crosses were not made.

FREQUENCY OF ALLELIC LETHALS IN DIFFERENT POPULATIONS

The equilibrium frequency of a recessive autosomal lethal—that is, the frequency to which it will be allowed to accumulate in a population—is a function of its mutation rate as well as of the effective size of the population in which it occurs. In very large populations the equilibrium level is the square root of the mutation rate producing the lethal; in smaller populations this level is lower than in large ones (WRIGHT 1937). The data presented above indicate that certain lethals are more common than others in some localities This fact deserves careful scrutiny.

Apparently the simplest explanation of some lethals being encountered more frequently than others is that the former have higher mutation rates than the latter. To determine the mutation rates for individual lethals, however, is impracticable. The hypothesis must be tested indirectly. If the observed frequencies of lethals reflect solely, or even principally, their mutation rates, a lethal found frequently in one population may be ex-

pected to be common in other populations as well. Therefore, the probability of a zygote carrying two lethals having two allelic ones should be independent of whether the lethals are recovered from the same or from different populations. The allelism of lethals found in different populations must be studied. The experimental procedure is here the same as that used to determine the allelism of lethals from a single population. Each lethal-containing strain must be outcrossed to strains carrying all other lethals. The necessary series of crosses was carried out, except that whenever a lethal was known to occur more than once within a locality (table 2), only one lethal-bearing chromosome was tested. Since the technique of the detection of lethals in wild chromosomes did not discriminate between chromosomes containing a single lethal and those containing more than one lethal (DOBZHANSKY and QUEAL 1938b), a source of error was incurred. If some of the chromosomes harbor two or more (of course, non-allelic) lethals, chromosomes A and B and A and C may show allelic lethals, while B and C show none. Although no case of this sort was encountered among the intra-locality crosses where almost all possible tests were made, the possibility of error is not eliminated. Since, however, a considerable labor saving was thus effected, and since the error incurred is small (*cf.* p. 37), this course was deemed justified. The results obtained are summarized in table 3.

TABLE 3

Inter-locality tests.

LETHALS TESTED	FOUND ONCE	FOUND TWICE	FOUND THRICE	KINDS OF LETHALS
105	68	17	1	86

Table 3 shows that some of the lethals found in different localities proved to be allelic. An analysis of the data on which tables 2 and 3 are based reveals some further facts of interest. Among the 68 lethals that had no alleles in the inter-locality crosses (table 3), 60 had no alleles in the intra-locality ones either; to put it in a different way, these lethals were found once in the whole material examined. Of the remaining eight lethals of this group, six were found twice within a locality, and two were detected four times each (in Lida and Coso, respectively, table 2). Among the 17 pairs of lethals observed in the interlocality crosses (table 3), 12 were not observed more than once in any one locality, and five were each found twice in one and once in a different locality (that is, three times in the whole material). Finally, the triplet of the inter-locality crosses is a lethal that was found once in Panamint, once in Sheep Range, and twice in Lida. In other words, the lethals found four times each in Lida and in Coso (table 2)

were not recovered from any other populations, and among the 12 lethals each of which occurred twice within a population only five were found once more elsewhere. Among the 91 lethals which had no alleles in the intra-locality tests, 28 were recovered from samples from two localities each. It follows that the lethals found repeatedly within a population are not significantly commoner in the species at large than are the apparently rare lethals while *vice versa* the lethals recovered from samples from two or three localities show no tendency to accumulate in any particular sample.

That some lethals attain higher frequencies within some populations than they have in the species at large may be demonstrated by still another method. A total of 4913 crosses were arranged in which each parent carried a lethal recovered from a different population. Among these, only 20 crosses, or 0.407 ± 0.061 percent, showed allelic lethals (table 3). The comparable figure for the intercrosses in which the parents carried lethals from the same population is 3.11 ± 0.42 percent. The difference is significant statistically. This fact may be restated as follows: A certain amount of heterosis would accrue to the species if the matings were random instead of being (as they largely are) confined to individuals inhabiting the same locality.

These observations militate against the assumption that the distribution of lethals in the populations of *Drosophila pseudoobscura* is governed exclusively, or even largely, by their mutation rates. It will be shown below that the facts at hand are more simply accounted for on the assumption that the population of this species inhabiting the Death Valley region is segregated into partly isolated colonies, each with a limited effective breeding size.

METHODS OF DETERMINING THE MUTATION RATE

To approach an understanding of the dynamics of the lethal contents of natural populations, information regarding the rate of their origin by mutation is indispensable. In principle, the experimental detection of the lethals arisen by mutation is as simple as that of the lethals present outdoors. A male, or males, from the strain in which it is desired to determine the spontaneous mutation rate in the third chromosome is crossed to females homozygous for the recessives orange and purple. The F_1 males are outcrossed singly to females carrying *or Bl Sc pr* in one and the "Cuernavaca" inversion in the other of their third chromosomes. In the F_2, females and males showing *Bl* and *Sc*, but not *or* and *pr*, are selected and inbred. In the absence of lethals, the F_3 offspring consist of *Bl-Sc* and wild type individuals in a ratio approaching 2:1; if a lethal is present only *Bl-Sc* and no wild types appear; semilethals reduce the proportion of the wild types to somewhere between 16 and zero percent of the total output of flies.

The classification of the F_3 cultures is made by inspecting the culture bottles with the aid of a hand lens; normal cultures and those containing lethals are thus simply identified. If wild type flies are present but constitute obviously less than one-third of the total, a semilethal may be involved. Cultures of this kind are subjected to counting, and if the doubt is not removed, a further generation is raised. A recessive semilethal is arbitrarily defined as a gene which permits survival of less than one-half of the homozygotes (16 percent of all flies in the culture under the conditions of the experiment). This arbitrariness may seem to cast doubt on the reliability of the data. Fortunately, among the mutants found in nature as well as among the newly arisen ones, complete lethals and extreme semilethals are much more frequent than semilethals of medium strength. This, and also the fact that most semilethals produce visible external effects or slow down the development of the homozygotes, insures that few, if any, of the semilethals were missed and that the error thus incurred is not serious.

The appearance of a lethal in an F_3 culture may indicate either that a mutation has arisen in the spermatogenesis of the P_1 generation male or that that male was heterozygous for a lethal inherited from close or remote ancestors. In the former case a single or a few, and in the latter case one-half, of the cultures derived from a given P_1 generation male show lethals. To guard against any ambiguity arising from this source, at least 40 F_1 males were taken from each P_1 generation culture, and as many F_3 cultures were raised from them. Series in which one-half of the cultures contained lethals were discarded, and not in a single instance has doubt arisen regarding the classification of a series of crosses. In two series two allelic lethals have arisen in each, among more than 50 cultures. Here the mutation has evidently taken place at a stage of spermatogenesis prior to the formation of spermatids.

The technique just described will be referred to as the "direct" method. Its laboriousness is obvious. A variation of the same technique, an "accumulation" method, was used in a part of the experiments. Males from wild strains known to be free of lethals are crossed to *or pr* females. The F_1 males are back-crossed to *or pr* females. In the next generation wild type and *or pr* flies appear in a ration approaching 1:1; *several* wild type males are selected from each culture and back-crossed again to *or pr* females; such back-crosses are repeated in a series of consecutive generations. Since there is normally no crossing over in Drosophila males a marked wild chromosome is transmitted intact through a series of generations, being exposed in each generation to the risk of a lethal mutation taking place in it. The experiment can be completed at any stage, by outcrossing the males heterozygous for *or pr* to *or Bl Sc pr*/Cuernavaca females and inbreeding the resulting *Bl-Sc* progeny. Of course, only a single male from each line

is so tested. The presence or absence of lethals is established exactly as with the direct method. Since the *or Bl Sc pr* chromosome has the Standard gene arrangement, all the experiments are made with chromosomes having gene arrangements other than the Standard.

Aside from the difference in the amount of the labor involved, the direct and the accumulation methods differ in other respects, which make it desirable to use both methods for the determination of the mutation rates. Among the newly arisen lethals some may be completely recessive, while others may produce a decrease in the viability of the heterozygotes. The absence or rarity of incompletely recessive lethals in natural populations

TABLE 4

Mutation rates in certain strains from the Death Valley region. The first and second figures in each column indicate the numbers of the tested chromosome-generations and of the lethals obtained, respectively.

	EXPERIMENTS					
STRAIN	1ST	2ND	3RD	4TH	5TH	TOTAL
Lida—2	41—0	57—0	—	103—0	1014— 5	1215— 5
Lida—4	47—0	—	—	—	—	47— 0
Lida—9	49—0	—	—	—	—	49— 0
Whitney—30	44—0	—	—	—	—	44— 0
Whitney—43	48—0	—	—	—	—	48— 0
Coso—6	47—1	72—0	62—0	161—0	—	342— 1
Coso—86	45—0	62—0	—	92—0	—	199— 0
Coso—100	48—1	136—1	73—0	208—2	1030— 1	1495— 5
Panamint—24	41—0	26—0	—	85—0	1533— 2	1685— 2
Awavaz—1	45—0	17—0	—	115—0	1344— 3	1521— 3
Awavaz—4	48—0	137—0	—	68—0	308— 2	561— 2
Grapevine—12	44—0	38—0	—	150—0	1099— 4	1331— 4
Grapevine—34	41—0	36—0	40—0	110—0	973— 4	1200— 4
Grapevine—36	—	55—0	—	—	861— 3	916— 3
Charleston—71	46—0	164—1	109—0	107—1	725— 2	1151— 4
Charleston—88	45—0	—	—	—	—	45— 0
Sheep Range—15	46—0	103—0	—	102—0	1005— 5	1256— 5
Sheep Range—17	53—0	59—0	46—0	209—2	—	367— 2
Total	778—2	962—2	330—0	1510—5	9892—31	13472—40

does not necessarily mean that such lethals arise infrequently. Most of the lethals encountered outdoors had arisen many generations before their detection, and therefore natural selection had had time to eliminate most of the incompletely recessive ones. The same phenomenon of "filtering" of lethals takes place in the accumulation experiments, since several wild type males are used in every generation, and a competition ensues among the prospective fathers as well as among their offspring. The direct method makes the competition less severe; although the carriers of deleterious

changes are eliminated here as well, mildly unfavorable variants may escape destruction. A comparison of the mutation rates observed in the direct and in the accumulation experiments may, then, give information on the problem of recessivity of newly arisen lethals.

According to DEMEREC (1937), RHOADES (1938), and others, mutation rates may vary from strain to strain of the same species; the general mutability as well as the mutation rates in specific genes may be affected. STURTEVANT (1939) found an increase of the mutation rates following hybridization of races A and B of *Drosophila pseudoobscura*. In the direct experiments the P_1 generation males are either homozygous for a given wild third chromosome or have a genetic structure similar to that of their wild ancestors. In the accumulation experiments the mutability is studied in wild/*or pr* heterozygotes. The *or pr* strain, obtained from PROFESSOR P. CH. KOLLER of Edinburgh, has been kept in laboratories for many generations, and its ultimate geographic origin is probably a mixed one. Although it is difficult to decide whether the direct or the accumulation experiments resemble the natural conditions more closely, a comparison of their outcomes may be of interest.

A great uniformity of environmental condition was deliberately avoided, since such hardly obtains in the natural habitats of the fly. In the first and second series of experiments (table 4), the P_1 generation flies were raised and kept in an incubator at $24\frac{1}{2}$°C, while all other experiments were conducted in a room in which the temperature fluctuated from 19° to 25°, usually lower during the night than in the daytime, and occasionally rose as high as 27° for several days.

Every lethal obtained was tested by out-crossing it to the *or Bl Sc pr*/Cuernavaca strain. The need of this precaution follows from the fact that the Cuernavaca chromosome carries a lethal extracted from a wild Mexican strain. Accidental use of non-virgin *or Bl Sc pr*/Cuernavaca females may result in the detection of a lethal introduced by such a contamination. The possibility that a *bona fide* new lethal is allelic to that contained in the Cuernavaca chromosome is disregarded to avoid a source of error which is deemed more serious. The experiments extended from May 1938 to July 1939, inclusive.

THE RESULTS OF THE EXPERIMENTS ON THE MUTATION RATE

Eighteen strains from the Death Valley region (table 4), six from Mexico, and five from Guatemala (table 5) were used. With the Death Valley strains, the experiments from the first to the fourth involve the application of the direct method, while the fifth and the Mexican and Guatemalan ones are accumulation experiments. An examination of data shows that there is no indication of any differences in the mutability among the

strains from the same geographic region. If the Death Valley, Mexican, and Guatemalan strains are contrasted as groups, the heterogeneity X^2 turns out to be 0.48, which, for two degrees of freedom, has a probability between 0.80 and 0.70. In percent, the mutation rates for the Death Valley material are 0.297 ± 0.032, for Mexico 0.359 ± 0.062, and for Guatemala 0.284 ± 0.061. All the differences are less than their probable errors.

TABLE 5

Mutation rates in certain strains from Mexico and Guatemala.

STRAIN	CHROMOSOME-GENERATIONS TESTED AND MUTATIONS OBTAINED	STRAIN	CHROMOSOME-GENERATIONS TESTED AND MUTATIONS OBTAINED
Tehuacan	591— 2	Atitlan	166— 0
Puebla	864— 3	Totonicapan	809— 2
Amecameca	820— 3	Sacapulas	820— 3
Pachuca	783— 3	Momostenango	901— 1
Cuernavaca	771— 3	Huehuetenango	826— 4
Zitacuaro	348— 1		
Total Mexico	4177—15	Total Guatemala	3522—10

The outcomes of the direct and the accumulation experiments may now be examined. In the four direct experiments nine lethals were obtained among 3580, and in the accumulation experiment 31 lethals among 9892 tested chromosome-generations (table 4). Hence, the mutation rates are 0.251 ± 0.056 and 0.313 ± 0.038 percent, respectively. The difference is again less than the probable error. It must be concluded either that in our material very few incompletely recessive lethals arise and the hybridity fails to influence materially the mutation rate or that the effects of these two factors are equal in extent and opposite in sign.

In all the experiments a total of 65 mutations occurred in 21,171 chromosome-generations. The mutation rate is 0.307 ± 0.026 percent, or about three mutations per one thousand chromosome-generations. Among the 65 mutations, 51 were complete lethals and 14 were semilethals. Among the semilethals only one permits the survival of about 30 percent of the homozygotes, the remainder being close to complete lethality. This may be compared to the data on lethals in natural populations where in a total of 163 natural lethals 126 were complete lethals and 37 were semilethals (DOBZHANSKY and QUEAL 1938b, DOBZHANSKY 1939a). The agreement is satisfactory.

CONCENTRATION OF LETHALS IN DIFFERENT POPULATIONS

In the populations from the Death Valley region 15.29 ± 0.83 percent of the third chromosomes carry a lethal or a semilethal, while in the Mexi-

can and Guatemalan populations the frequency of such chromosomes is as high as 30.00 ± 2.82 percent (DOBZHANSKY 1939a). Since the number of the lethal-producing loci in the third chromosome is at least 289 (*cf.* p. 39), the average concentration of lethals per locus is about $q = 0.00053$ for the Death Valley and 0.00104 for Mexico and Guatemala (slightly higher if the chance of occurrence of two or more lethals in a chromosome is allowed for). The level to which lethals are allowed to accumulate in different natural populations is evidently not constant.

Since the mutation rates in the strains from Death Valley on one hand and in those from Mexico and Guatemala on the other were found to be equal, it follows that the above difference in the accumulation of lethals in natural populations is not a reflection of a difference in mutability. The objection may be raised, however, that mutation rates in natural habitats of the fly may differ from those observed in the laboratory. If a thoroughly sceptical attitude is adopted, this objection is unsurmountable. The following considerations, however, are relevant. Since the concentrations of lethals in Mexico and Gautemala are twice as high as in the Death Valley region, the mutation rates must differ by a factor of four. Mutation rates are known to be subject to environmental modification, particularly by temperature. However dissimilar are the respective geographical regions, it is difficult to visualize a combination of climatic factors to which so great a difference could reasonably be attributed. The mean temperature at which the experiments were conducted is probably a few degrees higher than in the natural habitats of the flies during the breeding season, but this is far from enough to account for the situation. ZUITIN (1938a, 1938b) claims to have shown that temperature changes accentuate the mutation rates in *Drosophila melanogaster*. Assuming this to be true, mutation rates in the laboratory should be lower, not higher, than outdoors.

A SUMMARY OF THE PRIMARY DATA AND A DISCUSSION OF THEIR RELIABILITY

The available information on the distribution of naturally occurring lethals in the third chromosome of *Drosophila pseudoobscura* and the rate of their origin by mutation permit certain inferences to be drawn regarding the structure of natural populations of this species. To facilitate the reader's task in following the mathematical analysis of this information, it may be useful at this point to sum up the primary data that will be used therein and to discuss their limitations. These primary data are: rate of occurrence of lethal mutations, 0.00297 ± 0.00032; frequency of chromosomes carrying lethals, 0.1529 ± 0.0083; chance of allelism between localities, 0.00407 ± 0.00061; chance of allelism within localities, 0.0311 ± 0.0042.

The probable errors of these frequencies (X) are based on the number of

observations (K) according to the formula $PE = 0.6745\sqrt{\frac{X(1-X)}{K}}$. Some of the primary data, however, may contain flaws that require consideration. The figure for mutation rate comes from experiments in the laboratory. As each observation is wholly independent of the others, their total number can properly be used in calculating the probable error. The other data were collected in nature, and questions of sampling become important. The 857 tested third chromosomes were collected in eleven localities. The number per locality varied from eight to 124 with an average of 77.9 (table 1). The collections in each locality came from 10 to 20 traps distributed over maximum distances of one-quarter to one mile. It appears that on the average some four to eight of the chromosomes came from flies collected in the same trap, but doubtless the number varied greatly.

The flies caught in the same trap could easily have come from the same parents and are thus especially likely to carry the same lethal. This consideration does not affect the estimates of the frequency of lethal chromosomes in the region or of the chance of allelism between localities. It would, however, affect the probable errors, since the derivation of the formula for the latter postulates complete independence of the observations. Any positive correlation between observations reduces the effective value of K. In the case of the chance of allelism within localities, correlation between observations not only affects the probable error but may also have a systematic effect. In the extreme case, if most of the allelic lethals in intra-locality crosses come from flies caught in the same trap, the data would be inadequate to prove that the lethals found within a locality (apart from the immediate neighborhood of a brood) are alleles more frequently than the lethals coming from different localities. This would seriously vitiate certain parts of the following analysis of the data. This danger, however, is removed by certain newer observations made since this paper was concluded. These observations will be published in detail later; essentially they show that the frequencies of allelism of lethals collected in a given locality simultaneously or collected at different times (from a month to several months apart) do not differ to any appreciable extent.

POPULATION CONSTANTS

Before attempting any interpretation of the data, it is desirable to make a list of the statistical properties or constants of the population of a locality of which account must be taken. At least three constants must be determined for even a first approach to adequate description of the breeding structure of a population in a locality.

N The effective size of the breeding population. This may be expected to be very much smaller than the apparent number not only be-

cause it includes only parents, but also because it is much more closely related to the number at the minimum of the seasonal cycle than to that at maximum (WRIGHT 1938b, 1940).

F The inbreeding coefficient of the locality. This is zero if there is random mating within the locality but positive if for any reason there is a tendency for mating of relatives in excess of that under random mating.

The zygotic frequencies are as follows for a pair of alleles with gene frequencies $[(1-q)a+qA]$ not taking account of effects of selection, mutation, etc. (WRIGHT 1921, 1922):

Zygote	*Frequency*	
AA	$(1 - F)q^2 + Fq$	
Aa	$2(1 - F)q(1 - q)$	(1)
aa	$(1 - F)(1 - q)^2 + F(1 - q)$	

It may be noted that this array of frequencies consists of a random breeding component, $(1-F)[qA+(1-q)a]^2$, and a component like that of an array of completely fixed lines, $F[qAA+(1-q)aa]$.

m The immigration index. This measures the extent to which the population of the locality in question is replaced in each generation by immigrants, representative of the species as a whole. Since most actual immigrants are likely to trace to adjacent localities of the same range, more or less resembling the given locality in gene frequency, these are actually equivalent in effect to a much smaller number drawn at random from the species. Thus the effective value of m may be expected to be considerably less than the actual proportion of replacement by immigrants.

No doubt the constants N, F, and m differ considerably among the localities, but it is not practicable to take cognizance of this in the analysis.

Following are the most important properties of the lethal genes:

n The number of loci in the third chromosome, subject to lethal mutation.

v The rate of mutation per locus per generation.

s The selective disadvantage of heterozygotes.

The preceding six constants are independent of each other. With a given breeding structure (N, F, m), and a given rate of mutation (v), and of selection against heterozygotes (s), the genes should exhibit a certain distribution of frequencies $\phi(q)$, characterized especially by the following constants:

$\bar{q}$ The mean gene frequency.

σ_q^2 The variance of the gene frequencies due merely to chance in populations of the given breeding structure.

There may be important differences among loci in mutation rate and in the amount of selection against heterozygotes. The most important consequence of this, for our purpose, seems to be in causing variability of the mean gene frequencies of the various lethals. The actual variance of the frequencies of all loci (σ_{qt}^2) must be treated as consisting of two components, the variance ($\sigma_{\bar{q}}^2$) due to differences in v and s, and the variance of genes with the same v and s (σ_q^2) due merely to chance.

$\sigma_{\bar{q}}^2$ The variance of gene frequencies due to differences in v and s.

The inbreeding coefficient (F) of a locality should be a function of the size of population N_i and the immigration index m_i of stations in cases in which the inbreeding is due to subdivision into partially isolated groups. The proportion of heterozygosis in the locality is $2q(1-q)-2\sigma_{qi}^2$ where σ_{qi}^2 is the variance of gene frequencies due to chance, either within or between stations (WAHLUND 1928, WRIGHT 1931). This is also given by $2q(1-q)(1-F)$ not considering the effects of selection etc., as noted above. As

$$\sigma_{qi}^2 = \frac{q(1-q)}{4N_im_i+1}$$

under the same conditions (WRIGHT 1931):

$$F = \frac{\sigma_{qi}^2}{q(1-q)} = \frac{1}{4N_im_i+1}. \qquad (2)$$

It should be noted, however, that F also applies where there is a tendency toward mating of close relatives not based on subdivision into distinguishable stations.

We have listed nine important population constants and so far have indicated the possibility of only six equations (four data, two relations of dependence). Obviously, determinate solution requires at least three additional data. Nevertheless certain conclusions may be obtained without going farther.

CHROMOSOMES WITH ONE AND WITH MORE THAN ONE LETHAL

The method used for the detection of lethal-bearing third chromosomes in the natural populations did not discriminate between chromosomes having one and those having more than one lethal. Knowing that 15.29 percent of the third chromosomes carry lethals, we must estimate how many of them have one, two, or three lethals. Only independently occurring

lethals will be considered, since the frequency of deficiencies involving many loci is low among the spontaneous lethals (SLIZYNSKI 1938, TIMOFEEFF-RESSOVSKY and ZIMMER 1939). The mean frequency of lethal genes, considered collectively is $n\bar{q}$. The distribution of frequencies in chromosomes should be according to the terms of the Poisson series $e^{-n\bar{q}(1-l)}$ where l symbolizes a lethal. The frequency of chromosomes lacking all lethals is thus $(1-0.1529)=e^{-n\bar{q}}$.

$$n\bar{q} = 0.166. \tag{3}$$

The expected frequencies of chromosomes with 0, 1, 2 and, 3 lethals are as follows:

	Percent of all chromosomes	Percent of lethal chromosomes
No lethals $e^{-n\bar{q}}$	84.71	
1 lethal $n\bar{q}\ e^{-n\bar{q}}$	14.06	92.0
2 lethals $\frac{(n\bar{q})^2}{2!} e^{-n\bar{q}}$	1.17	7.6
3 lethals $\frac{(n\bar{q})^3}{3!} e^{-n\bar{q}}$	0.06	0.4
	100.00	100.0

(4)

If a chromosome has only one lethal, all chromosomes with lethals that are allelic to it may be expected to be allelic to each other. But if the given chromosome has two lethals, in approximately 50 percent of the cases, two alleles of it should not behave as such with each other (exactly if both have single lethals, approximately otherwise). Similarly, if the given chromosome has three lethals in approximately two thirds of the cases, two lethal alleles of it should not behave as such with each other (exactly if both have single lethals, approximately otherwise). Taking the estimates of the frequencies of chromosomes with one, two, and three lethals in the present material, it appears that in only about 4.06 percent of the cases would two alleles of a given lethal chromosome fail to be alleles to each other $(7.6\times\frac{1}{2}+0.4\times 2/3=4.06)$.

This theory applies to allelic lethal chromosomes picked at random from the species as a whole and thus in general of independent mutational origin. In the case of allelic lethal chromosomes from the same locality, there is an excellent chance of common recent ancestry and thus of similarity in all lethals. Thus the error involved in testing only one of the allelic lethals found within localities should be considerably less than four percent and may be considered negligible.

INTERRELATIONS AMONG CONSTANTS n, v, $\bar{q}$, σ_q^2 AND $\sigma_{\bar{q}}^2$

In order to relate the data to the theoretical constants, it is first necessary to deduce genic relations from the observed data, referring to *chromosomes*. The expected frequencies of chromosomes with one, two, and three lethals are given above. Another equation is given at once by the observed mutation rate of normal third chromosomes:

$$nv = 0.00297 \tag{5}$$

We have next to relate the frequencies of allelism within and between localities to the constants. Let p be the proportion which lethals at a given locus constitute of all lethals (that is, $\sum p = 1$). The probability that a second lethal is at the same locus as one chosen first is obviously p. The average value of this probability, considering all loci is $\sum p^2$. But $\sigma_p^2 = (\sum p^2/n) - \bar{p}^2$ and $\bar{p} = \sum p/n = 1/n$. Thus the probability of allelism is $1/n + n\sigma_p^2$. The value of p is proportional to the frequency (q) of the gene in the population in question, but as

$$\sum q = n\bar{q}, \qquad p = \frac{q}{n\bar{q}}, \qquad \sigma_p^2 = \sigma_{qt}^2/(n\bar{q})^2.$$

The data give us the probability of allelism of lethal chromosomes rather than of lethal genes. As noted above, however, it is likely that within localities most allelic lethal chromosomes are alike in all lethal genes. The figure 0.0311 thus will be taken to apply approximately to genes as well as to chromosomes (apart from the consideration of randomness of sampling within localities, taken up later).

Probability of allelism within localities is thus as follows, recalling that $\sigma_{qt}^2 = \sigma_q^2 + \sigma_{\bar{q}}^2$:

$$\frac{1}{n} + \frac{n(\sigma_q^2 + \sigma_{\bar{q}}^2)}{(n\bar{q})^2} = 0.0311. \tag{6}$$

Correction is theoretically important in the case of allelic lethal chromosomes from remote populations, since these are likely to be independent with respect to mutational origin of all contained lethal genes. As indicated above, the proportions of chromosomes with K lethals are according to the successive terms of

$$\frac{e^{-n\bar{q}}}{1 - e^{-n\bar{q}}}\left[n\bar{q} + \frac{(n\bar{q})^2}{2!} + \frac{(n\bar{q})^3}{3!} \cdots \frac{(n\bar{q})^K}{K!} \cdots\right].$$

Pairs of chromosomes with K_1 and K_2 lethals, respectively, have K_1K_2 times the chance of allelism of a random pair of lethal genes. The ratio of the probability of allelism for lethal chromosomes to that for lethal genes is thus:

$$\left[\frac{e^{-n\bar{q}}}{1-e^{-n\bar{q}}}\right]^2\left[n\bar{q}+2\frac{(n\bar{q})^2}{2!}+3\frac{(n\bar{q})^3}{3!}\cdots K\frac{(n\bar{q})^K}{K!}\cdots\right]^2 = \left[\frac{n\bar{q}}{1-e^{-n\bar{q}}}\right]^2. \tag{7}$$

In the present case this equals $[0.166/0.153]^2$. The probability of allelism of lethal genes from different localities may thus be taken as $0.00407 \times [0.153/0.166]^2 = 0.00346$.

The variance ($\sigma_{\bar{q}}^2$) of gene frequencies due to real differences in $\bar{q}$ among different lethals should apply equally among genes from the same locality and from different ones. But that due to restriction of size of populations (σ_q^2) would be very small in a sample collected from a larger region. Even if the localities in the present material were fully representative of their entire ranges and the total region were merely the sum of the eleven ranges sampled, σ_q^2 for the total would be only one eleventh of its value for a single range. But since the localities are very much smaller than the ranges and there are more than eleven ranges in the Death Valley region, the interlocality tests really apply to a population many more than eleven times the size of populations of a locality, and in these tests the term corresponding to σ_q^2 may be treated as negligible without important error.

$$n(\sigma_q^2 + \sigma_{\bar{q}}^2 + \bar{q}^2) = 0.0311\ (n\bar{q})^2 = 0.000{,}857 \tag{8}$$

$$n(\sigma_{\bar{q}}^2 + \bar{q}^2) = 0.00346(n\bar{q})^2 = 0.000{,}095 \tag{9}$$

$$n\sigma_q^2 = 0.0276\ (n\bar{q})^2 = 0.000{,}762 \tag{10}$$

If v and s are the same for all loci, $\sigma_{\bar{q}}^2 = 0$, and $n = 289$ (from (9). This gives a minimum estimate of the number of loci in the third chromosome of *Drosophila pseudoobscura* subject to mutations that are lethal. The actual number is presumably considerably larger (see p. 47). Below are three possible sets of values of the constants considered so far (calculated from equations 5, 3, 9 and 10 using three assumed values of n). It may be noted that $\sigma_{\bar{q}}$ is maximum if $n = 578$, accepting the empirical chance of allelism within localities at face value.

n	v	$\bar{q}$	$\sigma_{\bar{q}}$	σ_q
289	10.3×10^{-6}	0.000574	0	0.00162
500	5.9×10^{-6}	0.000332	0.000284	0.00123
1000	3.0×10^{-6}	0.000166	0.000260	0.00087

THE EFFECTS OF SELECTION AND INBREEDING

In any population at equilibrium the rates of the origin of new lethals by mutation and of the elimination of those already present by selection

must be approximately equal. The elimination rate, expected from the chance union of lethal bearing chromosomes in a random breeding population, turns out to be 0.073 percent or 0.0007 (*cf.* p. 26). Yet, the observed mutation rate is 0.00297, or more than four times as great as the computed elimination rate, a point to which DR. A. H. STURTEVANT called our attention. Although both figures are subject to large error, the discrepancy is too striking to be attributed to chance. This is the more so since the elimination and mutation rates computed from the data of DUBININ and collaborators (1936) for Caucasian populations of *Drosophila melanogaster* proved to be even more strikingly unequal (*cf.* DOBZHANSKY 1939b). There are two ways of accounting for this discrepancy. There may be selection against the lethals in heterozygotes, or there may be enough inbreeding within localities to cause more frequent union of like lethals than postulated above, or both. It should be noted that inbreeding due to limitation of the size of population in the localities sampled has no effect here. We are concerned with s and F but not N.

An estimate of the amount of selection (S) against lethal bearing chromosomes in the heterozygous state that is required to balance gains and losses can be arrived at from the primary data (noting that S referring to chromosomes is not quite the same as s referring to genes). Let C and Cl represent normal and lethal bearing chromosomes, respectively, $Q(=.153)$ the frequency of the latter, $P(=0.0311)$ the chance of allelism, $V(=0.00297)$ the rate of origin of lethals, and W the selective value of a zygote. The selective value of C/C being 1, that of the heterozygotes (C/Cl) is $1-S$ and of zygotes resulting from the union of two lethals is $(1-2S)$ in the $(1-P)$ cases in which these are not allelic, 0 in the P cases in which they are allelic.

Zygote	*Frequency*	W	
C/C	$(1-Q)^2$	1	$\overline{W} = 1 - 2SQ - PQ^2 + 2SPQ^2$
C/Cl	$2Q(1-Q)$	$1-S$	$\frac{d\overline{W}}{dQ} = -2S - 2PQ + 4SPQ$
Cl/Cl	Q^2	$(1-P)(1-2S)$	

The rate of change of the frequency of lethal bearing chromosomes is given by the following formula (WRIGHT 1937):

$$\Delta Q = V(1-Q) + \frac{Q(1-Q)}{2W}\frac{d\overline{W}}{dQ}.$$

If $S=0$, the rate of change due to mutation comes out $V(1-Q)=0.00252$, while that due to elimination of homozygotes is $-PQ^2(1-Q)/(1-PQ^2)$

$= -0.00062$, less than one fourth enough to balance the former, as noted in a slightly different form by DR. STURTEVANT.

Putting $\Delta Q = 0$, the solution for S yields the following:

$$S = \frac{V - PQ^2(1 + V)}{Q(1 + 2V) - 2PQ^2(1 + V)} = 0.0147.$$

Thus it does not require much selection against heterozygotes to balance gains and losses. This figure refers to lethal chromosomes. Taking into account our estimate that there may be more than one lethal gene in a chromosome, it must be multiplied by 0.153/0.166 to estimate the average selection against lethal genes, giving $s = 0.0135$.

To investigate the simultaneous effects of inbreeding and selection against heterozygotes, it is necessary to deal with gene frequencies. The following formulae in which l refers to a lethal gene, hold sufficiently accurately if both F and s are small.

Zygote	*Frequency*	W	
+/+	$(1 - q)^2(1 - F) + (1 - q)F$	1	
+/l	$2q(1 - q)(1 - F)$	$1 - s$	(11)
l/l	$q^2(1 - F) + qF$	0	

In the presence of inbreeding, selection pressure is given by the formula

$$\frac{q(1 - q)}{\overline{W}}\left[(1 - F)\frac{d\overline{W}_R}{2dq} + F\frac{d\overline{W}_I}{dq}\right] \qquad (12)$$

where $\overline{W}_R$ is the mean selective value of the random bred component of the array of frequencies (*cf.* (1)) and $\overline{W}_I$ is that of the inbred component (WRIGHT in manuscript).

$$\overline{W}_R = 1 - 2sq(1-q) - q^2, \qquad \frac{d\overline{W}_R}{dq} = -2q - 2s + 4sq$$

$$\overline{W}_I = 1 - q, \qquad \frac{d\overline{W}_I}{dq} = -1$$

$$\overline{W} = 1 - 2sq(1-q)(1-F) - q^2(1-F) - qF$$

$$\Delta q = -\frac{q(1-q)}{\overline{W}}\left[(s + F - Fs + q(1 - 2s)(1-F)\right] + v(1-q) - m(q - \bar{q}). \qquad (13)$$

In a population of limited size it is not to be expected that each gene will always exhibit the frequency at which, in the long run, gains and losses balance (that is, at which $\Delta q = 0$). Its frequency varies at random in the distribution curve ϕ (q) characterized by the constants $\bar{q}$ and σ_q^2. This *distribution*, as exhibited by a large number of genes with the same v and

s, should, however, remain constant, implying among other things, that the values of Δq for all such genes will cancel out. Thus we must write $\int_0^1 \Delta q \phi(q) dq = 0$, in contrast with the simpler equation, $\Delta Q = 0$ which sufficed for lethal bearing chromosomes. In evaluating this equation, it may be noted that

$$\int_0^1 \phi(q)dq = 1, \qquad \int_0^1 q\phi(q)dq = \bar{q}, \qquad \int_0^1 q^2\phi(q)dq = \sigma_q^2 + \bar{q}^2,$$

that $\int_0^1 q^3\phi(q)dq$ is negligibly small in the case of lethals, and that $\overline{W}$ is sufficiently close to 1 to be ignored where it appears in the denominator in the formula for selection pressure. Writing $C_1 = (s + F - Fs)$ and $C_2 = (1 - 2s)(1 - F)$,

$$\int_0^1 \Delta q \phi(q) dq = v - (C_1 + v)\bar{q} + (C_1 - C_2)(\sigma_q^2 + \bar{q}^2) = 0. \qquad (14)$$

If $F = 0$, $C_1 = s$, $C_2 = 1 - 2s$,

$$s = \frac{v(1 - \bar{q}) - (\sigma_q^2 + \bar{q}^2)}{\bar{q} - 3(\sigma_q^2 + \bar{q}^2)}. \qquad (15)$$

While evaluation depends on the number of loci assumed, it varies only slightly with different assumptions. Thus, if $n = 289$, $s = 0.0129$. If $n = 500$, $s = 0.0132$, if $n = 1000$, $s = 0.0133$.

The formula deduced from the balance of gains and losses of lethal bearing chromosomes, agrees with the above except for second order terms, and the estimate of s based on it (.0135) is in good agreement considering that slightly different assumptions and approximations are involved.

If $s = 0$, $C_1 = F$, $C_2 = 1 - F$,

$$F = \frac{v(1 - \bar{q}) - (\sigma_q^2 + \bar{q}^2)}{\bar{q} - 2(\sigma_q^2 + \bar{q}^2)}. \qquad (16)$$

As this differs from the formula for s (15) only in a small term, the values are closely similar. If $n = 289$, $F = 0.0129$, if $n = 500$, $F = 0.0131$, if $n = 1000$, $F = 0.0133$. The coefficient $F = 0.013$ would be accounted for if five percent of the matings are between brother and sister ($F = 0.25$), 95 percent random, or if there is subdivision such that $N_i m_i = 19$.

The primary data give no basis for separating the effects of selection against heterozygotes (s) and of heterogeneity within localities (F). We may write with sufficient accuracy:

$$s + F = 0.013. \qquad (17)$$

An attempt, however, has been made to attack the problem of the incomplete recessivity of lethals (the value of s) experimentally (DOB-

ZHANZKY and QUEAL 1938b, DOBZHANSKY 1939a). Flies carrying one *or Bl Sc pr* and one wild third chromosome derived from natural populations of the Death Valley region were mated. Two classes of such matings were arranged involving chromosomes from the same locality. In some of them, both wild chromosomes were known to be free of lethals, while in the others they carried non-allelic lethals. These crosses may be represented as follows, letting S′ be the selection against +/*or Bl Sc pr* and S that against heterozygotes for the "natural" lethals.

No lethals in wild chromosomes			*Different lethals in wild chromosomes*		
Constitution	*Frequency*	W	*Constitution*	*Frequency*	W
+/+	$\frac{1}{4}$	1	l_1/l_2	$\frac{1}{4}$	$1 - 2S$
+/*Bl Sc*	$\frac{1}{2}$	$1 - S'$	*l*/*Bl Sc*	$\frac{1}{2}$	$1 - S - S'$
Bl Sc/*Bl Sc*	$\frac{1}{4}$	0	*Bl Sc*/*Bl Sc*	$\frac{1}{4}$	0
Proportion of +/+			*Proportion of* l_1/l_2		
$0.3618 \pm 0.0031 = \frac{\frac{1}{4}}{\frac{3}{4} - \frac{2}{4}S'}$			$0.3472 \pm 0.0023 = \frac{\frac{1}{4}(1 - 2S)}{\frac{3}{4} - S - \frac{2}{4}S'}$		

Solving these as simultaneous equations, $S' = 0.118$, $S = 0.066$. Taken at face value, there is a clearly significant difference between the experiments, and the selection coefficient (S) relating to heterozygous lethals is several times as great as the value required by the primary data, assuming the same coefficient for all loci and no heterogeneity within localities. The authors note, however, that the experiments were not conducted simultaneously and that the observed difference may be environmental. The experiment, accordingly, was repeated on a larger scale with all possible precautions, this time using Mexican and Guatemalan materials (DOBZHANSKY 1939a). The proportion of the wild type flies in the experiment without lethals was 33.59 ± 0.22 percent and in that involving non-allelic lethals 33.11 ± 0.27 percent. The difference is not significant, and the productivity per bottle also did not differ significantly (160.25 ± 4.22 and 163.11 ± 3.50 flies, respectively). However, if the figures are taken at face value, there is a selection of 2.1 percent against each lethal bearing chromosome (1.2 percent against *Bl Sc*). Even this figure is more than required to account for the apparent discrepancy between the mutation and elimination rates of lethals. These data accordingly do not advance us in interpreting the indeterminate equation $s + F = 0.013$. The situation illustrates the extreme difficulty of a direct attack on questions of selection.

It is also possible that while $s = 0$ for practically all the lethals found in nature, it may be very high in the great majority of the lethal mutations. The only data bearing on this question are the rates of mutation observed in the direct and the accumulation experiments. As already noted, the

percentage was actually higher (0.00313 ± 0.00038) in the accumulation experiment than in the direct experiment (0.00251 ± 0.00056), although not significantly so. As far as it goes, the evidence is thus against the hypothesis of a rapid elimination of some three-fourths of the lethal mutations.

One of the possible approaches to the problem of inbreeding within localities and neighborhoods as a cause of the excess elimination of lethals is through studies on the variations in the gene arrangement in the third chromosome. As shown by DOBZHANSKY and QUEAL (1938a), KOLLER (1939) and DOBZHANSKY (unpublished), six different gene arrangements, of which three are common, occur in the natural populations of the Death Valley region. Some data are available on the differentiation of populations in stations, localities, and ranges, as well as on changes in the composition of populations in time. An analysis of these and related data is reserved for a future publication, but two points may be mentioned here for the sake of completeness. The first is that the estimates of the value of F derived from these data are of an order of magnitude compatible with the equation (17). Second, a possible effect of a selection against homozygotes for the gene arrangements (*cf.* STURTEVANT and MATHER 1938) on the distribution and elimination of lethals has been considered and found to be slight.

EFFECTIVE SIZE OF POPULATION AND AMOUNT OF CROSSBREEDING

The greater chance of finding an allele of a given lethal within the same locality than elsewhere demonstrates that the population number of localities is not indefinitely large. It should be possible to determine something about the breeding structure from the estimate of the chance variability of gene frequencies (σ_q^2).

The general formula for the distribution of gene frequencies is as follows where Δq is the rate of change of gene frequency due to systematic pressures (mutation, selection, immigration) and $\sigma^2_{\Delta q}$ is the sampling variance ($\sigma^2_{\Delta q} = q(1-q)/2N$) in a random breeding diploid population (WRIGHT 1938a).

$$\phi(q) = (C/\sigma^2_{\Delta q})e^{-2\int(\Delta q/\sigma^2_{\Delta q})\,dq}. \qquad (18)$$

In the present case this reduces to the following expression (attributing the apparent lack of balance of gains and losses of lethals to imperfect dominance rather than heterogeneity):

$$\phi(q) = C[1 - 2sq - (1-2s)q^2]^{2N}q^{4N[m\bar{q}+v]-1}(1-q)^{4Nm(1-\bar{q})-1}. \qquad (19)$$

For sufficiently small values of N, the selection term differs little from 1 and thus has little effect on the form of the distribution in spite of the

great importance of selection in determining the low value of $\bar{q}$ in the species as a whole. Omitting the selection term,

$$\phi(q) = \frac{\Gamma[4N(m+v)]}{\Gamma[4N(m\bar{q}+v)]\Gamma[4Nm(1-\bar{q})]} q^{4N(m\bar{q}+v)-1}(1-q)^{4Nm(1-q)-1}. \quad (20)$$

This distribution has the mean $\bar{q}$ and variance $\sigma_q^2 = \bar{q}(1-\bar{q})/[4N(m+v)+1]$

$$N(m+v) = \frac{1}{4}\left[\frac{\bar{q}(1-\bar{q})}{\sigma_q^2} - 1\right]. \quad (21)$$

Ignoring v, negligibly small in comparison with m in any case in which the latter is of interest, the result is practically independent of the estimate of number of loci. Substituting the values of $n\bar{q}$ from (3), of $n\sigma_q^2$ from (10) and treating $1-\bar{q}$ as 1 (maximum value of $\bar{q}$ is 0.000574),

$$Nm = 54. \quad (22)$$

As the effect of selection is to reduce the variance of gene frequencies, it would require a smaller value of N for a given m to give the observed variance. Thus 54 is too large an estimate even when N is small and may be very much too large if N is large. From consideration of the variability of localities in another respect (inversions in the third chromosome, to be discussed in another paper), Nm cannot be less than 5.

The frequencies of gene frequencies $f(q) = \phi(q)\delta q$ at intervals $\delta q = 1/2N$ (or more at higher values of q) have been calculated for various values of n, N, and m, assuming $v = 0.00297/n$, $\bar{q} = 0.166/n$, $s = 0.013$ in order to find sets of values which satisfy approximately the calculated variance ($\sigma_q^2 = 0.000762/n$), as well as the mean. The frequency f(0) with which lethals are absent from a locality can not however be calculated from the formula of the curve. The approximate formula $f(0) = f(1/2N)/4N(m\bar{q}+v)$ was used (WRIGHT 1937). Calculations have not been made for values of m larger than 0.05 since the derivation of the formula for the curve requires that $(\Delta q)^2$ be negligible and the above formula for f(0) becomes inaccurate for larger values.

The results indicate that the effective size of population of a locality could not have been larger than 2500, (assuming that $Nm = 5$ which, as shown by the preliminary results of the analysis of the data on the variation in gene arrangement, is probably too low). If $N = 1000$, m must be about 0.04 (Nm about 40) in order to satisfy the conditions, largely irrespective of the estimate of number of loci. It is probable that $N = 500$, m about 0.08 would satisfy them equally well.

So far we have assumed a uniform value of s (or of F). It is desirable to consider the effect of extreme differences in the amount of selection against heterozygotes. Assume that some of the lethal mutations observed in the

laboratory have no effect whatever when heterozygous, while others have such severe effects that under natural conditions they leave few offspring. Obviously this hypothesis is not in accord with the data (see above) but it is of interest as an extreme case. The population constants are as follows:

Breeding structure N, $F=0$, m

Completely recessive lethals n_1, v_1, $s_1=0$, $\bar{q}_1$, $\sigma_{q_1}^2$

Heterozygotes non-productive in nature n_2, v_2, s_2, $\bar{q}_2$, $\sigma_{q_2}^2$.

Under the postulated conditions, there would be very few mutations of the second class in nature, and $\bar{q}_2^2$ and $\sigma_{q_2}^2$ would be negligibly small. The following equations can be written (*cf.* (5) (3), (8), (9) and (14)).

$$n_1v_1 + n_2v_2 = 0.00297 \tag{23}$$

$$n_1\bar{q}_1 + n_2\bar{q}_2 = 0.166 \tag{24}$$

$$n_1(\sigma_{q_1}^2 + \bar{q}_1^2) + n_2(\sigma_{q_2}^2 + \bar{q}_2^2) = 0.0311 \times 0.166^2 \tag{25}$$

$$n_1\bar{q}_1^2 + n_2\bar{q}_2^2 = 0.00346 \times 0.166^2 \tag{26}$$

$$\int_0^1 \Delta q_1\phi(q_1)dq_1 = v_1(1-\bar{q}_1) - (\sigma_{q_1}^2 + \bar{q}_1^2) = 0,$$

$$v_1 = \sigma_{q_1}^2 + \bar{q}_1^2 \text{ approx.} \tag{27}$$

$$\int_0^1 \Delta q_2\phi(q_2)dq_2 = v_2(1-\bar{q}_2) - s_2\bar{q}_2 - (1-3s_2)(\sigma_{q_2}^2 + \bar{q}_2^2) = 0,$$

$$s_2 = v_2/\bar{q}_2 \text{ approx.} \tag{28}$$

It makes little difference what value is assigned s_2, provided it is large (whether $s_2=0.25$, $\bar{q}_2=4v_2$; $s_2=.50$, $\bar{q}_2=2v_2$, or $s_2=1$, dominant lethal in nature $\bar{q}_2=v_2$, there is no appreciable accumulation in nature). As the limiting case we shall assume $s_2=1$. We shall also assume that $v_2=v_1$.

The values of all the constants except N and m can now be found by solving the equations.

$$n_1 = 282 \qquad v_1 = v_2 = 0.000,003,03$$

$$n_2 = 698 \qquad \bar{q}_1 = 0.000,581$$

$$n = 980 \qquad \sigma_{q_1}^2 = 0.000,002,69$$

The frequency distribution of the complete recessives is:

$$\phi(q_1) = C(1-q_1^2)^{2N}q_1^{4N(m\bar{q}_1+v)-1}(1-q_1)^{4Nm(1-\bar{q}_1)-1}. \tag{29}$$

To find the maximum possible value of N for a locality put $m=0$. The mean and variance can readily be obtained (WRIGHT 1937).

$$\bar{q}_1 = \frac{\Gamma(2Nv+\frac{1}{2})}{\sqrt{2N}\,\Gamma(2Nv)} \text{ approximately} \tag{30}$$

$$\sigma_{q_1}^2 = v - \bar{q}_1^2 \text{ approximately.} \tag{31}$$

By trial it is found that N is close to 6500. This must be far above the actual effective size of populations of localities, since these are certainly not completely isolated and the extreme assumptions made here ($F=0$ and $s_1=0$ for the lethals in nature) increase the estimate.

If there were complete panmixia within completely isolated ranges and the flies collected from one locality could be considered as a random sample from the range, the estimate 6500 could be applied to the whole range. But the conditions are not sufficiently realized to give this interpretation any value. If, as is undoubtedly the case, there is considerable exchange of population between localities of the same range, N for the locality must be much less than 6500, even assuming $F=0$, $s_1=0$. The term $(1-q_1^2)^{2N}$ approaches 1, and as before, the variance of the distribution indicates $Nm=54$. The approach is closer than with $F+s=0.013$ assumed before.

DISCUSSION

In the foregoing analysis, we have recognized nine major population constants and have attempted evaluation as far as possible from four primary data and two equations connecting $\bar{q}$ and σ_q^2 with the others. This left three degrees of indeterminacy. The solution of the six equations can be expressed as follows:

$$nv = 0.00297 \tag{32}$$

$$n\bar{q} = 0.166 \tag{33}$$

$$1/n + 36.3n\sigma_{\bar{q}}^2 = 0.00346 \tag{34}$$

$$n\sigma_q^2 = 0.000,762 \tag{35}$$

$$s + F = 0.013 \text{ (if } \sigma_s^2 = 0) \tag{36}$$

$$\left.\begin{aligned} Nm &= 40 \pm \text{ (if N lies between 80 and 1000)} \\ Nm &= 15 \pm \text{ (if N is 2000)} \\ Nm &= 5 \pm \text{ (if N is 2500)} \end{aligned}\right\} \tag{37}$$

A minimum estimate of n, number of loci in the third chromosome giving rise to lethals by mutation, was found to be 289 from (34) putting $\sigma_{\bar{q}}^2=0$. Unfortunately, this estimate is predicated on certain assumptions which lack experimental verification, namely that all the loci producing lethals have the same mutation rates (v) and that the lethals have no deleterious effects in heterozygotes. If the mutation rates are variable, the figure 289 is an underestimate. Obviously there could be an indefinitely large number of loci, of which all but a small percentage have such exceedingly low mutation rates that the respective mutants are unlikely to be found either in mutation experiments or in nature. Not only dominant lethals, but also those producing appreciable deleterious effects in heterozygotes are eliminated probably very rapidly outdoors; the lethals that are retained in

natural populations must be very nearly, or completely, recessive. The possibility remains, however, that some of the genes in the third chromosome produce only dominant or semi-dominant lethal changes. Such genes are evidently not included in the estimate.

A lethal may represent a change within a single locus, a gene, or it may be structural change in the chromosome, such as a deficiency for one or several adjacent genes. Indeed, SLIZYNSKI (1938) and others have shown that, at least among the X-ray induced lethals in the X chromosome of *Drosophila melanogaster*, many are demonstrably deficiencies including usually a single or a few discs in the salivary gland chromosome. Among the spontaneous lethals the frequency of deficiencies is apparently lower than among the induced ones. Although no cytological study of the lethals in *Drosophila pseudoobscura* has been attempted, the situation in this species is presumably similar to that in *melanogaster*. Since several non-identical but overlapping deficiencies may behave as alleles, the allelism of lethals is not a proof of their identity.

Assuming that each lethal is due to a change in or to a deficiency for a single gene, and that there is a one-to-one relation between genes and discs in the salivary gland chromosomes, a cytological estimate of the number of genes in a chromosome may be attempted. The map of the third chromosome of *Drosophila pseudoobscura* published by DOBZHANSKY and STURTEVANT (1938) shows approximately 554 discs, but the authors state explicitly that some of the discs have probably been missed. The right limb of the second chromosome of *Drosophila melanogaster*, which is supposed to be homologous to the third chromosome of *pseudoobscura*, has 1136 discs (according to the best existing map published by BRIDGES and BRIDGES, 1939), and hence a maximum estimate of n would be of this order. Assignment of the minimum value, $n = 289$, gives a value for average mutation rate per locus, v, of about 10^{-5}. If the maximum value for n is accepted, v lies between 10^{-5} and 10^{-6}, which can hardly be considered excessively low for average mutability (TIMOFEEFF-RESSOVSKY 1937).

The estimate of the population number, N, depends largely on the observed difference in the chance of allelism of lethals found in the same and in different localities. As already noted, there is nothing in the data which forbids the hypothesis that the chance of allelism for chromosomes from different traps of the same locality is the same as that between localities. In this limiting case $\sigma_q^2 = 0$. The apparent discrepancy between mutation and elimination rates becomes nine times as great as before, yet it requires only that $s+F = 0.018$ to balance it. In this case only the first three of the six equations (32–37) still stand, and no upper limit can be set to the value of Nm, except from consideration of the relation between the number of flies tested by one trap and the number in the whole "local-

ity," not given by the numerical data at hand. It must be reiterated, however, that this difficulty is removed by newer observations, to be published elsewhere, which show that the chances of allelism of lethals collected in the same station simultaneously and at different times are practically identical.

In conclusion it may be stated that we entertain no illusions regarding the lack of precision in the results obtained. At best it is hoped that the order of magnitude of certain variables effective in population dynamics has been arrived at. But at a certain stage of the development of a scientific subject even such rough estimates may be useful to guide further work. Although population dynamics has been discussed for decades, mainly in connection with evolutionary speculations, few attempts have ever been made to put the discussion on a basis of experimentally determined quantities.

SUMMARY

1. Samples of the population of race A of *Drosophila pseudoobscura* were made in eleven localities, each in a separate mountain forest, in the Death Valley region. Tests of 857 wild third chromosomes showed that 15.29 $\pm$0.83 percent of them carried one or more lethals or semilethals. This is significantly less than found in similar samples from Mexico (28.1 $\pm$3.3 percent) or from Guatemala (34.2 $\pm$5.2 percent). Yet the rates of occurrence of lethal mutations in stocks from these regions were substantially the same (Death Valley 0.00297 $\pm$0.00032, Mexico 0.00359 $\pm$0.00062, Guatemala 0.00284 $\pm$0.00061). In tests of pairs of lethals found in the same locality in Death Valley, 3.11 $\pm$0.42 percent were found to be allelic. On the other hand, only 0.407 $\pm$0.061 percent were found to be allelic in tests of all pairs of lethals from different localities.

2. In attempting to interpret these figures, it is pointed out that account must be taken of at least nine properties of the populations and of the loci: the effective population number N, the inbreeding coefficient F, the immigration index m, the number of loci in the third chromosome subject to lethal mutation n, the rate of mutation per locus per generation v, the possible selective disadvantage of heterozygotes s, the mean $\bar{q}$, and the variance $\sigma_{\bar{q}}^2$, expected among genes with a given v and s under a given breeding structure (N, F, m), and finally an additional contribution to variance σ_q^2, expected among genes which differ in the values of v and s.

3. A determinative solution for these nine constants is obviously impossible on the basis of four primary data and two relations of dependence ($\bar{q}$, σ^2_q). Evaluations of v, $\bar{q}$, σ_q^2 and $\sigma_{\bar{q}}^2$, however, depend merely on evaluation of n, according to the following four indeterminate equations:

(a) $nv = 0.00297$ (c) $n\sigma_q^2 = 0.000762$

(b) $n\bar{q} = 0.166$ (d) $1/n + 36.3n\sigma_{\bar{q}}^2 = 0.00346.$

The last of these gives $n = 289$ as a minimum estimate of the number of mutable loci (assuming $\sigma_{\bar{q}}^2 = 0$). Consideration of the number of discs in the chromosome as seen in salivary gland cells suggests $n = 1000$ as a rough maximum estimate. The probable range of values of v, $\bar{q}$, σ_q^2, and $\sigma_{\bar{q}}^2$ can be deduced at once from this probable range of values of n.

4. The observed mutation rate is more than four times as great as necessary to account for the elimination of lethals on the assumption that these are completely recessive ($s = 0$) and that there is random mating within localities ($F = 0$). The possibility that about three fourths of the lethal mutations are lost at once in nature has been shown to be remote.

5. If s is assumed to be the same for all loci, the balance of gains and losses of lethals in nature can be accounted for according to the indeterminate equation (e), $s + F = 0.013$. The data are as yet insufficient to discriminate between the roles of s and F in this equation.

6. The difference between the chances of allelism within and between localities can be accounted for by an indeterminate relation between N and m:

(f)
$$Nm = 40 \pm \text{ if N lies between 80 and 1000}$$
$$Nm = 15 \pm \text{ if } N = 2000$$
$$Nm = 5 \pm \text{ if } N = 2500.$$

The variability in the frequencies of chromosome aberrations (not discussed in detail in the present article) indicates that Nm can hardly be less than five. It is concluded that the effective size of population in localities (collections from 10 to 20 traps along $\frac{1}{4}$ to 1 mile) is less than 2500. A more exact determination of this value must await further data.

LITERATURE CITED

BRIDGES, C. B., and BRIDGES, P. N., 1939 A new map of the second chromosome. J. Hered. **30**: 475–476.

DEMEREC, M., 1937 Frequency of spontaneous mutation in certain stocks of *Drosophila melanogaster*. Genetics **22**: 469–478.

DOBZHANSKY, TH., 1939a Genetics of natural populations. IV. Mexican and Guatemalan populations of *Drosophila pseudoobscura*. Genetics **24**: 391–412.

1939b Experimental studies on genetics of free-living populations of Drosophila. Biol. Rev. **14**: 339–368.

DOBZHANSKY, TH., and QUEAL, M. L., 1938a Genetics of natural populations. I. Chromosome variations in populations of *Drosophila pseudoobscura* inhabiting isolated mountain ranges. Genetics **23**: 239–251.

1938b Genetics of natural populations. II. Genic variations in populations of *Drosophila pseudoobscura* inhabiting isolated mountain ranges. Genetics **23**: 463–484.

DOBZHANSKY, TH., and STURTEVANT, A. H., 1938 Inversions in the chromosomes of *Drosophila pseudoobscura*. Genetics **23**: 28–64.

DUBININ, N. P., HEPTNER, M. A., DEMIDOVA, Z. A., and DJACHKOVA, L. I., 1936 Genetic constitution and gene dynamics of wild populations of *Drosophila melanogaster*. Biol. Z. Moskau **5**: 939–976.

KOLLER, P. CH., 1939 Genetics of natural populations. III. Gene arrangements in populations of *Drosophila pseudoobscura* from contiguous localities. Genetics **24:** 22–33.

RHOADES, M. M., 1938 Effect of the *Dt* gene on the mutability of the *a1* allele in maize. Genetics **23:** 377–397.

SLIZYNSKI, B. M., 1938 Salivary chromosome studies of lethals in *Drosophila melanogaster*. Genetics **23:** 283–290.

STURTEVANT, A. H., 1937 Autosomal lethals in wild populations of *Drosophila pseudoobscura*. Biol. Bull. **73:** 542–551.

1939 High mutation frequency induced by hybridization. Proc. Nat. Acad. Sci. **27:** 308–310.

STURTEVANT, A. H., and MATHER, K., 1938 The interrelations of inversions, heterosis and recombination. Amer. Nat. **72:** 447–452.

TIMOFEEFF-RESSOVSKY, N. W., 1937 Experimentelle Mutationsforschung in der Vererbungslehre. Dresden: Theodor Steinkopf.

TIMOFEEFF-RESSOVSKY, N. W., and ZIMMER, K. G., 1939 Mutationsauslösung durch Röntgenbestrahlung unter verschiedener Temperatur bei *Drosophila melanogaster*. Biol. Zbl. **59:** 358–362.

WAHLUND, S., 1928 Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas **11:** 65–106.

WRIGHT, S., 1921 Systems of mating. Genetics **6:** 11–178.

1922 The effects of inbreeding and crossbreeding on guinea pigs. III. Crosses between highly inbred lines. Bull. U. S. Dept. Agric. No. 1121, 55 pp.

1931 Evolution in Mendelian populations. Genetics **16:** 97–159.

1937 The distribution of gene frequencies in populations. Proc. Nat. Acad. Sci. **23:** 307–320.

1938a The distribution of gene frequencies under irreversible mutation. Proc. Nat. Acad. Sci. **24:** 253–259.

1938b Size of population and breeding structure in relation to evolution. Science **87:** 430–431.

1940 Breeding structure of populations in relation to speciation. Amer. Nat. **74:** 232–248.

ZUITIN, A. I., 1938a The influence of the change of the thermal regime upon the frequency of occurrence of lethal mutations in *Drosophila melanogaster*. C. R. (Doklady) Acad. Sci. U.R.S.S. **21:** 53–55.

1938b The combined effects of change of the thermal regime and the subsequent temperature contrast upon the frequency of lethal mutations in *Drosophila melanogaster*. C. R. (Doklady) Acad. Sci. U.R.S.S. **21:** 56–58.